Discontinuous Galerkin Methods[1]

Daniele Di Pietro[a] and Alexandre Ern[b]

[a] IFP Energies Nouvelles, 1 & 4 avenue du bois Préau
92852 Rueil-Malmaison Cedex, France
dipietrd@ifpen.fr

Université Paris-Est, CERMICS, Ecole des Ponts
6 et 8 avenue Blaise Pascal, 77455 Marne-la-Vallée cedex 2, France
ern@cermics.enpc.fr

1 Introduction

Discontinuous Galerkin (dG) methods can be viewed as finite element methods allowing for discontinuities in the discrete trial and test spaces. Localizing test functions to single mesh elements and introducing numerical fluxes at interfaces, they can also be viewed as finite volume methods in which the approximate solution is represented on each mesh element by a polynomial function and not only by a constant function. From a practical viewpoint, working with discontinuous discrete spaces leads to compact discretization stencils and, at the same time, offers a substantial amount of flexibility, making the approach appealing for multi-domain and multi-physics simulations. Moreover, basic conservation principles can be incorporated into the method. Applications of dG methods cover a vast realm in engineering sciences. Examples can be found, e.g., in the conference proceedings edited by Cockburn, Karniadakis, and Shu [39]. There is also an increasing number of open source libraries implementing dG methods. A non exhaustive list includes `deal.II` [9], `Dune` [12], `FEniCS` [71], `freeFEM` [45], `libmesh` [63], and `Life` [74].

A brief historical perspective

Although dG methods have existed in various forms for more than thirty years, they have experienced a vigorous development only over the last decade, as illustrated in Figure 1.

The first dG method to approximate first-order PDEs has been introduced by Reed and Hill in 1973 [75] in the context of steady neutron transport, while the first analysis for steady first-order PDEs was performed by Lesaint and Raviart in 1974 [65, 66, 67]. The error estimate was improved by Johnson and Pitkäranta in 1986 [61] who established an order of convergence in the L^2-norm of $(k + \frac{1}{2})$ if polynomials of degree k are used and the exact solution is smooth enough. A few years later, dG methods were extended to time-dependent hyperbolic PDEs by Chavent and Cockburn [28] using the forward Euler scheme for time discretization together with limiters. The order of accuracy was improved by Cockburn and Shu [41, 42] using explicit Runge–Kutta schemes for time discretization, while a convergence proof to the entropy solution was obtained by Jaffré, Johnson, and Szepessy [60]. Extensions are discussed in a series of papers by Cockburn, Shu, and coworkers; see, e.g., [33, 40, 44].

For PDEs with diffusion, dG methods originated from the work of Nitsche on boundary-penalty methods in the early seventies [69, 70] and the use of Interior Penalty (IP) techniques to weakly enforce continuity conditions imposed on the solution or its derivatives across interfaces,

[1]Lecture Notes for the Spring School on Numerical Fluid Mechanics, Roscoff June 2011. A more elaborate and thorough presentation of the material can be found in the book *Mathematical Aspects of Discontinuous Galerkin Methods* by D. Di Pietro and A. Ern, volume 69 of SMAI Mathématiques & Applications, Springer, 2012.

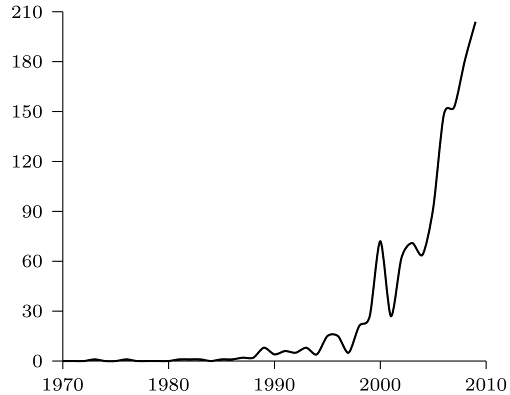

Figure 1: Yearly number of entries with the keyword ‘discontinuous Galerkin’ in the MathSciNet database

as in the work of Babuška [6], Babuška and Zlámal [7], Douglas and Dupont [47], Baker [8], Wheeler [82], and Arnold [2]. In the late nineties, following an approach more closely related to hyperbolic problems, dG methods were formulated using numerical fluxes by considering the mixed formulation of the diffusion term. Examples include the work of Bassi and Rebay [10] on the compressible Navier–Stokes equations and that of Cockburn and Shu [43] on convection-diffusion systems, leading to a new thrust in the development of dG methods. A unified analysis of dG methods for the Poisson problem can be found in the work of Arnold, Brezzi, Cockburn, and Marini [3], while a unified analysis encompassing both elliptic and hyperbolic PDEs in the framework of Friedrichs' systems has been derived by Ern and Guermond [50, 51, 52].

Overview

Section 2 introduces the basic concepts to formulate and analyze dG methods, namely (i) the basic ingredients related to meshes and polynomials to build discrete functional spaces and, in particular, broken polynomial spaces, (ii) the three key properties for the convergence analysis of dG methods in the context of nonconforming finite elements, namely discrete stability, consistency, and boundedness, (iii) the basic analysis tools, in particular inverse and trace inequalities needed to assert discrete stability and boundedness, together with optimal polynomial approximation results, thereby leading to the concept of admissible mesh sequences. We focus on mesh refinement as the main parameter to achieve convergence. Convergence analysis using, e.g., high-degree polynomials is possible; important tools in this direction can be found, in the context of dG methods, in the recent textbook of Hesthaven and Warburton [59].

Section 3 focuses on the steady advection-reaction equation as a simple first-order model problem. Therein, we indetify some key ideas to design dG methods. Two methods are analyzed, which correspond in the finite volume terminology to the use of centered and upwind fluxes.

Section 4 is concerned with the Poisson problem as the basic model problem with diffusion. We first present a heuristic derivation and a convergence analysis to smooth solutions using the Symmetric Interior Penalty (SIP) dG method of Arnold [2]. Then, we introduce the concept of discrete gradients and present some important applications, including the link with the mixed dG approach and the local formulation of the discrete problem using numerical fluxes.

Section 5 is devoted to incompressible flows. Focusing first on the steady Stokes equations, we examine how the divergence-free constraint on the velocity field can be tackled using dG methods. We detail the analysis of equal-order approximations using both discontinuous velocities and pressures, whereby pressure jumps need to be penalized, and then briefly discuss alternative formulations avoiding the need for pressure jump penalty. The next step is the discretization of the nonlinear convection term in the momentum equation. To this purpose, we derive a discrete trilinear form that leads to the correct kinetic energy balance, using the so-called Temam's device to handle the fact that discrete velocities are only weakly divergence-free.

2 Basic concepts

This section introduces the basic concepts to build discontinuous Galerkin (dG) methods.

2.1 The domain Ω

To simplify the presentation, we focus, throughout this lectures notes, on polyhedra.

Definition 2.1 (Polyhedron in $\mathbb{R}^d$). *We say that the set P is a* polyhedron *in $\mathbb{R}^d$ if P is an open, connected, bounded subset of $\mathbb{R}^d$ such that its boundary ∂P is a finite union of parts of hyperplanes, say $\{H_i\}_{1\le i\le n_\Omega}$. Moreover, for all $1 \le i \le n_\Omega$, at each point in the interior of $\partial P \cap H_i$, the set P is assumed to lie on only one side of its boundary.*

Assumption 2.2 (Domain Ω). *The domain Ω is a polyhedron in $\mathbb{R}^d$. The* boundary *of Ω is denoted by $\partial\Omega$ and its (unit)* outward normal*, which is defined a.e. on $\partial\Omega$, by* n.

The advantage of Assumption 2.2 is that polyhedra can be exactly covered by a mesh consisting of polyhedral elements. PDEs posed over domains with curved boundary can also be approximated by dG methods using, e.g., isoparametric finite elements to build the mesh near curved boundaries as described, e.g., by Ciarlet [29, p. 224] and Brenner and Scott [15, p. 117].

2.2 Meshes

The first step is to discretize the domain Ω using a mesh. Various types of meshes can be considered. We examine first the most familiar case, that of simplicial meshes. Such meshes should be familiar to the reader since they are one of the key ingredients to build continuous finite element spaces.

Definition 2.3 (Simplex). *Given a family $\{a_0,\dots,a_d\}$ of $(d+1)$ points in $\mathbb{R}^d$ such that the vectors $\{a_1-a_0,\dots,a_d-a_0\}$ are linearly independent, the interior of the convex hull of $\{a_0,\dots,a_d\}$ is called a non-degenerate* simplex *of $\mathbb{R}^d$, and the points $\{a_0,\dots,a_d\}$ are called its* vertices.

By its definition, a non-degenerate simplex is an open subset of $\mathbb{R}^d$. In dimension 1, a non-degenerate simplex is an interval, in dimension 2 a triangle, and in dimension 3 a tetrahedron. The unit simplex of $\mathbb{R}^d$ is the set

$$S_d := \left\{(x_1,\dots,x_d)\in\mathbb{R}^d;\ \forall i\in\{1,\dots,d\},\ x_i>0;\ x_1+\ldots+x_d<1\right\}.$$

Any non-degenerate simplex of $\mathbb{R}^d$ is the image of the unit simplex by a bijective affine transformation of $\mathbb{R}^d$.

Definition 2.4 (Simplex faces). *Let S be a non-degenerate simplex with vertices $\{a_0,\dots,a_d\}$. For each $i\in\{0,\dots,d\}$, the convex hull of $\{a_0,\dots,a_d\}\setminus\{a_i\}$ is called a* face *of the simplex S.*

Thus, a non-degenerate simplex has $(d+1)$ faces, and, by construction, a simplex face is a closed subset of $\mathbb{R}^d$. A simplex face has zero d-dimensional Hausdorff measure, but positive $(d-1)$-dimensional Hausdorff measure. In dimension 2, a simplex face is also called an edge, while in dimension 1, a simplex face is a point and its 0-dimensional Hausdorff measure is conventionally set to 1.

Definition 2.5 (Simplicial mesh). *A* simplicial mesh *$\mathcal{T}$ of the domain Ω is a finite collection of disjoint non-degenerate simplices $\mathcal{T}=\{T\}$ forming a partition of Ω,*

$$\overline{\Omega} = \bigcup_{T\in\mathcal{T}} \overline{T}. \tag{1}$$

Each $T\in\mathcal{T}$ is called a mesh element.

While simplicial meshes are quite convenient in the context of continuous finite elements, dG methods more easily accommodate general meshes.

Definition 2.6 (General mesh). *A* general mesh $\mathcal{T}$ *of the domain* Ω *is a finite collection of disjoint polyhedra* $\mathcal{T} = \{T\}$ *forming a partition of* Ω *as in* (1). *Each* $T \in \mathcal{T}$ *is called a* mesh element.

Obviously, a simplicial mesh is just a particular case of a general mesh.

Definition 2.7 (Element diameter, meshsize). *Let* $\mathcal{T}$ *be a (general) mesh of the domain* Ω. *For all* $T \in \mathcal{T}$, h_T *denotes the* diameter *of* T, *and the* meshsize *is defined as the real number*

$$h := \max_{T \in \mathcal{T}} h_T.$$

We use the notation $\mathcal{T}_h$ *for a mesh* $\mathcal{T}$ *with meshsize* h.

Definition 2.8 (Element outward normal). *Let* $\mathcal{T}_h$ *be a mesh of the domain* Ω *and let* $T \in \mathcal{T}_h$. *We define* n_T *a.e. on* ∂T *as the (unit)* outward normal *to* T.

Faces of a *single* polyhedral mesh element can be defined. Such faces are not needed in what follows, and we prefer to leave them undefined to avoid confusion with the important concept of mesh faces introduced in §2.3. (Mesh faces depend on the way neighboring mesh elements come into contact.)

2.3 Mesh faces, averages, and jumps

The concepts of mesh faces, averages, and jumps play a central role in the design and analysis of dG methods.

Definition 2.9 (Mesh faces). *Let* $\mathcal{T}_h$ *be a mesh of the domain* Ω. *We say that a (closed) subset* F *of* $\overline{\Omega}$ *is a* mesh face *if* F *has positive* $(d-1)$*-dimensional Hausdorff measure (in dimension 1, this means that* F *is nonempty) and if either one of the two following conditions is satisfied:*

(i) *There are distinct mesh elements* T_1 *and* T_2 *such that* $F = \partial T_1 \cap \partial T_2$*; in such a case,* F *is called an* interface.

(ii) *There is* $T \in \mathcal{T}_h$ *such that* $F = \partial T \cap \partial \Omega$*; in such a case,* F *is called a* boundary face.

Interfaces are collected in the set $\mathcal{F}_h^i$, *and boundary faces are collected in the set* $\mathcal{F}_h^b$. *Henceforth, we set*

$$\mathcal{F}_h := \mathcal{F}_h^i \cup \mathcal{F}_h^b.$$

Moreover, for any mesh element $T \in \mathcal{T}_h$, *the set*

$$\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$$

collects the mesh faces composing the boundary of T. *The maximum number of mesh faces composing the boundary of mesh elements is denoted by*

$$N_\partial := \max_{T \in \mathcal{T}_h} \mathrm{card}(\mathcal{F}_T). \tag{2}$$

Finally, for any mesh face $F \in \mathcal{F}_h$, *we define the set*

$$\mathcal{T}_F := \{T \in \mathcal{T}_h \mid F \subset \partial T\}, \tag{3}$$

and observe that $\mathcal{T}_F$ *consists of two mesh elements if* $F \in \mathcal{F}_h^i$ *and of one mesh element if* $F \in \mathcal{F}_h^b$.

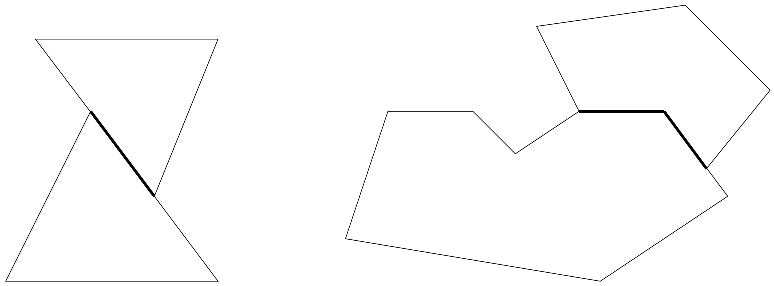

Figure 2: Examples of interface for a simplicial mesh (left) and a general mesh (right)

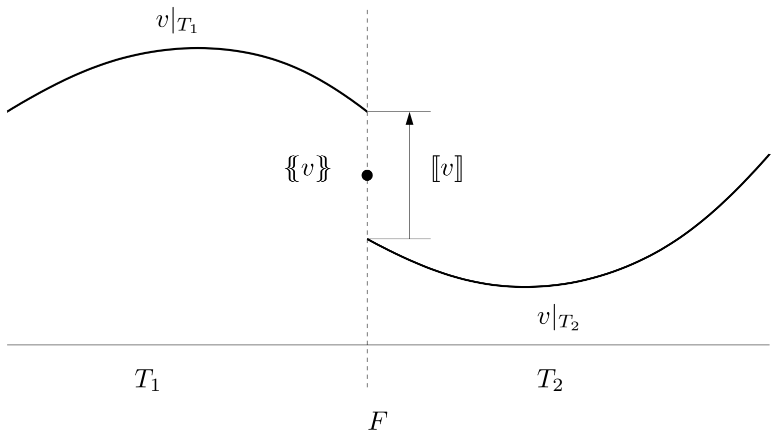

Figure 3: One-dimensional example of average and jump operators; the face reduces to a point separating two adjacent intervals

Figure 2 depicts an interface between two mesh elements belonging to a simplicial mesh (left) or to a general mesh (right). We observe that in the case of simplicial meshes, interfaces are always parts of hyperplanes, but this is not necessarily the case for general meshes containing nonconvex polyhedra. We now define averages and jumps across interfaces of piecewise smooth functions; cf. Figure 3 for a one-dimensional illustration.

Definition 2.10 (Interface averages and jumps). *Let v be a scalar-valued function defined on Ω and assume that v is smooth enough to admit on all $F \in \mathcal{F}_h^i$ a possibly two-valued trace. This means that, for all $T \in \mathcal{T}_h$, the restriction $v|_T$ of v to the open set T can be defined up to the boundary ∂T. Then, for all $F \in \mathcal{F}_h^i$ and a.e. $x \in F$, the* average *of v is defined as*

$$\{\!\{v\}\!\}_F(x) := \frac{1}{2}\Big(v|_{T_1}(x) + v|_{T_2}(x)\Big),$$

and the jump *of v as*

$$[\![v]\!]_F(x) := v|_{T_1}(x) - v|_{T_2}(x).$$

When v is vector-valued, the above average and jump operators act componentwise on the function v. Whenever no confusion can arise, the subscript F and the variable x are omitted, and we simply write $\{\!\{v\}\!\}$ and $[\![v]\!]$.

Definition 2.11 (Face normals). *For all $F \in \mathcal{F}_h$ and a.e. $x \in F$, we define the (unit)* normal *n_F to F at x as*

(i) n_{T_1}, *the unit normal to F at x pointing from T_1 to T_2 if $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$; the orientation of n_F is arbitrary depending on the choice of T_1 and T_2, but kept fixed in what follows.*

(ii) n, *the unit outward normal to Ω at x if $F \in \mathcal{F}_h^b$.*

k	$d=1$	$d=2$	$d=3$
0	1	1	1
1	2	3	4
2	3	6	10
3	4	10	20

Table 1: Dimension of the polynomial space $\mathbb{P}_d^k$ for $d \in \{1,2,3\}$ and $k \in \{0,1,2,3\}$

2.4 Broken polynomial spaces

After having built a mesh of the domain Ω, the second step in the construction of discrete function spaces consists in choosing a certain functional behavior within each mesh element. For the sake of simplicity, we restrict ourselves to polynomial functions; more general cases can also be accommodated (see, e.g., Yuan and Shu [83]). The resulting spaces, consisting of piecewise polynomial functions, are termed *broken polynomial spaces*.

Let $k \geq 0$ be an integer. We focus for simplicity on the simplest polynomial space consisting of polynomials of d variables of total degree at most k. Letting

$$A_d^m := \left\{\alpha \in \mathbb{N}^d \mid |\alpha|_{\ell^1} \leq m\right\}, \qquad |\alpha|_{\ell^1} := \sum_{i=1}^d \alpha_i, \tag{4}$$

this polynomial space is defined as

$$\mathbb{P}_d^k := \left\{ p : \mathbb{R}^d \ni x \mapsto p(x) \in \mathbb{R} \mid \exists \{\gamma_\alpha\}_{\alpha \in A_d^k} \in \mathbb{R}^{\operatorname{card}(A_d^k)} \text{ s.t. } p(x) = \sum_{\alpha \in A_d^k} \gamma_\alpha x^\alpha \right\},$$

with the convention that, for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $x^\alpha := \prod_{i=1}^d x_i^{\alpha_i}$. The dimension of the vector space $\mathbb{P}_d^k$ is

$$\dim(\mathbb{P}_d^k) = \operatorname{card}(A_d^k) = \binom{k+d}{k} = \frac{(k+d)!}{k!d!}. \tag{5}$$

The first few values of $\dim(\mathbb{P}_d^k)$ are listed in Table 1.

We consider the broken polynomial space

$$\mathbb{P}_d^k(\mathcal{T}_h) := \left\{v \in L^2(\Omega) \mid \forall T \in \mathcal{T}_h,\, v|_T \in \mathbb{P}_d^k(T)\right\}, \tag{6}$$

where $\mathbb{P}_d^k(T)$ is spanned by the restriction to T of polynomials in $\mathbb{P}_d^k$. It is clear that

$$\dim(\mathbb{P}_d^k(\mathcal{T}_h)) = \operatorname{card}(\mathcal{T}_h) \times \dim(\mathbb{P}_d^k),$$

since the restriction of a function $v \in \mathbb{P}_d^k(\mathcal{T}_h)$ to each mesh element can be chosen independently of its restriction to other elements.

2.5 Abstract nonconforming error analysis

The goal of this section is to present the key ingredients for the error analysis when approximating linear model problems by dG methods. The error analysis presented in this section is derived in the spirit of Strang's Second Lemma [77] (see also Ern and Guermond [49, §2.3]). The three ingredients are (i) discrete stability, (ii) (strong) consistency, and (iii) boundedness.

2.5.1 Well-posedness for linear model problems

Let X and Y be two Banach spaces equipped with their respective norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ and assume that Y is reflexive. In many applications, X and Y are actually Hilbert spaces. We recall that $\mathcal{L}(X,Y)$ is the vector space spanned by bounded linear operators from X to Y, and that this space is equipped with the usual norm

$$\|A\|_{\mathcal{L}(X,Y)} := \sup_{v\in X\setminus\{0\}} \frac{\|Av\|_Y}{\|v\|_X} \qquad \forall A \in \mathcal{L}(X,Y).$$

We are interested in the abstract linear model problem

$$\text{Find } u \in X \text{ s.t. } a(u,w) = \langle f, w\rangle_{Y',Y} \text{ for all } w \in Y, \tag{7}$$

where $a \in \mathcal{L}(X\times Y,\mathbb{R})$ is a bounded bilinear form, $f \in Y' := \mathcal{L}(Y,\mathbb{R})$ is a bounded linear form, and $\langle\cdot,\cdot\rangle_{Y',Y}$ denotes the duality pairing between Y' and Y.

Problem (7) is said to be *well-posed* if it admits one and only one solution $u \in X$. The key result for asserting well-posedness is the so-called Banach–Nečas–Babuška (BNB) Theorem. We stress that this result provides *necessary and sufficient* conditions for well-posedness.

Theorem 2.12 (Banach–Nečas–Babuška (BNB))**.** *Let X be a Banach space and let Y be a reflexive Banach space. Let $a \in \mathcal{L}(X\times Y,\mathbb{R})$ and let $f\in Y'$. Then, problem* (7) *is well-posed if and only if:*

(i) *there is $C_{\rm sta} > 0$ such that*

$$\forall v \in X, \qquad C_{\rm sta}\|v\|_X \le \sup_{w\in Y\setminus\{0\}} \frac{a(v,w)}{\|w\|_Y}, \tag{8}$$

(ii) *For all $w \in Y$,*

$$(\forall v \in X,\ a(v,w)=0) \Longrightarrow (w=0). \tag{9}$$

Moreover, the following a priori estimate holds true:

$$\|u\|_X \le \frac{1}{C_{\rm sta}}\|f\|_{Y'}.$$

Remark 2.13 (Inf-sup condition)*.* Condition (8) is often called an inf-sup condition since it is equivalent to

$$C_{\rm sta} \le \inf_{v\in X\setminus\{0\}} \sup_{w\in Y\setminus\{0\}} \frac{a(v,w)}{\|v\|_X\|w\|_Y}.$$

A simpler, yet less general, condition to assert the well-posedness of (7) is provided by the Lax–Milgram Lemma [64]. In this setting, X is a Hilbert space, $Y = X$, and a coercivity property is invoked.

Lemma 2.14 (Lax–Milgram)**.** *Let X be a Hilbert space, let $a\in\mathcal{L}(X\times X,\mathbb{R})$, and let $f\in X'$. Then, problem* (7) *is well-posed if the bilinear form a is* coercive *on X, that is, if there is $C_{\rm sta}>0$ such that*

$$\forall v\in X, \qquad C_{\rm sta}\|v\|_X^2 \le a(v,v).$$

Moreover, the following a priori estimate holds true:

$$\|u\|_X \le \frac{1}{C_{\rm sta}}\|f\|_{X'}.$$

2.5.2 The discrete problem

Let $V_h \subset L^2(\Omega)$ denote a finite-dimensional function space; typically, V_h is a broken polynomial space. We are interested in the discrete problem

$$\text{Find } u_h \in V_h \text{ s.t. } a_h(u_h,w_h) = l_h(w_h) \text{ for all } w_h \in V_h, \tag{10}$$

with discrete bilinear form a_h defined (so far) only on $V_h \times V_h$ and discrete linear form l_h defined on V_h. We observe that we consider the so-called standard Galerkin approximation where the discrete trial and test spaces coincide. Moreover, since functions in V_h can be discontinuous across mesh elements, $V_h \not\subset X$ and $V_h \not\subset Y$ in general; cf., e.g., Lemma 2.34. In the terminology of finite elements, we say that the approximation is *nonconforming*.

We are concerned with model problems where $Y \hookrightarrow L^2(\Omega)$ with dense and continuous injection. Identifying $L^2(\Omega)$ with its topological dual space $L^2(\Omega)'$ by means of the Riesz–Fréchet representation theorem, we are thus in the situation where

$$Y \hookrightarrow L^2(\Omega) \equiv L^2(\Omega)' \hookrightarrow Y',$$

with dense and continuous injections. For simplicity, we assume that the datum f is in $L^2(\Omega)$, so that the right-hand side of the model problem (7) becomes $(f, w)_{L^2(\Omega)}$, while the right-hand side of the discrete problem (10) becomes

$$l_h(w_h) = (f, w_h)_{L^2(\Omega)}.$$

2.5.3 Discrete stability

To formulate discrete stability, we introduce a norm, say $|||\cdot|||$, defined (at least) on V_h.

Definition 2.15 (Discrete stability)**.** *We say that the discrete bilinear form a_h enjoys* discrete stability *on V_h if there is $C_{\mathrm{sta}} > 0$, independent of h, such that*

$$\forall v_h \in V_h, \qquad C_{\mathrm{sta}} |||v_h||| \le \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{|||w_h|||}. \tag{11}$$

Property (11) is referred to as a discrete inf-sup condition since it is equivalent to

$$C_{\mathrm{sta}} \le \inf_{v_h \in V_h \setminus \{0\}} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{|||v_h|||\,|||w_h|||}.$$

An important fact is that (11) is a necessary and sufficient condition for discrete well-posedness.

Lemma 2.16 (Discrete well-posedness)**.** *The discrete problem* (10) *is well-posed if and only if the discrete inf-sup condition* (11) *holds true.*

We observe that discrete well-posedness is equivalent to only one condition, namely (11), while two conditions appear in the continuous case. This is because, in finite dimension, injectivity is equivalent to bijectivity.

A sufficient, and often easily verified, condition for discrete stability is coercivity. This property can be stated as follows: There is $C_{\mathrm{sta}} > 0$ such that

$$\forall v_h \in V_h, \qquad C_{\mathrm{sta}} |||v_h|||^2 \le a_h(v_h, v_h). \tag{12}$$

Discrete coercivity implies the discrete inf-sup condition (11) since, for all $v_h \in V_h \setminus \{0\}$,

$$C_{\mathrm{sta}} |||v_h||| \le \frac{a_h(v_h, v_h)}{|||v_h|||} \le \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(v_h, w_h)}{|||w_h|||}.$$

Property (12) is the discrete counterpart of that invoked in the Lax–Milgram Lemma.

2.5.4 Consistency

For the time being, we consider a rather strong form of consistency, namely that the exact solution u satisfies the discrete equations in (10). To formulate consistency, it is thus necessary to plug the exact solution into the first argument of the discrete bilinear form a_h, and this may not be possible in general since the discrete bilinear form a_h is so far defined on $V_h \times V_h$ only. Therefore, we assume that there is a subspace $X_* \subset X$ such that the exact solution u belongs to X_* and such that the discrete bilinear form a_h can be extended to $X_* \times V_h$ (it is not possible in general to extend a_h to $X \times V_h$). Consistency can now be formulated as follows.

Definition 2.17 (Consistency)**.** *We say that the discrete problem* (10) *is* consistent *if for the exact solution* $u \in X_*$,

$$a_h(u, w_h) = l_h(w_h) \qquad \forall w_h \in V_h. \tag{13}$$

Remark 2.18 (Galerkin orthogonality). Consistency is equivalent to the usual *Galerkin orthogonality property* often considered in the context of finite element methods. Indeed, (13) holds true if and only if

$$a_h(u - u_h, w_h) = 0 \qquad \forall w_h \in V_h.$$

2.5.5 Boundedness

The last ingredient in the error analysis is boundedness. We introduce the vector space

$$X_{*h} := X_* + V_h,$$

and observe that the *approximation error* $(u - u_h)$ belongs to this space. We aim at measuring the approximation error using the discrete stability norm $|||\cdot|||$. Therefore, we assume in what follows that this norm can be extended to the space X_{*h}. In the present setting, we want to assert boundedness in the product space $X_{*h} \times V_h$, and not just in $V_h \times V_h$. It turns out that in most situations, it is not possible to assert boundedness using only the discrete stability norm $|||\cdot|||$. This is the reason why we introduce a second norm, say $|||\cdot|||_*$.

Definition 2.19 (Boundedness)**.** *We say that the discrete bilinear form* a_h *is* bounded *in* $X_{*h} \times V_h$ *if there is* C_{bnd}*, independent of* h*, such that*

$$\forall (v, w_h) \in X_{*h} \times V_h, \qquad |a_h(v, w_h)| \le C_{\text{bnd}} |||v|||_* |||w_h|||,$$

for a norm $|||\cdot|||_*$ *defined on* X_{*h} *and such that, for all* $v \in X_{*h}$, $|||v||| \le |||v|||_*$.

2.5.6 Error estimate

We can now state the main result of this section.

Theorem 2.20 (Abstract error estimate)**.** *Let* u *solve* (7) *with* $f \in L^2(\Omega)$*. Let* u_h *solve* (10)*. Let* $X_* \subset X$ *and assume that* $u \in X_*$*. Set* $X_{*h} = X_* + V_h$ *and assume that the discrete bilinear form* a_h *can be extended to* $X_{*h} \times V_h$*. Let* $|||\cdot|||$ *and* $|||\cdot|||_*$ *be two norms defined on* X_{*h} *and such that, for all* $v \in X_{*h}$, $|||v||| \le |||v|||_*$*. Assume that discrete stability, consistency, and boundedness hold true. Then, the following error estimate holds true:*

$$|||u - u_h||| \le C \inf_{y_h \in V_h} |||u - y_h|||_*, \tag{14}$$

with $C = 1 + C_{\text{sta}}^{-1} C_{\text{bnd}}$.

Proof. Let $y_h \in V_h$. Owing to discrete stability and consistency,

$$|||u_h - y_h||| \le C_{\text{sta}}^{-1} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(u_h - y_h, w_h)}{|||w_h|||} = C_{\text{sta}}^{-1} \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h(u - y_h, w_h)}{|||w_h|||}.$$

Hence, owing to boundedness,

$$|||u_h - y_h||| \le C_{\text{sta}}^{-1} C_{\text{bnd}} |||u - y_h|||_*.$$

Estimate (14) then results from the triangle inequality, the fact that $|||u - y_h||| \le |||u - y_h|||_*$, and that y_h is arbitrary in V_h. □

2.6 Admissible mesh sequences

The goal of this section is to derive some technical, yet important, tools to analyze the convergence of dG methods as the meshsize goes to zero. We are thus led to consider a mesh sequence

$$\mathcal{T}_{\mathcal{H}} := (\mathcal{T}_h)_{h\in\mathcal{H}},$$

where $\mathcal{H}$ denotes a countable subset of $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$ having 0 as only accumulation point. The analysis tools are, on the one hand, inverse and trace inequalities that are instrumental to assert discrete stability and boundedness uniformly in h and, on the other hand, optimal polynomial approximation properties so as to infer from error estimates of the form (14) h-convergence rates for the approximation error whenever the exact solution is smooth enough.

2.6.1 Shape and contact regularity

A useful concept encountered in the context of conforming finite element methods is that of matching simplicial meshes.

Definition 2.21 (Matching simplicial mesh)**.** *We say that $\mathcal{T}_h$ is a* matching simplicial mesh *if it is a simplicial mesh and if for any $T \in \mathcal{T}_h$ with vertices $\{a_0, \ldots, a_d\}$, the set $\partial T \cap \partial T'$ for any $T' \in \mathcal{T}_h$, $T' \neq T$, is the convex hull of a (possibly empty) subset of $\{a_0, \ldots, a_d\}$.*

For instance, in dimension 2, the set $\partial T \cap \partial T'$ for two distinct elements of a matching simplicial mesh is either empty, or a common vertex, or a common edge of the two elements. We now turn to the matching simplicial submesh of a general mesh.

Definition 2.22 (Matching simplicial submesh)**.** *Let $\mathcal{T}_h$ be a general mesh. We say that $\mathfrak{S}_h$ is a* matching simplicial submesh *of $\mathcal{T}_h$ if*

(i) *$\mathfrak{S}_h$ is a matching simplicial mesh,*

(ii) *for all $T' \in \mathfrak{S}_h$, there is only one $T \in \mathcal{T}_h$ such that $T' \subset T$,*

(iii) *for all $F' \in \mathfrak{F}_h$, the set collecting the mesh faces of $\mathfrak{S}_h$, there is at most one $F \in \mathcal{F}_h$ such that $F' \subset F$.*

The simplices in $\mathfrak{S}_h$ are called subelements*, and the mesh faces in $\mathfrak{F}_h$ are called* subfaces*. We set, for all $T \in \mathcal{T}_h$,*

$$\mathfrak{S}_T := \{T' \in \mathfrak{S}_h \mid T' \subset T\},$$
$$\mathfrak{F}_T := \{F' \in \mathfrak{F}_h \mid F' \subset \partial T\}.$$

We also set, for all $F \in \mathcal{F}_h$,

$$\mathfrak{F}_F := \{F' \in \mathfrak{F}_h \mid F' \subset F\}.$$

Figure 4 illustrates the matching simplicial submesh for two polygonal mesh elements, say T_1 and T_2, that come into contact. The triangular subelements composing the sets $\mathfrak{S}_{T_1}$ and $\mathfrak{S}_{T_2}$ are indicated by dashed lines. We observe that the mesh face $F = \partial T_1 \cap \partial T_2$ (highlighted in bold) is not a part of a hyperplane and that the set $\mathfrak{F}_F$ contains two subfaces.

Definition 2.23 (Shape and contact regularity)**.** *We say that the mesh sequence $\mathcal{T}_{\mathcal{H}}$ is* shape- and contact-regular *if for all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathfrak{S}_h$ such that*

(i) *the mesh sequence $\mathfrak{S}_{\mathcal{H}}$ is shape-regular in the usual sense of Ciarlet* [29]*, meaning that there is a parameter $\varrho_1 > 0$, independent of h, such that, for all $T' \in \mathfrak{S}_h$,*

$$\varrho_1 h_{T'} \leq h^{\flat}_{T'},$$

where $h_{T'}$ is the diameter of T' and $h^{\flat}_{T'}$ the diameter of the largest ball inscribed in T',

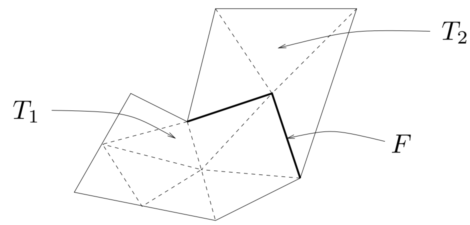

Figure 4: Two polygonal mesh elements that come into contact with corresponding subelements indicated by dashed lines and interface indicated in bold

(ii) *there is a parameter* $\varrho_2 > 0$, *independent of* h, *such that, for all* $T \in \mathcal{T}_h$ *and for all* $T' \in \mathfrak{S}_T$,

$$\varrho_2 h_T \leq h_{T'}.$$

Henceforth, the parameters ϱ_1 *and* ϱ_2 *are called the* mesh regularity parameters *and are collectively denoted by the symbol* ϱ. *Finally, if* $\mathcal{T}_h$ *is itself matching and simplicial, then* $\mathfrak{S}_h = \mathcal{T}_h$ *and the only requirment is shape-regularity with parameter* $\varrho_1 > 0$ *independent of* h.

The two conditions in Definition 2.23 allow one to control the shape of the elements in $\mathcal{T}_h$ and the way these elements come into contact. Indeed, let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$ and all $T \in \mathcal{T}_h$,

1. $\mathrm{card}(\mathfrak{S}_T)$ is bounded uniformly in h;
2. $\mathrm{card}(\mathcal{F}_T)$, $\mathrm{card}(\mathfrak{F}_T)$, and N_∂ are bounded uniformly in h;
3. all $F \in \mathcal{F}_T$, $\delta_F \geq \varrho_1 \varrho_2 h_T$, where δ_F denotes the *diameter* of F, and this implies that the diamters of neighboring mesh elements are uniformly comparable.

2.6.2 Inverse and trace inequalities

Lemma 2.24 (Inverse inequality)**.** *Let* $\mathcal{T}_\mathcal{H}$ *be a shape- and contact-regular mesh sequence with regularity parameters* ϱ. *Then, for all* $h \in \mathcal{H}$, *all* $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$, *and all* $T \in \mathcal{T}_h$,

$$\|\nabla v_h\|_{[L^2(T)]^d} \leq C_{\mathrm{inv}} h_T^{-1} \|v_h\|_{L^2(T)}, \tag{15}$$

where C_{inv} *only depends on* ϱ, d, *and* k.

Lemma 2.25 (Discrete trace inequality)**.** *Let* $\mathcal{T}_\mathcal{H}$ *be a shape- and contact-regular mesh sequence with regularity parameters* ϱ. *Then, for all* $h \in \mathcal{H}$, *all* $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$, *all* $T \in \mathcal{T}_h$, *and all* $F \in \mathcal{F}_T$,

$$h_T^{1/2} \|v_h\|_{L^2(F)} \leq C_{\mathrm{tr}} \|v_h\|_{L^2(T)}, \tag{16}$$

where C_{tr} *only depends on* ϱ, d, *and* k.

2.6.3 Polynomial approximation

To infer from estimate (14) a convergence rate in h for the *approximation error* $(u - u_h)$ measured in the $|\!|\!|\cdot|\!|\!|$-norm when the exact solution u is smooth enough, we need to estimate the right-hand side given by

$$\inf_{y_h \in V_h} |\!|\!| u - y_h |\!|\!|_*,$$

when V_h is typically the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined by (6); other broken polynomial spaces can be considered. Since $u_h \in V_h$, we infer from (14) that

$$\inf_{y_h \in V_h} |\!|\!| u - y_h |\!|\!| \leq |\!|\!| u - u_h |\!|\!| \leq C \inf_{y_h \in V_h} |\!|\!| u - y_h |\!|\!|_*. \tag{17}$$

Definition 2.26 (Optimality, quasi-optimality, and suboptimality of the error estimate). *We say that the error estimate* (17) *is*

(i) optimal *if* $|||\cdot||| = |||\cdot|||_*$,

(ii) quasi-optimal *if the two norms are different, but the lower and upper bounds in* (17) *converge, for smooth enough* u, *at the same convergence rate as* $h \to 0$,

(iii) suboptimal *if the upper bound converges at a slower rate 'than the lower bound.*

The analysis of the upper bound $\inf_{y_h \in V_h} |||u - y_h|||_*$ depends on the polynomial approximation properties that can be achieved in the broken polynomial space V_h. The approximation error is measured using Sobolev norms that are defined in §2.7. In what follows, π_h denotes the $L^2(\Omega)$-*orthogonal projection* onto V_h, that is, $\pi_h : L^2(\Omega) \to V_h$ is defined so that, for all $v \in L^2(\Omega)$, $\pi_h v \in V_h$ with

$$(\pi_h v, y_h)_{L^2(\Omega)} = (v, y_h)_{L^2(\Omega)} \qquad \forall y_h \in V_h. \tag{18}$$

We observe that the restriction of $\pi_h v$ to a given mesh element $T \in \mathcal{T}_h$ can be computed independently from other mesh elements. For instance, if $V_h = \mathbb{P}^k_d(\mathcal{T}_h)$, we obtain that, for all $T \in \mathcal{T}_h$, $\pi_h v|_T \in \mathbb{P}^k_d(T)$ is such that

$$(\pi_h v|_T, \xi)_{L^2(T)} = (v, \xi)_{L^2(T)} \qquad \forall \xi \in \mathbb{P}^k_d(T).$$

Definition 2.27 (Optimal polynomial approximation). *We say that the mesh sequence* $\mathcal{T}_\mathcal{H}$ *has* optimal polynomial approximation *properties if, for all* $h \in \mathcal{H}$, *all* $T \in \mathcal{T}_h$, *all polynomial degree* k, *all* $s \in \{0, \dots, k+1\}$, *and all* $v \in H^s(T)$, *there holds*

$$|v - \pi_h v|_{H^m(T)} \leq C'_{\mathrm{app}} h_T^{s-m} |v|_{H^s(T)} \qquad \forall m \in \{0, \dots, s\}, \tag{19}$$

where C_{app} *is independent of both* T *and* h. *Moreover, for all* $F \in \mathcal{F}_T$, *there holds*

$$\|v - \pi_h v\|_{L^2(F)} \leq C'_{\mathrm{app}} h_T^{s-1/2} |v|_{H^s(T)},$$

and if $s \geq 2$,

$$\|\nabla(v - \pi_h v)|_T \cdot \mathrm{n}_T\|_{L^2(F)} \leq C''_{\mathrm{app}} h_T^{s-3/2} |v|_{H^s(T)},$$

where C'_{app} *and* C''_{app} *are independent of both* T *and* h.

Definition 2.28 (Admissible mesh sequences). *We say that the mesh sequence* $\mathcal{T}_\mathcal{H}$ *is* admissible *if it is shape- and contact-regular and if it has optimal polynomial approximation properties.*

On general meshes, asserting optimal polynomial approximation is a delicate question since this property depends on the shape of mesh elements. In practice, meshes are generated by successive refinements of an initial mesh, and the shape of mesh elements depends on the refinement procedure. It is convenient to identify sufficient conditions on the mesh sequence $\mathcal{T}_\mathcal{H}$ to assert optimal polynomial approximation in broken polynomial spaces. One approach is based on the star-shaped property with respect to a ball (see, e.g., Brenner and Scott [15, Chapter 4]).

Definition 2.29 (Star-shaped property with respect to a ball). *We say that a polyhedron* P *is* star-shaped with respect to a ball *if there is a ball* $\mathfrak{B}_P \subset P$ *such that, for all* $x \in P$, *the convex hull of* $\{x\} \cup \mathfrak{B}_P$ *is included in* $\overline{P}$.

Figure 5 displays two polyhedra. The one on the left is star-shaped with respect to the ball indicated in black. Instead, the one on the right is not star-shaped with respect to any ball.

Lemma 2.30 (Mesh sequences with star-shaped property). *Let* $\mathcal{T}_\mathcal{H}$ *be a shape- and contact-regular mesh sequence. Assume that, for all* $h \in \mathcal{H}$ *and all* $T \in \mathcal{T}_h$, *the mesh element* T *is* star-shaped with respect to a ball *with uniformly comparable diameter with respect to* h_T. *Then, the mesh sequence* $\mathcal{T}_\mathcal{H}$ *is admissible.*

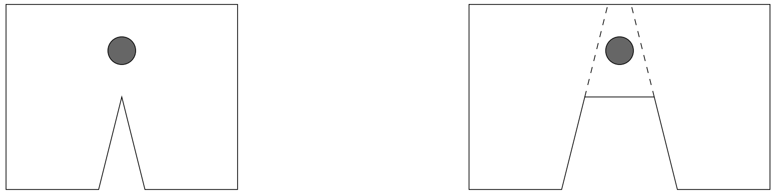

Figure 5: Example (left) and counter-example (right) of a polyhedron which is star-shaped with respect to a ball

Another sufficient condition ensuring optimal polynomial approximation, but somewhat less general than the star-shaped property, is that of finitely shaped mesh sequences. A simple example is that of shape- and contact-regular mesh sequences whose elements are either simplices or parallelotopes in $\mathbb{R}^d$.

Lemma 2.31 (Finitely shaped mesh sequences)**.** *Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence. Assume that $\mathcal{T}_\mathcal{H}$ is* finitely shaped *in the sense that there is a finite set $\widehat{\mathcal{R}} = \{\widehat{T}\}$ whose elements are reference polyhedra in $\mathbb{R}^d$ and such that, for all $h \in \mathcal{H}$, each $T \in \mathcal{T}_h$ is the image of a reference polyhedron in $\widehat{\mathcal{R}}$ by an affine bijective map F_T. Then, the mesh sequence $\mathcal{T}_\mathcal{H}$ is admissible.*

2.7 Some background on functional analysis

In this section, we briefly present two important classes of function spaces, namely Lebesgue and Sobolev spaces. We only state the basic properties of such spaces, and we refer the reader to Evans [53, Chapter 5] or Brézis [16, Chapters 8 and 9] for further background. We also introduce broken Sobolev spaces.

2.7.1 Lebesgue spaces

We consider functions $v : \Omega \to \mathbb{R}$ that are Lebesgue measurable and we denote by $\int_\Omega v$ the (Lebesgue) integral of v over Ω. Let $1 \leq p \leq \infty$ be a real number. We set

$$\|v\|_{L^p(\Omega)} := \left(\int_\Omega |v|^p \right)^{1/p} \qquad 1 \leq p < \infty,$$

and

$$\begin{aligned} \|v\|_{L^\infty(\Omega)} &:= \sup\operatorname{ess}\{|v(x)| \text{ a.e. } x \in \Omega\} \\ &= \inf\{M > 0 \mid |v(x)| \leq M \text{ a.e. } x \in \Omega\}. \end{aligned}$$

In either case, we define the *Lebesgue space*

$$L^p(\Omega) := \{v \text{ Lebesgue measurable} \mid \|v\|_{L^p(\Omega)} < \infty\}.$$

Equipped with the norm $\|\cdot\|_{L^p(\Omega)}$, $L^p(\Omega)$ is a Banach space for all $1 \leq p \leq \infty$ (see Evans [53, p. 249] or Brézis [16, p. 150]). In the particular case $p = 2$, $L^2(\Omega)$ is a (real) Hilbert space when equipped with the scalar product

$$(v, w)_{L^2(\Omega)} := \int_\Omega vw.$$

The Cauchy–Schwarz inequality states that, for all $v, w \in L^2(\Omega)$,

$$(v, w)_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)}.$$

2.7.2 Sobolev spaces

On the Cartesian basis of $\mathbb{R}^d$ with coordinates $(x_1, \dots, x_d)$, the symbol ∂_i with $i \in \{1, \dots, d\}$ denotes the distributional partial derivative with respect to x_i. For a d-uple $\alpha \in \mathbb{N}^d$, $\partial^\alpha v$ denotes the distributional derivative $\partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} v$ of v, with the convention that $\partial^{(0,\dots,0)} v = v$.

Let $m \geq 0$ be an integer. We define the *Sobolev space*

$$H^m(\Omega) = \left\{ v \in L^2(\Omega) \mid \forall \alpha \in A_d^m,\ \partial^\alpha v \in L^2(\Omega) \right\},$$

with A_d^m defined by (4). $H^m(\Omega)$ is a Hilbert space when equipped with the scalar product

$$(v, w)_{H^m(\Omega)} := \sum_{\alpha \in A_d^m} (\partial^\alpha v, \partial^\alpha w)_{L^2(\Omega)},$$

leading to the norm and seminorm

$$\|v\|_{H^m(\Omega)} := \left(\sum_{\alpha \in A_d^m} \|\partial^\alpha v\|_{L^2(\Omega)}^2 \right)^{1/2}, \qquad |v|_{H^m(\Omega)} := \left(\sum_{\alpha \in \overline{A}_d^m} \|\partial^\alpha v\|_{L^2(\Omega)}^2 \right)^{1/2}.$$

The seminorm is obtained by restricting the summation to the set $\overline{A}_d^m := \left\{ \alpha \in \mathbb{N}^d \mid |\alpha|_{\ell^1} = m \right\}$, that is, by keeping only the derivatives of global order m. To allow for a more compact notation in the case $m = 1$, we consider the gradient $\nabla v = (\partial_1 v, \dots, \partial_d v)^t$ with values in $\mathbb{R}^d$, yielding

$$(v, w)_{H^1(\Omega)} = (v, w)_{L^2(\Omega)} + (\nabla v, \nabla w)_{[L^2(\Omega)]^d}.$$

Boundary values of functions in the Sobolev space $H^1(\Omega)$ can be given a meaning (at least) in $L^2(\partial\Omega)$. More precisely (see, e.g., Brenner and Scott [15, Chap. 1]), there is C such that

$$\|v\|_{L^2(\partial\Omega)} \leq C \|v\|_{L^2(\Omega)}^{1/2} \|v\|_{H^1(\Omega)}^{1/2} \qquad \forall v \in H^1(\Omega). \tag{20}$$

2.7.3 Broken Sobolev spaces and broken gradient

Let $\mathcal{T}_h$ be a mesh of the domain Ω. For any mesh element $T \in \mathcal{T}_h$, the Sobolev spaces $H^m(T)$ can be defined as above by replacing Ω by T. We then define the broken Sobolev spaces

$$H^m(\mathcal{T}_h) := \left\{ v \in L^2(\Omega) \mid \forall T \in \mathcal{T}_h,\ v|_T \in H^m(T) \right\}, \tag{21}$$

(22)

where $m \geq 0$ is an integer. It is natural to define a broken gradient operator acting on the broken Sobolev space $H^1(\mathcal{T}_h)$. In particular, this operator also acts on broken polynomial spaces.

Definition 2.32 (Broken gradient)**.** *The* broken gradient $\nabla_h : H^1(\mathcal{T}_h) \to [L^2(\Omega)]^d$ *is defined such that, for all* $v \in H^1(\mathcal{T}_h)$,

$$\forall T \in \mathcal{T}_h, \qquad (\nabla_h v)|_T := \nabla(v|_T). \tag{23}$$

In what follows, we drop the index h in the broken gradient when this operator appears inside an integral over a fixed mesh element $T \in \mathcal{T}_h$.

It is important to observe that the usual Sobolev spaces are subspaces of the broken Sobolev spaces, and that on the usual Sobolev spaces, the broken gradient coincides with the distributional gradient.

Lemma 2.33 (Broken gradient on usual Sobolev spaces)**.** *Let $m \geq 0$. There holds $H^m(\Omega) \subset H^m(\mathcal{T}_h)$. Moreover, for all $v \in H^1(\Omega)$, $\nabla_h v = \nabla v$ in $[L^2(\Omega)]^d$.*

The reverse inclusion of Lemma 2.33 does not hold true in general (except obviously for $m = 0$). The reason is that functions in the broken Sobolev space $H^1(\mathcal{T}_h)$ can have nonzero jumps across interfaces, while functions in the usual Sobolev space $H^1(\Omega)$ have zero jumps across interfaces. We now give a precise statement of this important result.

Lemma 2.34 (Characterization of $H^1(\Omega)$)**.** *A function $v \in H^1(\mathcal{T}_h)$ belongs to $H^1(\Omega)$ if and only if*

$$[\![v]\!] = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{24}$$

3 Advection-reaction

The steady advection-reaction equation with homogeneous inflow boundary condition

$$\beta\cdot\nabla u + \mu u = f \quad \text{in } \Omega, \tag{25a}$$

$$u = 0 \quad \text{on } \partial\Omega^-, \tag{25b}$$

is one of the simplest model problems based on a linear, scalar, steady first-order PDE. Here, the unknown function u is scalar-valued and represents, e.g., a solute concentration; β is the $\mathbb{R}^d$-valued advective velocity, μ the reaction coefficient, f the source term, and $\partial\Omega^-$ denotes the inflow part of the boundary of Ω, namely

$$\partial\Omega^- := \{x \in \partial\Omega \mid \beta(x)\cdot\mathrm{n}(x) < 0\}. \tag{26}$$

The goal of this section is to design and analyze dG methods to approximate the model problem (25). Since dG methods are essentially tailored to approximate PDEs in an L^2-setting where discrete stability is enhanced by suitable least-squares penalties, the most natural weak formulation at the continuous level is that based on the concept of graph space. Moreover, we formulate the boundary condition (25b) weakly in the continuous problem since this is the way boundary conditions are enforced in dG methods. Then, we present a step-by-step derivation of suitable dG bilinear forms that match the discrete stability, consistency, and boundedness properties outlined in §2.5 for nonconforming finite element error analysis. We also discuss an alternative viewpoint using local (elementwise) problems and numerical fluxes. Two dG methods are analyzed, resulting from the use of so-called centered or upwind fluxes.

3.1 Assumptions on the data

We assume that

$$\mu \in L^\infty(\Omega), \qquad \beta \in [\mathrm{Lip}(\Omega)]^d, \tag{27}$$

where $\mathrm{Lip}(\Omega)$ denotes the space spanned by Lipschitz continuous functions, that is, $v \in \mathrm{Lip}(\Omega)$ means that there is L_v such that, for all $x, y \in \Omega$, $|v(x) - v(y)| \le L_v|x-y|$ where $|x-y|$ denotes the Euclidean norm of $(x-y)$ in $\mathbb{R}^d$. The quantity L_v is called the *Lipschitz module* of v. In what follows, we set $L_\beta := \max_{1\le i\le d} L_{\beta_i}$. In addition to (27), we assume that there is a real number $\mu_0 > 0$ such that

$$\Lambda := \mu - \frac{1}{2}\nabla\cdot\beta \ge \mu_0 \text{ a.e. in } \Omega. \tag{28}$$

Concerning the source term f, we assume that

$$f \in L^2(\Omega).$$

Finally, we recall that Ω is a polyhedron in $\mathbb{R}^d$ (cf. Definition 2.1). This assumption is solely made to facilitate the meshing of Ω.

We consider a reference time τ_{c} and a reference velocity β_{c} defined as

$$\tau_{\mathrm{c}} := \{\max(\|\mu\|_{L^\infty(\Omega)}, L_\beta)\}^{-1}, \qquad \beta_{\mathrm{c}} := \|\beta\|_{[L^\infty(\Omega)]^d}. \tag{29}$$

Since μ and L_β scale as the reciprocal of a time, τ_c can be interpreted as the (fastest) time scale in the problem. Moreover, β_c represents the maximum velocity. We observe that τ_c is finite since $\|\mu\|_{L^\infty(\Omega)} = L_\beta = 0$ implies $\Lambda = 0$ which contradicts (28). We keep track of the parameters τ_c and β_c in the convergence analysis of dG approximations. This allows us to work with norms consisting of terms having the same physical dimension. Keeping track of these parameters is also useful when dealing with singularly perturbed regimes. For simplicity, the reader can assume that both parameters are of order unity and discard them in what follows.

3.2 The continuous setting

Our first goal is to specify the functional space in which the solution to the model problem (25) is sought. Let $C_0^\infty(\Omega)$ denote the space of infinitely differentiable functions with compact support in Ω and recall that this space is dense in $L^2(\Omega)$. For a function $v \in L^2(\Omega)$, the statement $\beta\cdot\nabla v \in L^2(\Omega)$ means that the linear form

$$C_0^\infty(\Omega) \ni \varphi \longmapsto -\int_\Omega v\nabla\cdot(\beta\varphi) \in \mathbb{R}$$

is bounded in $L^2(\Omega)$, that is, there is C_v such that

$$\forall \varphi \in C_0^\infty(\Omega), \qquad \int_\Omega v\nabla\cdot(\beta\varphi) \le C_v\|\varphi\|_{L^2(\Omega)}.$$

The function $\beta\cdot\nabla v$ is then defined as the function representing this linear form in $L^2(\Omega)$ by means of the Riesz–Fréchet theorem.

Definition 3.1 (Graph space). *The* graph space *is defined as*

$$V := \left\{ v \in L^2(\Omega) \mid \beta\cdot\nabla v \in L^2(\Omega) \right\}, \tag{30}$$

and is equipped with the natural scalar product: For all $v, w \in V$,

$$(v, w)_V := (v, w)_{L^2(\Omega)} + (\beta\cdot\nabla v, \beta\cdot\nabla w)_{L^2(\Omega)}, \tag{31}$$

and the associated graph norm $\|v\|_V = (v, v)_V^{1/2}$.

Proposition 3.2 (Hilbertian structure of graph space). *The graph space V defined by* (30) *and equipped with the scalar product* (31) *is a Hilbert space.*

The next step is to specify mathematically the meaning of the boundary condition (25b). To this purpose, we need to investigate the trace on $\partial\Omega$ of functions in the graph space V. Our aim is to give a meaning to such traces in the space

$$L^2(|\beta\cdot\mathrm{n}|; \partial\Omega) := \left\{ v \text{ is measurable on } \partial\Omega \mid \int_{\partial\Omega} |\beta\cdot\mathrm{n}|v^2 < \infty \right\}. \tag{32}$$

Recalling definition (26) of the inflow boundary, we also define the outflow boundary as

$$\partial\Omega^+ := \{x \in \partial\Omega \mid \beta(x)\cdot\mathrm{n}(x) > 0\},$$

and following [50], we assume that the inflow and outflow boundaries are well-separated, namely

$$\mathrm{dist}(\partial\Omega^-, \partial\Omega^+) := \min_{(x,y)\in\partial\Omega^-\times\partial\Omega^+} |x - y| > 0.$$

The following result is very important since it allows us to define traces of functions belonging to the graph space and to use an integration by parts formula.

Lemma 3.3 (Traces and integration by parts)**.** *In the above framework, the trace operator*

$$\gamma : C^0(\overline{\Omega}) \ni v \longmapsto \gamma(v) := v|_{\partial\Omega} \in L^2(|\beta{\cdot}\mathrm{n}|;\partial\Omega)$$

extends continuously to V, meaning that there is C_γ such that, for all $v \in V$,

$$\|\gamma(v)\|_{L^2(|\beta\cdot\mathrm{n}|;\partial\Omega)} \le C_\gamma \|v\|_V.$$

Moreover, the following integration by parts formula holds true: For all $v, w \in V$,

$$\int_\Omega [(\beta{\cdot}\nabla v)w + (\beta{\cdot}\nabla w)v + (\nabla{\cdot}\beta)vw] = \int_{\partial\Omega} (\beta{\cdot}\mathrm{n})vw. \tag{33}$$

For a real number x, we define its positive and negative parts respectively as

$$x^\oplus := \frac{1}{2}(|x| + x), \qquad x^\ominus := \frac{1}{2}(|x| - x). \tag{34}$$

We observe that both quantities are, by definition, nonnegative. We introduce the following bilinear form: For all $v, w \in V$,

$$a(v, w) := \int_\Omega \mu vw + \int_\Omega (\beta{\cdot}\nabla v)w + \int_{\partial\Omega} (\beta{\cdot}\mathrm{n})^\ominus vw. \tag{35}$$

This bilinear form is bounded in $V \times V$ owing to Lemma 3.3. Precisely, for all $v, w \in V$, the Cauchy–Schwarz inequality yields

$$|a(v, w)| \le (1 + \|\mu\|^2_{L^\infty(\Omega)})^{1/2} \|v\|_V \|w\|_{L^2(\Omega)} + C_\gamma \|v\|_V \|w\|_V.$$

Using the graph space V and the bilinear form a, the model problem (25) can be cast into the weak form

$$\text{Find } u \in V \text{ s.t. } a(u, w) = \int_\Omega fw \text{ for all } w \in V. \tag{36}$$

This problem turns out to be well-posed (cf. Theorem 3.6). Before addressing this, we examine in which sense does a solution to (36) solve the original problem (25). In particular, we observe that the boundary condition is weakly enforced in (36).

Proposition 3.4 (Characterization of the solution to (36))**.** *Assume that $u \in V$ solves* (36)*. Then,*

$$\beta{\cdot}\nabla u + \mu u = f \quad \text{a.e. in } \Omega, \tag{37}$$

$$u = 0 \quad \text{a.e. in } \partial\Omega^-. \tag{38}$$

An important (yet, not sufficient) ingredient for the well-posedness of the weak problem (36) is the L^2-coercivity of a in the graph space V.

Lemma 3.5 (L^2-coercivity of a)**.** *The bilinear form a defined by* (35) *is L^2-coercive on V, namely,*

$$\forall v \in V, \qquad a(v, v) \ge \mu_0 \|v\|^2_{L^2(\Omega)} + \int_{\partial\Omega} \frac{1}{2}|\beta{\cdot}\mathrm{n}|v^2. \tag{39}$$

Proof. This is a straightforward consequence of assumption (28) and of the integration by parts formula (33) since, for all $v \in V$,

$$\begin{aligned} a(v, v) &= \int_\Omega \left(\mu - \frac{1}{2}\nabla{\cdot}\beta\right) v^2 + \int_{\partial\Omega} \frac{1}{2}(\beta{\cdot}\mathrm{n})v^2 + \int_{\partial\Omega} (\beta{\cdot}\mathrm{n})^\ominus v^2 \\ &= \int_\Omega \Lambda v^2 + \int_{\partial\Omega} \frac{1}{2}|\beta{\cdot}\mathrm{n}|v^2 \ge \mu_0\|v\|^2_{L^2(\Omega)} + \int_{\partial\Omega} \frac{1}{2}|\beta{\cdot}\mathrm{n}|v^2, \end{aligned}$$

completing the proof. □

A consequence of Lemma 3.5 is that the weak problem (36) admits at most one solution. We are now in a position to state the main result of this theoretical section.

Theorem 3.6 (Well-posedness)**.** *Problem* (36) *is well-posed.*

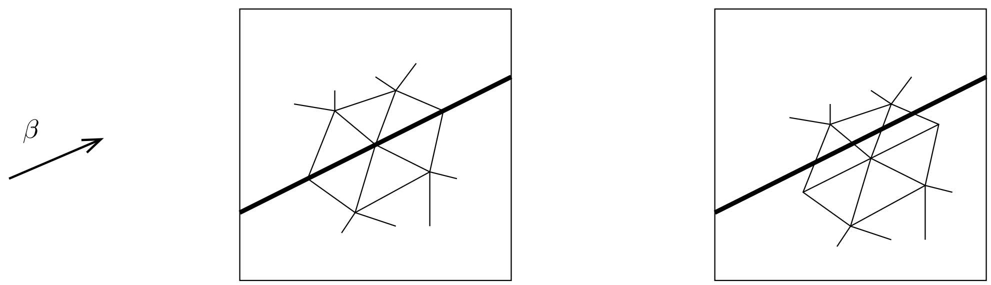

Figure 6: Fitted (left) and unfitted (right) simplicial mesh; the partition P_Ω consists of two polygons, and the exact solution can jump across the thick line

3.3 Centered fluxes

The goal of this section is to design and analyze the simplest dG method to approximate the model problem (36). Referring the reader to §2.5, the method is designed so as to be consistent, and a minimal discrete stability is ensured by L^2-coercivity. Using the terminology of Definition 2.26, the resulting error estimate turns out to be suboptimal. Alternatively, the method can be viewed as based on the use of centered fluxes.

We seek an approximate solution in the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined by (6). We assume $k \geq 1$ and that $\mathcal{T}_h$ belongs to an admissible mesh sequence. We set

$$V_h := \mathbb{P}_d^k(\mathcal{T}_h)$$

and consider the discrete problem:

$$\text{Find } u_h \in V_h \text{ s.t. } a_h(u_h, v_h) = \int_\Omega f v_h \text{ for all } v_h \in V_h,$$

for a discrete bilinear form a_h yet to be designed.

To analyze the method, we make a slightly more stringent regularity assumption on the exact solution u rather than just belonging to the graph space V. This assumption is needed to formulate the consistency of the method by directly plugging in the exact solution into the discrete bilinear form a_h. In particular, we need to consider the trace of the exact solution on each mesh face.

Assumption 3.7 (Regularity of exact solution and space V_*)**.** *We assume that there is a partition $P_\Omega = \{\Omega_i\}_{1 \leq i \leq N_\Omega}$ of Ω into disjoint polyhedra such that, for the exact solution u,*

$$u \in V_* := V \cap H^1(P_\Omega).$$

In the spirit of §2.5, *we set $V_{*h} := V_* + V_h$.*

Assumption 3.7 implies that, for all $T \in \mathcal{T}_h$, the restriction $u|_T$ has traces a.e. on each face $F \in \mathcal{F}_T$, and these traces belong to $L^2(F)$.

Lemma 3.8 (Jumps of u across interfaces)**.** *The exact solution $u \in V_*$ is such that, for all $F \in \mathcal{F}_h^i$,*

$$(\beta \cdot \mathrm{n}_F)[\![u]\!](x) = 0 \qquad \text{for a.e. } x \in F. \tag{40}$$

Remark 3.9 (Singularities of exact solution)*.* Condition (40) does not say anything on the jumps of the exact solution across interfaces to which the advective velocity β is tangential. We also observe that Assumption 3.7 does not require the mesh to be fitted to solution singularities, that is, both situations depicted in Figure 6 are admissible.

3.3.1 Heuristic derivation

The main idea in the design of the discrete bilinear form a_h is to mimic at the discrete level the L^2-coercivity that holds at the continuous level (cf. (39)), while, at the same time, ensuring consistency. Our starting point is a discrete bilinear form $a_h^{(0)}$ simply derived from the exact bilinear form a by replacing the exact gradient by the broken gradient (cf. (23) for its definition), namely, we define on $V_{*h} \times V_h$,

$$a_h^{(0)}(v, w_h) := \int_\Omega \Big\{ \mu v w_h + (\beta\cdot\nabla_h v) w_h \Big\} + \int_{\partial\Omega} (\beta\cdot\mathrm{n})^\ominus v w_h.$$

That $a_h^{(0)}$ yields consistency is clear since the exact solution satisfies (37) and (38).

Let us now focus on discrete coercivity. An important observation is that this property is not transferred from a to $a_h^{(0)}$. Indeed, integration by parts on each mesh element yields, for all $v_h \in V_h$,

$$\begin{aligned} a_h^{(0)}(v_h, v_h) &= \int_\Omega \Big\{ \mu v_h^2 + (\beta\cdot\nabla_h v_h) v_h \Big\} + \int_{\partial\Omega} (\beta\cdot\mathrm{n})^\ominus v_h^2 \\ &= \int_\Omega \mu v_h^2 + \sum_{T\in\mathcal{T}_h} \int_T (\beta\cdot\nabla v_h) v_h + \int_{\partial\Omega} (\beta\cdot\mathrm{n})^\ominus v_h^2 \\ &= \int_\Omega \Lambda v_h^2 + \sum_{T\in\mathcal{T}_h} \int_{\partial T} \frac{1}{2} (\beta\cdot\mathrm{n}_T) v_h^2 + \int_{\partial\Omega} (\beta\cdot\mathrm{n})^\ominus v_h^2, \end{aligned}$$

where we recall that $\Lambda = \mu - \frac{1}{2}\nabla\cdot\beta$ and that n_T denotes the outward normal to T on ∂T. The second term on the right-hand side can be reformulated as a sum over mesh faces. Indeed, exploiting the continuity of (the normal component of) β across interfaces leads to

$$\sum_{T\in\mathcal{T}_h} \int_{\partial T} \frac{1}{2} (\beta\cdot\mathrm{n}_T) v_h^2 = \sum_{F\in\mathcal{F}_h^i} \int_F \frac{1}{2} (\beta\cdot\mathrm{n}_F) [\![v_h^2]\!] + \sum_{F\in\mathcal{F}_h^b} \int_F \frac{1}{2} (\beta\cdot\mathrm{n}) v_h^2.$$

For all $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$, $v_i = v_h|_{T_i}$, $i \in \{1, 2\}$, there holds

$$\frac{1}{2} [\![v_h^2]\!] = \frac{1}{2}(v_1^2 - v_2^2) = \frac{1}{2}(v_1 - v_2)(v_1 + v_2) = [\![v_h]\!] \{\!\{ v_h \}\!\}.$$

As a result,

$$\begin{aligned} a_h^{(0)}(v_h, v_h) = &\int_\Omega \Lambda v_h^2 + \sum_{F\in\mathcal{F}_h^i} \int_F (\beta\cdot\mathrm{n}_F) [\![v_h]\!] \{\!\{ v_h \}\!\} \\ &+ \sum_{F\in\mathcal{F}_h^b} \int_F \frac{1}{2} (\beta\cdot\mathrm{n}) v_h^2 + \int_{\partial\Omega} (\beta\cdot\mathrm{n})^\ominus v_h^2, \end{aligned}$$

and combining the two rightmost terms, we arrive at

$$a_h^{(0)}(v_h, v_h) = \int_\Omega \Lambda v_h^2 + \sum_{F\in\mathcal{F}_h^i} \int_F (\beta\cdot\mathrm{n}_F) [\![v_h]\!] \{\!\{ v_h \}\!\} + \int_{\partial\Omega} \frac{1}{2} |\beta\cdot\mathrm{n}| v_h^2.$$

The second term on the right-hand side, involving interfaces, has no sign *a priori*. Therefore, it must be removed, and this can be achieved while maintaining consistency if we set, for all $(v, w_h) \in V_{*h} \times V_h$,

$$\begin{aligned} a_h^{\mathrm{cf}}(v, w_h) := &\int_\Omega \Big\{ \mu v w_h + (\beta\cdot\nabla_h v) w_h \Big\} + \int_{\partial\Omega} (\beta\cdot\mathrm{n})^\ominus v w_h \\ &- \sum_{F\in\mathcal{F}_h^i} \int_F (\beta\cdot\mathrm{n}_F) [\![v]\!] \{\!\{ w_h \}\!\}, \end{aligned} \tag{41}$$

since $(\beta{\cdot}\mathrm{n}_F)[\![u]\!] = 0$ for all $F \in \mathcal{F}_h^i$ owing to (40). The superscript indicates the use of centered fluxes, as detailed in §3.3.3.

We can now summarize the properties of the discrete bilinear form a_h^{cf} established so far. The coercivity of a_h^{cf} is expressed using the following norm defined on V_{*h}:

$$|\!|\!|v|\!|\!|_{\mathrm{cf}}^2 := \tau_{\mathrm{c}}^{-1}\|v\|_{L^2(\Omega)}^2 + \int_{\partial\Omega} \frac{1}{2}|\beta{\cdot}\mathrm{n}|v^2, \tag{42}$$

with the time scale τ_{c} defined by (29). We observe that $|\!|\!|\cdot|\!|\!|_{\mathrm{cf}}$ is indeed a norm since it controls the L^2-norm.

Lemma 3.10 (Consistency and discrete coercivity). *The discrete bilinear form a_h^{cf} defined by* (41)

(i) *is* consistent*, namely for the exact solution $u \in V_*$,*

$$a_h^{\mathrm{cf}}(u, v_h) = \int_\Omega f v_h \qquad \forall v_h \in V_h,$$

(ii) *is* coercive *on V_h with respect to the $|\!|\!|\cdot|\!|\!|_{\mathrm{cf}}$-norm, namely*

$$\forall v_h \in V_h, \qquad a_h^{\mathrm{cf}}(v_h, v_h) \geq C_{\mathrm{sta}}|\!|\!|v_h|\!|\!|_{\mathrm{cf}}^2,$$

with $C_{\mathrm{sta}} := \min(1, \tau_{\mathrm{c}}\mu_0)$.

Before proceeding further, we record an equivalent expression of the discrete bilinear form a_h^{cf} obtained after integrating by parts the advective derivative in each mesh element. This expression is useful when introducing the notion of fluxes in §3.3.3 and when analyzing the dG method based on upwinding in §3.4. For all $(v, w_h) \in V_{*h} \times V_h$, there holds

$$\begin{aligned} a_h^{\mathrm{cf}}(v, w_h) = &\int_\Omega \Big\{(\mu - \nabla{\cdot}\beta)v w_h - v(\beta{\cdot}\nabla_h w_h)\Big\} + \int_{\partial\Omega} (\beta{\cdot}\mathrm{n})^{\oplus} v w_h \\ &+ \sum_{F\in\mathcal{F}_h^i} \int_F (\beta{\cdot}\mathrm{n}_F)\{\!\{v\}\!\}[\![w_h]\!]. \end{aligned} \tag{43}$$

3.3.2 Error estimates

We consider the discrete problem:

$$\text{Find } u_h \in V_h \text{ s.t. } a_h^{\mathrm{cf}}(u_h, v_h) = \int_\Omega f v_h \text{ for all } v_h \in V_h. \tag{44}$$

This problem is well-posed owing to the discrete coercivity of a_h^{cf} on V_h. Our goal is to estimate the approximation error $(u - u_h)$ in the $|\!|\!|\cdot|\!|\!|_{\mathrm{cf}}$-norm. The convergence analysis is performed in the spirit of Theorem 2.20. Owing to Lemma 3.10, it only remains to address the boundedness of the discrete bilinear form a_h^{cf}. To this purpose, we define on V_{*h} the norm

$$|\!|\!|v|\!|\!|_{\mathrm{cf},*}^2 = |\!|\!|v|\!|\!|_{\mathrm{cf}}^2 + \sum_{T\in\mathcal{T}_h} \tau_{\mathrm{c}}\|\beta{\cdot}\nabla v\|_{L^2(T)}^2 + \sum_{T\in\mathcal{T}_h} \tau_{\mathrm{c}}\beta_{\mathrm{c}}^2 h_T^{-1}\|v\|_{L^2(\partial T)}^2,$$

with time scale τ_{c} and reference velocity β_{c} defined by (29). There holds

$$\forall (v, w_h) \in V_{*h} \times V_h, \qquad a_h^{\mathrm{cf}}(v, w_h) \leq C_{\mathrm{bnd}}|\!|\!|v|\!|\!|_{\mathrm{cf},*}|\!|\!|w_h|\!|\!|_{\mathrm{cf}},$$

with C_{bnd} independent of h and of the data μ and β.

Theorem 3.11 (Error estimate and convergence rate)**.** *Let u solve (36) and let u_h solve (44) where a_h^{cf} is defined by (41) and $V_h = \mathbb{P}_d^k(\mathcal{T}_h)$ with $k \geq 1$ and $\mathcal{T}_h$ belongs to an admissible mesh sequence. Then, there holds*

$$|||u - u_h|||_{\mathrm{cf}} \leq C \inf_{y_h \in V_h} |||u - y_h|||_{\mathrm{cf},*}, \tag{45}$$

with C independent of h and depending on the data only through the factor $\{\min(1, \tau_{\mathrm{c}}\mu_0)\}^{-1}$. Moreover, if $u \in H^{k+1}(\Omega)$,

$$|||u - u_h|||_{\mathrm{cf}} \leq C_u h^k, \tag{46}$$

with $C_u = C\|u\|_{H^{k+1}(\Omega)}$.

Estimate (46) yields the convergence of the dG approximation for $k \geq 1$. The result is not quasi-optimal, but suboptimal since the L^2-norm of the error should converge with order $(k+1)$ and the boundary contribution with order $(k + 1/2)$ if the exact solution is smooth enough. A sharper estimate is obtained in §3.4 using upwinding.

3.3.3 Numerical fluxes

It is instructive to consider an alternative viewpoint based on numerical fluxes. Because we are working with broken polynomial spaces, the discrete problem (44) admits a local formulation obtained by considering an arbitrary mesh element $T \in \mathcal{T}_h$ and an arbitrary polynomial $\xi \in \mathbb{P}_d^k(T)$. For a set $S \subset \Omega$, we denote by χ_S its characteristic function, namely

$$\chi_S(x) = \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Then, using the test function $v_h = \xi\chi_T$ in the discrete problem (44), observing that

$$[\![\xi\chi_T]\!] = \epsilon_{T,F}\xi \qquad \text{with} \qquad \epsilon_{T,F} := \mathrm{n}_T{\cdot}\mathrm{n}_F,$$

and owing to the expression (43) for the discrete bilinear form a_h^{cf}, we infer

$$\int_T \Big\{(\mu - \nabla{\cdot}\beta)u_h\xi - u_h(\beta{\cdot}\nabla\xi)\Big\} + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h)\xi = \int_T f\xi, \tag{47}$$

where the *numerical fluxes* $\phi_F(u_h)$ are given by

$$\phi_F(u_h) := \begin{cases} (\beta{\cdot}\mathrm{n}_F)\{\!\{u_h\}\!\} & \text{if } F \in \mathcal{F}_h^i, \\ (\beta{\cdot}\mathrm{n})^{\oplus}u_h & \text{if } F \in \mathcal{F}_h^b. \end{cases}$$

The numerical fluxes $\phi_F(u_h)$ are called *centered fluxes* because the average value of u_h is used on each $F \in \mathcal{F}_h^i$. Since these fluxes are *single-valued* and since for all $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$, $\epsilon_{T_1,F} + \epsilon_{T_2,F} = 0$, the local formulation (47) is *conservative* in the sense that whatever "flows" out of a mesh element through one of its faces "flows" into the neighboring element through that face. Finally, taking $\xi \equiv 1$ in (47) leads to the usual balance formulation encountered in finite volume methods, namely

$$\int_T (\mu - \nabla{\cdot}\beta)u_h + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h) = \int_T f.$$

A useful concept in practical implementations is that of stencil.

Definition 3.12 (Stencil)**.** *For a given element $T \in \mathcal{T}_h$, we define the elementary* stencil *$\mathcal{S}(a_h^{\mathrm{cf}};T)$ associated with the bilinear form a_h^{cf} as*

$$\mathcal{S}(a_h^{\mathrm{cf}};T) := \left\{T' \in \mathcal{T}_h \mid \exists q \in \mathbb{P}_d^k(T), \exists r \in \mathbb{P}_d^k(T'), a_h^{\mathrm{cf}}(q\chi_T, r\chi_{T'}) \neq 0\right\},$$

where χ_T and $\chi_{T'}$ denote characteristic functions.

Owing to the local formulation (47), the *stencil* of a given element $T \in \mathcal{T}_h$ consists of T itself and its neighbors in the sense of faces. For instance, on a matching simplicial mesh, the stencil contains $(d+2)$ mesh elements; cf. Figure 7 for a two-dimensional illustration.

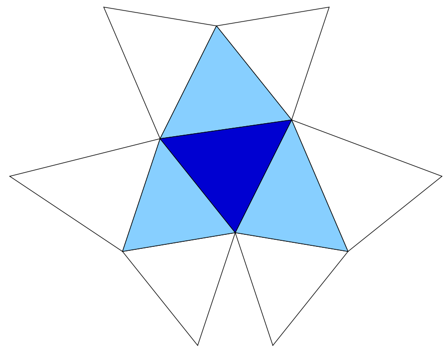

Figure 7: Example of stencil of an element $T \in \mathcal{T}_h$ when $\mathcal{T}_h$ is a matching triangular mesh; the mesh element is highlighted in dark, and its three neighbors, which all belong to the stencil, are highlighted in light; the other triangles do not belong to the stencil

3.4 Upwinding

The goal of this section is to strengthen the stability of the dG bilinear form so as to arrive at quasi-optimal error estimates in the sense of Definition 2.26. This goal is achieved by penalizing in a least-squares sense the interface jumps of the discrete solution. In terms of fluxes, this approach can be interpreted as upwinding. We keep assumptions (27) and (28) on the data μ and β as well as Assumption 3.7 on the regularity of the exact solution u, but the polynomial degree k is here such that $k \geq 0$. For $k = 0$, the dG method considered in this section coincides with a finite volume approximation with upwinding.

The idea of presenting dG methods with upwinding through a suitable penalty of interface jumps has been highlighted recently by Brezzi, Marini, and Süli [20]. Therein, a quasi-optimal error estimate on the L^2-error and the jumps is derived, hinging on discrete coercivity to establish stability. To tighten the error estimate further by including an optimal bound on the advective derivative of the error, a discrete inf-sup condition is needed; this condition, stated in §3.4.2, has been derived by Johnson and Pitkäranta [61].

3.4.1 Tightened stability using penalties

We consider the new bilinear form

$$a_h^{\mathrm{upw}}(v_h, w_h) := a_h^{\mathrm{cf}}(v_h, w_h) + s_h(v_h, w_h), \tag{48}$$

with the stabilization bilinear form

$$s_h(v_h, w_h) = \sum_{F \in \mathcal{F}_h^i} \int_F \frac{\eta}{2} |\beta \cdot \mathrm{n}_F| [\![v_h]\!] [\![w_h]\!], \tag{49}$$

where $\eta > 0$ is a user-dependent parameter. Specifically, using (41),

$$\begin{aligned} a_h^{\mathrm{upw}}(v_h, w_h) := & \int_\Omega \Big\{ \mu v_h w_h + (\beta \cdot \nabla_h v_h) w_h \Big\} + \int_{\partial\Omega} (\beta \cdot \mathrm{n})^{\ominus} v_h w_h \\ & - \sum_{F \in \mathcal{F}_h^i} \int_F (\beta \cdot \mathrm{n}_F) [\![v_h]\!] \{\!\{ w_h \}\!\} + \sum_{F \in \mathcal{F}_h^i} \int_F \frac{\eta}{2} |\beta \cdot \mathrm{n}_F| [\![v_h]\!] [\![w_h]\!], \end{aligned} \tag{50}$$

or, equivalently, using (43),

$$\begin{aligned} a_h^{\mathrm{upw}}(v_h, w_h) = & \int_\Omega \Big\{ (\mu - \nabla \cdot \beta) v_h w_h - v_h (\beta \cdot \nabla_h w_h) \Big\} + \int_{\partial\Omega} (\beta \cdot \mathrm{n})^{\oplus} v_h w_h \\ & + \sum_{F \in \mathcal{F}_h^i} \int_F (\beta \cdot \mathrm{n}_F) \{\!\{ v_h \}\!\} [\![w_h]\!] + \sum_{F \in \mathcal{F}_h^i} \int_F \frac{\eta}{2} |\beta \cdot \mathrm{n}_F| [\![v_h]\!] [\![w_h]\!]. \end{aligned} \tag{51}$$

We observe that the discrete bilinear forms a_h^{cf} and a_h^{upw} lead to the same stencil. The numerical flux associated with the discrete bilinear form a_h^{upw} depends on the penalty parameter η. Choosing $\eta = 1$ is particularly interesting since it leads to the usual upwind fluxes in the context of finite volume schemes. More generally, the discrete bilinear form a_h^{upw} is henceforth referred to as the *upwind dG bilinear form.*

We consider the discrete problem:

$$\text{Find } u_h \in V_h \text{ s.t. } a_h^{\mathrm{upw}}(u_h, v_h) = \int_\Omega f v_h \text{ for all } v_h \in V_h. \tag{52}$$

We first examine the consistency and discrete coercivity of the upwind dG bilinear form. Recalling definition (42) of the discrete coercivity norm $|\!|\!|\cdot|\!|\!|_{\mathrm{cf}}$ considered for centered fluxes, we now assert coercivity with respect to the following stronger norm, also defined on V_{*h}:

$$|\!|\!|v|\!|\!|_{\mathrm{uw}\flat}^2 := |\!|\!|v|\!|\!|_{\mathrm{cf}}^2 + \sum_{F \in \mathcal{F}_h^i} \int_F \frac{\eta}{2} |\beta \cdot \mathrm{n}_F| [\![v]\!]^2. \tag{53}$$

Lemma 3.13 (Consistency and discrete coercivity)**.** *The upwind dG bilinear form a_h^{upw} defined by* (48)–(49)

(i) *is* consistent*, namely for the exact solution $u \in V_*$,*

$$a_h^{\mathrm{upw}}(u, v_h) = \int_\Omega f v_h \qquad \forall v_h \in V_h,$$

(ii) *is* coercive *on V_h with respect to the $|\!|\!|\cdot|\!|\!|_{\mathrm{uw}\flat}$-norm, namely*

$$\forall v_h \in V_h, \qquad a_h^{\mathrm{upw}}(v_h, v_h) \geq C_{\mathrm{sta}} |\!|\!|v_h|\!|\!|_{\mathrm{uw}\flat}^2,$$

with $C_{\mathrm{sta}} = \min(1, \tau_{\mathrm{c}} \mu_0)$ as in Lemma 3.10*.*

The discrete coercivity of a_h^{upw} on V_h implies the well-posedness of the discrete problem (52).

3.4.2 Error estimates based on inf-sup stability

Recalling the definition (53) of the $|\!|\!|\cdot|\!|\!|_{\mathrm{uw}\flat}$-norm, we introduce the stronger norm

$$|\!|\!|v|\!|\!|_{\mathrm{uw}\sharp}^2 := |\!|\!|v|\!|\!|_{\mathrm{uw}\flat}^2 + \sum_{T \in \mathcal{T}_h} \beta_{\mathrm{c}}^{-1} h_T \|\beta \cdot \nabla v\|_{L^2(T)}^2.$$

Lemma 3.14 (Discrete inf-sup condition)**.** *Assume $h \leq \beta_{\mathrm{c}} \tau_{\mathrm{c}}$. There is $C'_{\mathrm{sta}} > 0$, independent of h, μ, and β, such that*

$$\forall v_h \in V_h, \qquad C'_{\mathrm{sta}} C_{\mathrm{sta}} |\!|\!|v_h|\!|\!|_{\mathrm{uw}\sharp} \leq \sup_{w_h \in V_h \setminus \{0\}} \frac{a_h^{\mathrm{upw}}(v_h, w_h)}{|\!|\!|w_h|\!|\!|_{\mathrm{uw}\sharp}},$$

with $C_{\mathrm{sta}} = \min(1, \tau_{\mathrm{c}} \mu_0)$ as in Lemma 3.13*.*

To formulate a boundedness result, we define the following norm:

$$|\!|\!|v|\!|\!|_{\mathrm{uw}\sharp,*}^2 := |\!|\!|v|\!|\!|_{\mathrm{uw}\sharp}^2 + \sum_{T \in \mathcal{T}_h} \beta_{\mathrm{c}} \left(h_T^{-1} \|v\|_{L^2(T)}^2 + \|v\|_{L^2(\partial T)}^2 \right).$$

There holds

$$\forall (v, w_h) \in V_{*h} \times V_h, \qquad |a_h^{\mathrm{upw}}(v, w_h)| \leq C_{\mathrm{bnd}} |\!|\!|v|\!|\!|_{\mathrm{uw}\sharp,*} |\!|\!|w_h|\!|\!|_{\mathrm{uw}\sharp},$$

with C independent of h, μ, and β.

Theorem 3.15 (Error estimate and convergence rate)**.** *Let u solve* (36) *and let u_h solve* (52) *where a_h^{upw} is defined by* (50) *and $V_h = \mathbb{P}_d^k(\mathcal{T}_h)$ with $k \geq 1$ and $\mathcal{T}_h$ belongs to an admissible mesh sequence. Then, there holds*

$$\|u - u_h\|_{\mathrm{uw}\sharp} \leq C \inf_{y_h \in V_h} \|u - y_h\|_{\mathrm{uw}\sharp,*}, \tag{54}$$

with C independent of h and depending on the data only through the factor $\{\min(1, \tau_c\mu_0)\}^{-1}$. *Moreover, if $u \in H^{k+1}(\Omega)$,*

$$\|u - u_h\|_{\mathrm{uw}\sharp} \leq C_u h^{k+1/2}, \tag{55}$$

with $C_u = C\|u\|_{H^{k+1}(\Omega)}$.

Estimate (55) improves estimate (46) by a factor $h^{1/2}$ for the L^2-norm and since it provides a quasi-optimal convergence estimate for the advective derivative.

3.4.3 Numerical fluxes

To conclude this section, we examine how the additional penalty term on the interface jumps modifies the numerical fluxes. Proceeding as in §3.3.3, we obtain the following local formulation: For all $T \in \mathcal{T}_h$ and all $\xi \in \mathbb{P}_d^k(T)$,

$$\int_T \Big\{(\mu - \nabla{\cdot}\beta)u_h\xi - u_h(\beta{\cdot}\nabla\xi)\Big\} + \sum_{F \in \mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F(u_h)\xi = \int_T f\xi, \tag{56}$$

where the numerical fluxes now take the form

$$\phi_F(u_h) = \begin{cases} \beta{\cdot}\mathrm{n}_F \{\!\{u_h\}\!\} + \frac{1}{2}\eta|\beta{\cdot}\mathrm{n}_F|[\![u_h]\!] & \text{if } F \in \mathcal{F}_h^i, \\ (\beta{\cdot}\mathrm{n})^{\oplus} u_h & \text{if } F \in \mathcal{F}_h^b. \end{cases}$$

The choice $\eta = 1$ leads to the so-called *upwind fluxes*

$$\phi_F(u_h) = \begin{cases} \beta{\cdot}\mathrm{n}_F u_h^{\uparrow} & \text{if } F \in \mathcal{F}_h^i, \\ (\beta{\cdot}\mathrm{n})^{\oplus} u_h & \text{if } F \in \mathcal{F}_h^b, \end{cases}$$

where $u_h^{\uparrow} = u_h|_{T_1}$ if $\beta{\cdot}\mathrm{n}_F > 0$ and $u_h^{\uparrow} = u_h|_{T_2}$ otherwise (recall that $F = \partial T_1 \cap \partial T_2$ and that n_F points from T_1 toward T_2). The upwind fluxes can also be written as

$$\phi_F(u_h) = \begin{cases} (\beta{\cdot}\mathrm{n}_F)^{\oplus} u_h|_{T_1} - (\beta{\cdot}\mathrm{n}_F)^{\ominus} u_h|_{T_2} & \text{if } F \in \mathcal{F}_h^i, \\ (\beta{\cdot}\mathrm{n})^{\oplus} u_h & \text{if } F \in \mathcal{F}_h^b. \end{cases}$$

4 Diffusion

We consider the Poisson problem with homogeneous Dirichlet boundary condition

$$-\triangle u = f \quad \text{in } \Omega, \tag{57a}$$
$$u = 0 \quad \text{on } \partial\Omega, \tag{57b}$$

and source term $f \in L^2(\Omega)$.

4.1 The continuous setting

The weak formulation of (57) is classical:

$$\text{Find } u \in V \text{ s.t. } a(u, v) = \int_\Omega fv \text{ for all } v \in V, \tag{58}$$

with energy space $V = H_0^1(\Omega) := \{v \in H^1(\Omega) \mid v|_{\partial\Omega} = 0\}$ and bilinear form

$$a(u, v) := \int_\Omega \nabla u \cdot \nabla v. \tag{59}$$

Recalling the *Poincaré inequality* (see, e.g., Evans [53, p. 265] or Brézis [16, p. 174]) stating that there is C_Ω such that, for all $v \in H_0^1(\Omega)$,

$$\|v\|_{L^2(\Omega)} \leq C_\Omega \|\nabla v\|_{[L^2(\Omega)]^d}, \tag{60}$$

we infer that the bilinear form a is coercive on V. Therefore, owing to the Lax–Milgram Lemma, the weak problem (58) is well-posed.

The PDE (57a) can be rewritten in *mixed form* as a system of first-order PDEs:

$$\sigma + \nabla u = 0 \quad \text{in } \Omega, \tag{61a}$$

$$\nabla \cdot \sigma = f \quad \text{in } \Omega. \tag{61b}$$

Definition 4.1 (Potential and diffusive flux)**.** *In the context of the mixed formulation* (61)*, the scalar-valued function u is termed the* potential *and the vector-valued function $\sigma := -\nabla u$ is termed the* diffusive flux*.*

The derivation of dG methods to approximate the model problems (57) on a given mesh $\mathcal{T}_h$ hinges on the fact that the jumps of the potential and of the normal component of the diffusive flux vanish across interfaces. To allow for a more compact notation, we define boundary averages and jumps.

Definition 4.2 (Boundary averages and jumps)**.** *For a smooth enough function v, for all $F \in \mathcal{F}_h^b$, and for a.e. $x \in F$, we define the* average and jump *of v as*

$$\{\!\{v\}\!\}_F(x) = [\![v]\!]_F(x) := v(x).$$

The subscript as well as the dependence on x are omitted unless necessary.

For simplicity, we enforce a somewhat strong regularity assumption on the exact solution.

Assumption 4.3 (Regularity of exact solution and space V_*)**.** *We assume that the exact solution u is such that*

$$u \in V_* := V \cap H^2(\Omega).$$

In the spirit of §2.5*, we set $V_{*h} := V_* + V_h$.*

Lemma 4.4 (Jumps of potential and diffusive flux)**.** *Assume $u \in V_*$. Then, there holds*

$$[\![u]\!] = 0 \qquad \forall F \in \mathcal{F}_h, \tag{62a}$$

$$[\![\sigma]\!] \cdot \mathrm{n}_F = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{62b}$$

4.2 Symmetric Interior Penalty

Our goal is to approximate the solution of the model problem (58) using dG methods in the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined by (6). We set

$$V_h := \mathbb{P}_d^k(\mathcal{T}_h),$$

with polynomial degree $k \geq 1$ and where $\mathcal{T}_h$ belongs to an admissible mesh sequence. The focus of this section is on a specific dG method, the Symmetric Interior Penalty (SIP) method introduced by Arnold [2].

4.2.1 Heuristic derivation

To derive a suitable discrete bilinear form, we loosely follow the same path of ideas as in Section 3 aiming at a discrete bilinear form that satisfies the consistency requirement (13) and enjoys discrete coercivity. Moreover, we add a (consistent) term to recover, at the discrete level, the symmetry of the continuous problem.

We begin localizing gradients to mesh elements in the exact bilinear form a, that is, we set, for all $v_h, w_h \in V_h$,

$$a_h^{(0)}(v_h, w_h) := \int_\Omega \nabla_h v_h \cdot \nabla_h w_h = \sum_{T \in \mathcal{T}_h} \int_T \nabla v_h \cdot \nabla w_h.$$

To examine the consistency requirement (13), we integrate by parts on each mesh element. This leads to

$$a_h^{(0)}(v_h, w_h) = -\sum_{T \in \mathcal{T}_h} \int_T (\triangle v_h) w_h + \sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla v_h \cdot \mathrm{n}_T) w_h.$$

The second term on the right-hand side can be reformulated as a sum over mesh faces in the form

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla v_h \cdot \mathrm{n}_T) w_h = \sum_{F \in \mathcal{F}_h^i} \int_F [\![(\nabla_h v_h) w_h]\!] \cdot \mathrm{n}_F + \sum_{F \in \mathcal{F}_h^b} \int_F (\nabla v_h \cdot \mathrm{n}_F) w_h,$$

since for all $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$, $\mathrm{n}_F = \mathrm{n}_{T_1} = -\mathrm{n}_{T_2}$. Moreover,

$$[\![(\nabla_h v_h) w_h]\!] = \{\!\{ \nabla_h v_h \}\!\} [\![w_h]\!] + [\![\nabla_h v_h]\!] \{\!\{ w_h \}\!\},$$

since letting $a_i = (\nabla v_h)|_{T_i}$, $b_i = w_h|_{T_i}$, $i \in \{1, 2\}$, yields

$$\begin{aligned} [\![(\nabla_h v_h) w_h]\!] &= a_1 b_1 - a_2 b_2 \\ &= \tfrac{1}{2}(a_1 + a_2)(b_1 - b_2) + (a_1 - a_2)\tfrac{1}{2}(b_1 + b_2) \\ &= \{\!\{ \nabla_h v_h \}\!\} [\![w_h]\!] + [\![\nabla_h v_h]\!] \{\!\{ w_h \}\!\}. \end{aligned}$$

As a result, and accounting for boundary faces using Definition 4.2, yields

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla v_h \cdot \mathrm{n}_T) w_h = \sum_{F \in \mathcal{F}_h} \int_F \{\!\{ \nabla_h v_h \}\!\} \cdot \mathrm{n}_F [\![w_h]\!] + \sum_{F \in \mathcal{F}_h^i} \int_F [\![\nabla_h v_h]\!] \cdot \mathrm{n}_F \{\!\{ w_h \}\!\}.$$

Hence,

$$\begin{aligned} a_h^{(0)}(v_h, w_h) = & -\sum_{T \in \mathcal{T}_h} \int_T (\triangle v_h) w_h + \sum_{F \in \mathcal{F}_h} \int_F \{\!\{ \nabla_h v_h \}\!\} \cdot \mathrm{n}_F [\![w_h]\!] \\ & + \sum_{F \in \mathcal{F}_h^i} \int_F [\![\nabla_h v_h]\!] \cdot \mathrm{n}_F \{\!\{ w_h \}\!\}. \end{aligned} \tag{63}$$

To plug the exact solution u into the above expression, we extend the bilinear form $a_h^{(0)}$ to $V_{*h} \times V_h$ and set $v_h = u$ in (63). A consequence of (62b) is that, for all $w_h \in V_h$,

$$a_h^{(0)}(u, w_h) = \int_\Omega f w_h + \sum_{F \in \mathcal{F}_h} \int_F (\nabla u \cdot \mathrm{n}_F) [\![w_h]\!].$$

In order to match the consistency requirement (13), we are prompted to modify $a_h^{(0)}$ as follows: For all $(v, w_h) \in V_{*h} \times V_h$,

$$a_h^{(1)}(v, w_h) := \int_\Omega \nabla_h v \cdot \nabla_h w_h - \sum_{F \in \mathcal{F}_h} \int_F \{\!\{ \nabla_h v \}\!\} \cdot \mathrm{n}_F [\![w_h]\!].$$

It is clear that $a_h^{(1)}$ is consistent in the sense of (13), i.e., for all $w_h \in V_h$,

$$a_h^{(1)}(u, w_h) = \int_\Omega f w_h.$$

A desirable property of the discrete bilinear form is to preserve the original symmetry of the exact bilinear form. Indeed, symmetry can simplify the solution of the resulting linear system and furthermore, it is a natural ingredient to derive optimal L^2-norm error estimates (cf. §4.2.4). In view of this remark, we set, for all $(v, w_h) \in V_{*h} \times V_h$,

$$a_h^{\text{cs}}(v, w_h) := \int_\Omega \nabla_h v \cdot \nabla_h w_h - \sum_{F \in \mathcal{F}_h} \int_F \left(\{\!\{\nabla_h v\}\!\} \cdot \mathrm{n}_F [\![w_h]\!] + [\![v]\!] \{\!\{\nabla_h w_h\}\!\} \cdot \mathrm{n}_F \right), \tag{64}$$

so that a_h^{cs} is symmetric on $V_h \times V_h$ The bilinear form a_h^{cs} remains consistent owing to (62a). The superscript in a_h^{cs} indicates the consistency and symmetry achieved so far. For future use, we record the following equivalent expression of a_h^{cs} resulting from (63),

$$\begin{aligned} a_h^{\text{cs}}(v, w_h) = & - \sum_{T \in \mathcal{T}_h} \int_T (\triangle v) w_h + \sum_{F \in \mathcal{F}_h^i} \int_F [\![\nabla_h v]\!] \cdot \mathrm{n}_F \{\!\{w_h\}\!\} \\ & - \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \{\!\{\nabla_h w_h\}\!\} \cdot \mathrm{n}_F. \end{aligned} \tag{65}$$

The last requirement to match is discrete coercivity on the broken polynomial space V_h with respect to a suitable norm. The difficulty with the discrete bilinear form a_h^{cs} defined by (64) is that, for all $v_h \in V_h$,

$$a_h^{\text{cs}}(v_h, v_h) = \|\nabla_h v_h\|^2_{[L^2(\Omega)]^d} - 2 \sum_{F \in \mathcal{F}_h} \int_F \{\!\{\nabla_h v_h\}\!\} \cdot \mathrm{n}_F [\![v_h]\!],$$

and the second term on the right-hand side has no *a priori* sign. To achieve discrete coercivity, we add to a_h^{cs} a term penalizing interface and boundary jumps, namely we set, for all $(v, w_h) \in V_{*h} \times V_h$,

$$a_h^{\text{sip}}(v, w_h) := a_h^{\text{cs}}(v, w_h) + s_h(v, w_h), \tag{66}$$

with the stabilization bilinear form

$$s_h(v, w_h) := \sum_{F \in \mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![v]\!] [\![w_h]\!], \tag{67}$$

where $\eta > 0$ is a user-dependent parameter and h_F a local length scale associated with the mesh face $F \in \mathcal{F}_h$. We observe that, owing to (62a), adding the bilinear form s_h to a_h^{cs} does not alter the consistency and symmetry achieved so far. Moreover, Lemma 4.10 below shows that, provided the penalty parameter η is large enough, the discrete bilinear form a_h^{sip} enjoys discrete coercivity on V_h.

We now present a simple choice for the local length scale h_F. Other choices are possible; cf. Remark 4.6.

Definition 4.5 (Local length scale h_F)**.** *For all $F \in \mathcal{F}_h$, in dimension $d \geq 2$, we set h_F to be equal to the diameter of the face F, while, in dimension 1, we set $h_F := \min(h_{T_1}, h_{T_2})$ if $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$ and $h_F := h_T$ if $F \in \mathcal{F}_h^b$ with $F = \partial T \cap \partial \Omega$. In all cases, for a mesh element $T \in \mathcal{T}_h$, h_T denotes its diameter (cf. Definition 2.7).*

Remark 4.6 (Local length scale h_F). Other choices are possible for the local length scale h_F weighting the face penalties in the stabilization bilinear form s_h, e.g., the choice $h_F = \{\!\{h\}\!\} := \frac{1}{2}(h_{T_1} + h_{T_2})$ for all $F \in \mathcal{F}_h^i$, or the choice $h_F = \frac{\{\!\{|T|_d\}\!\}}{|F|_{d-1}}$ (that is, the mean value of the d-dimensional Hausdorff measures of the neighboring elements divided by the $(d-1)$-dimensional Hausdorff measure of the face, recalling that for $d = 1$, $|F|_0 = 1$). Incidentally, we observe that modifying the choice for the local length scale impacts the value of the minimal threshold on the penalty parameter η for which discrete coercivity is achieved.

Combining (66) with (67) yields, for all $(v, w_h) \in V_{*h} \times V_h$,

$$a_h^{\mathrm{sip}}(v, w_h) = \int_\Omega \nabla_h v \cdot \nabla_h w_h - \sum_{F \in \mathcal{F}_h} \int_F (\{\!\{\nabla_h v\}\!\} \cdot \mathrm{n}_F [\![w_h]\!] + [\![v]\!] \{\!\{\nabla_h w_h\}\!\} \cdot \mathrm{n}_F) + \sum_{F \in \mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![v]\!][\![w_h]\!], \tag{68}$$

or, equivalently using (65),

$$a_h^{\mathrm{sip}}(v, w_h) = -\sum_{T \in \mathcal{T}_h} \int_T (\triangle v) w_h + \sum_{F \in \mathcal{F}_h^i} \int_F [\![\nabla_h v]\!] \cdot \mathrm{n}_F \{\!\{w_h\}\!\} - \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \{\!\{\nabla_h w_h\}\!\} \cdot \mathrm{n}_F + \sum_{F \in \mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![v]\!][\![w_h]\!]. \tag{69}$$

Henceforth, a_h^{sip} is called the SIP bilinear form. In the present context, interior penalty means interior as well as boundary penalties.

Definition 4.7 (Consistency, symmetry, and penalty terms)**.** *The second, third, and fourth terms on the right-hand side of* (68) *are respectively called* consistency*,* symmetry*, and* penalty *terms.*

4.2.2 The discrete problem

The discrete problem is

$$\text{Find } u_h \in V_h \text{ s.t. } a_h^{\mathrm{sip}}(u_h, v_h) = \int_\Omega f v_h \text{ for all } v_h \in V_h. \tag{70}$$

Lemma 4.10 below states that provided the penalty parameter η is large enough, the SIP bilinear form is coercive on V_h. Thus, owing to the Lax–Milgram Lemma, the discrete problem (70) is well-posed. Moreover, a straightforward consequence of the above derivation is consistency.

Lemma 4.8 (Consistency)**.** *Assume $u \in V_*$. Then, for all $v_h \in V_h$,*

$$a_h^{\mathrm{sip}}(u, v_h) = \int_\Omega f v_h.$$

Remark 4.9 (Stencil). With an eye toward implementation, we identify the elementary stencil (cf. Definition 3.12) associated with the SIP bilinear form. For all $T \in \mathcal{T}_h$, the stencil of the volume contribution is just the element T, while the stencil associated with the consistency, symmetry, and penalty terms consists of T and its neighbors in the sense of faces. Thus, the elementary stencil is that depicted in Figure 7.

4.2.3 Basic energy-error estimate

Let u solve the weak problem (58) and let u_h solve the discrete problem (70). The aim of this section is to estimate the approximation error $(u - u_h)$. The convergence analysis is performed in the spirit of Theorem 2.20. We recall that the space V_* is specified in Assumption 4.3 and that $V_{*h} = V_* + V_h$.

We aim at asserting discrete coercivity using the following norm: For all $v \in V_{*h}$,

$$|\!|\!| v |\!|\!|_{\mathrm{sip}} := \left(\|\nabla_h v\|^2_{[L^2(\Omega)]^d} + |v|^2_{\mathrm{J}} \right)^{1/2}, \tag{71}$$

with the *jump seminorm*

$$|v|_{\mathrm{J}} := (\eta^{-1} s_h(v, v))^{1/2} = \left(\sum_{F \in \mathcal{F}_h} \frac{1}{h_F} \|[\![v]\!]\|^2_{L^2(F)} \right)^{1/2}. \tag{72}$$

We observe that $|\!|\!|\cdot|\!|\!|_{\mathrm{sip}}$ is indeed a norm on V_{*h}, and even on the broken Sobolev space $H^1(\mathcal{T}_h)$. The only nontrivial property to check is whether, for all $v \in H^1(\mathcal{T}_h)$, $|\!|\!|v|\!|\!|_{\mathrm{sip}} = 0$ implies $v = 0$. Clearly, $|\!|\!|v|\!|\!|_{\mathrm{sip}} = 0$ implies $\|\nabla_h v\|_{[L^2(\Omega)]^d} = 0$ and $|v|_{\mathrm{J}} = 0$. The first property yields $\nabla_h v = 0$ so that v is piecewise constant. The second property implies that the interface and boundary jumps of v vanish. Hence, $v = 0$.

We can now turn to the discrete coercivity of the SIP bilinear form. We recall that N_∂, defined by (2), denotes the maximum number of mesh faces composing the boundary of a generic mesh element and that this quantity is bounded uniformly in h.

Lemma 4.10 (Discrete coercivity). *For all $\eta > \underline{\eta} := C_{\mathrm{tr}}^2 N_\partial$ where C_{tr} results from the discrete trace inequality* (16) *and the parameter N_∂ is defined by* (2)*, the SIP bilinear form defined by* (68) *is coercive on V_h with respect to the $|\!|\!|\cdot|\!|\!|_{\mathrm{sip}}$-norm, i.e.,*

$$\forall v_h \in V_h, \qquad a_h^{\mathrm{sip}}(v_h, v_h) \geq C_\eta |\!|\!|v_h|\!|\!|_{\mathrm{sip}}^2,$$

with $C_\eta := (\eta - C_{\mathrm{tr}}^2 N_\partial)(1+\eta)^{-1}$.

We define on V_{*h} the norm

$$|\!|\!|v|\!|\!|_{\mathrm{sip},*} := \left(|\!|\!|v|\!|\!|_{\mathrm{sip}}^2 + \sum_{T \in \mathcal{T}_h} h_T \|\nabla v|_T \cdot \mathrm{n}_T\|_{L^2(\partial T)}^2 \right)^{1/2}. \tag{73}$$

There is C_{bnd}, independent of h, such that

$$\forall (v, w_h) \in V_{*h} \times V_h, \qquad a_h^{\mathrm{sip}}(v, w_h) \leq C_{\mathrm{bnd}} |\!|\!|v|\!|\!|_{\mathrm{sip},*} |\!|\!|w_h|\!|\!|_{\mathrm{sip}}. \tag{74}$$

Theorem 4.11 ($|\!|\!|\cdot|\!|\!|_{\mathrm{sip}}$-norm error estimate and convergence rate). *Let $u \in V_*$ solve* (58)*. Let u_h solve* (70) *with a_h^{sip} defined by* (68) *and penalty parameter as in Lemma* 4.10*. Then, there is C, independent of h, such that*

$$|\!|\!|u - u_h|\!|\!|_{\mathrm{sip}} \leq C \inf_{v_h \in V_h} |\!|\!|u - v_h|\!|\!|_{\mathrm{sip},*}. \tag{75}$$

Moreover, if $u \in H^{k+1}(\Omega)$,

$$|\!|\!|u - u_h|\!|\!|_{\mathrm{sip}} \leq C_u h^k, \tag{76}$$

with $C_u = C\|u\|_{H^{k+1}(\Omega)}$.

4.2.4 L^2-norm error estimate

To derive an optimal L^2-norm error estimate, it is possible to resort to a duality argument (the so-called Aubin–Nitsche argument [4]) under the following assumption.

Definition 4.12 (Elliptic regularity). *We say that* elliptic regularity *holds true for the model problem* (58) *if there is C_{ell}, only depending on Ω, such that, for all $\psi \in L^2(\Omega)$, the solution to the problem:*

$$\text{Find } \zeta \in H_0^1(\Omega) \text{ s.t. } a(\zeta, v) = \int_\Omega \psi v \text{ for all } v \in H_0^1(\Omega),$$

is in V_* and satisfies

$$\|\zeta\|_{H^2(\Omega)} \leq C_{\mathrm{ell}} \|\psi\|_{L^2(\Omega)}.$$

Elliptic regularity can be asserted if, for instance, the polygonal domain Ω is convex; see Grisvard [57].

Theorem 4.13 (L^2-norm error estimate). *Let $u \in V_*$ solve* (58)*. Let u_h solve* (70) *with a_h^{sip} defined by* (68)*. Assume elliptic regularity. Then, there is C, independent of h, such that*

$$\|u - u_h\|_{L^2(\Omega)} \leq C h |\!|\!|u - u_h|\!|\!|_{\mathrm{sip},*}. \tag{77}$$

Therefore, under the hypotheses of Theorem 4.11 *and if $u \in H^{k+1}(\Omega)$,*

$$\|u - u_h\|_{L^2(\Omega)} \leq C_u h^{k+1}, \tag{78}$$

with $C_u = C\|u\|_{H^{k+1}(\Omega)}$.

Estimate (78) is optimal. We emphasize that the symmetry of a_h^{sip} is used in the proof of Theorem 4.13.

4.3 Liftings and discrete gradients

Liftings are operators that map scalar-valued functions defined on mesh faces to vector-valued functions defined on mesh elements. In the context of dG methods, liftings act on interface and boundary jumps. They were introduced by Bassi, Rebay, Mariotti, Pedinotti, and Savini [10, 11] in the context of compressible flows and analyzed by Brezzi, Manzini, Marini, Pietra, and Russo [18, 19] in the context of the Poisson problem (see also Perugia and Schötzau [73] for the hp-analysis). Liftings have many useful applications. They can be combined with the broken gradient to define discrete gradients. Discrete gradients play an important role in the design and analysis of dG methods. Indeed, they can be used to formulate the discrete problem locally on each mesh element using numerical fluxes. Moreover, they are instrumental in the derivation of discrete functional analysis results, that, in turn, play a central role in the convergence analysis to minimal regularity solutions (see Di Pietro and Ern [46]). Liftings can also be employed to define the stabilization bilinear form [11], yielding a more convenient lower bound for the penalty parameter η.

4.3.1 Main definitions

As before, we assume that the mesh $\mathcal{T}_h$ belongs to an admissible mesh sequence. For any mesh face $F \in \mathcal{F}_h$ and for any integer $l \geq 0$, we define the (local) lifting operator

$$\mathrm{r}_F^l : L^2(F) \longrightarrow [\mathbb{P}_d^l(\mathcal{T}_h)]^d$$

as follows: For all $\varphi \in L^2(F)$,

$$\int_\Omega \mathrm{r}_F^l(\varphi)\cdot\tau_h = \int_F \{\!\{\tau_h\}\!\}\cdot\mathrm{n}_F\varphi \qquad \forall \tau_h \in [\mathbb{P}_d^l(\mathcal{T}_h)]^d. \tag{79}$$

We observe that the support of $\mathrm{r}_F^l(\phi)$ consists of the one or two mesh elements of which F is part of the boundary; using the set $\mathcal{T}_F$ defined by (3) yields

$$\mathrm{supp}(\mathrm{r}_F^l) = \bigcup_{T\in\mathcal{T}_F} \overline{T}. \tag{80}$$

Moreover, whenever the mesh face F is a portion of a hyperplane (this happens, for instance, when working with simplicial meshes or with general meshes consisting of convex elements), $\mathrm{r}_F^l(\varphi)$ is colinear to the normal vector n_F.

For any integer $l \geq 0$ and for any function $v \in H^1(\mathcal{T}_h)$, we define the (global) lifting of its interface and boundary jumps as

$$\mathrm{R}_h^l([\![v]\!]) := \sum_{F\in\mathcal{F}_h} \mathrm{r}_F^l([\![v]\!]) \in [\mathbb{P}_d^l(\mathcal{T}_h)]^d, \tag{81}$$

being implicitly understood that r_F^l acts on the function $[\![v]\!]_F$ (which is in $L^2(F)$ since $v \in H^1(\mathcal{T}_h)$).

For any integer $l \geq 0$, we define the discrete gradient operator

$$G_h^l : H^1(\mathcal{T}_h) \longrightarrow [L^2(\Omega)]^d,$$

as follows: For all $v \in H^1(\mathcal{T}_h)$,

$$G_h^l(v) := \nabla_h v - \mathrm{R}_h^l([\![v]\!]). \tag{82}$$

4.3.2 Reformulation of the SIP bilinear form

Let $l \in \{k-1, k\}$ and set, as in §4.2, $V_h = \mathbb{P}_d^k(\mathcal{T}_h)$ where $k \geq 1$ and $\mathcal{T}_h$ belongs to an admissible mesh sequence. The bilinear form a_h^{cs} can be equivalently written as follows: For all $v_h, w_h \in V_h$,

$$a_h^{\mathrm{cs}}(v_h, w_h) = \int_\Omega \nabla_h v_h \cdot \nabla_h w_h - \int_\Omega \nabla_h v_h \cdot \mathrm{R}_h^l([\![w_h]\!]) - \int_\Omega \nabla_h w_h \cdot \mathrm{R}_h^l([\![v_h]\!]). \tag{83}$$

This results from definitions (79) and (81) and the fact that $\nabla_h v_h$ and $\nabla_h w_h$ are in $[\mathbb{P}_d^l(\mathcal{T}_h)]^d$ since $l \geq k-1$, so that, for all $F \in \mathcal{F}_h$,

$$\int_F \{\!\{\nabla_h v_h\}\!\} \cdot \mathrm{n}_F [\![w_h]\!] = \int_\Omega \nabla_h v_h \cdot \mathrm{r}_F^l([\![w_h]\!]).$$

Starting from (83) and using the definition (82) of the discrete gradient, we infer, for all $v_h, w_h \in V_h$,

$$a_h^{\mathrm{cs}}(v_h, w_h) = \int_\Omega G_h^l(v_h) \cdot G_h^l(w_h) - \int_\Omega \mathrm{R}_h^l([\![v_h]\!]) \cdot \mathrm{R}_h^l([\![w_h]\!]).$$

As a result, recalling that the SIP bilinear form considered in §4.2 is such that $a_h^{\mathrm{sip}} = a_h^{\mathrm{cs}} + s_h$ with s_h defined by (67), we obtain, for all $v_h, w_h \in V_h$,

$$a_h^{\mathrm{sip}}(v_h, w_h) = \int_\Omega G_h^l(v_h) \cdot G_h^l(w_h) + \hat{s}_h^{\mathrm{sip}}(v_h, w_h), \tag{84}$$

with

$$\hat{s}_h^{\mathrm{sip}}(v_h, w_h) := \sum_{F \in \mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![v_h]\!][\![w_h]\!] - \int_\Omega \mathrm{R}_h^l([\![v_h]\!]) \cdot \mathrm{R}_h^l([\![w_h]\!]). \tag{85}$$

The most natural choice for l appears to be $l = k-1$ since the broken gradient is in $[\mathbb{P}_d^{k-1}(\mathcal{T}_h)]^d$. The choice $l = k$ can facilitate the implementation of the method in that it allows one to use the same polynomial basis for computing the liftings and assembling the matrix.

The interest in using discrete gradients to formulate dG methods has been recognized recently in various contexts, e.g., by Lew, Neff, Sulsky, and Ortiz [68] and Ten Eyck and Lew [80] for linear and nonlinear elasticity, Buffa and Ortner [21] and Burman and Ern [22] for nonlinear variational problems, and the authors [46] for the Navier–Stokes equations.

It is interesting to notice that, for all $v_h \in V_h$,

$$a_h^{\mathrm{sip}}(v_h, v_h) \geq \|G_h^l(v_h)\|_{[L^2(\Omega)]^d}^2 + (\eta - C_{\mathrm{tr}}^2 N_\partial)|v_h|_{\mathrm{J}}^2.$$

In view of this result, the expression (84) for a_h^{sip} consists of two terms, both yielding a nonnegative contribution whenever $w_h = v_h$ and, as in Lemma 4.10, $\eta > C_{\mathrm{tr}}^2 N_\partial$. The first term can be seen as the discrete counterpart of the exact bilinear form a (such that $a(v, w) = \int_\Omega \nabla v \cdot \nabla w$) and provides a control on the discrete gradient in $[L^2(\Omega)]^d$. The role of the second term is to strengthen the discrete stability of the method.

Remark 4.14 (Extension to broken Sobolev spaces). We emphasize that the definition (84) of a_h^{sip} is equivalent to (68) only at the discrete level. Differences occur when extending the definitions (68) and (84) to larger spaces, e.g., broken Sobolev spaces. The SIP bilinear form defined by (68) cannot be extended to the minimum regularity space $H^1(\Omega)$ because traces of gradients on mesh faces are used. Instead, the bilinear form defined by (84) can be extended to the broken Sobolev space $H^1(\mathcal{T}_h)$. We denote this extension by $\tilde{a}_h^{\mathrm{sip}}$. Incidentally, $\tilde{a}_h^{\mathrm{sip}}$ is no longer consistent. For convergence analysis to smooth solutions, Strang's First Lemma (see [77] or, e.g., Braess [14, p. 106]) dedicated to nonconsistent finite element methods can be used, whereby the consistency error is estimated for $u \in H^{k+1}(\Omega)$ as follows: For all $v_h \in V_h$,

$$\tilde{a}_h^{\mathrm{sip}}(u - u_h, v_h) = \sum_{F \in \mathcal{F}_h} \int_F \{\!\{\nabla u - \pi_h(\nabla u)\}\!\} \cdot \mathrm{n}_F [\![v_h]\!] \leq C_u h^k |v_h|_{\mathrm{J}},$$

where π_h denotes the L^2-orthogonal projection onto V_h. As a result, the consistency error tends optimally to zero as the meshsize goes to zero.

4.3.3 Numerical fluxes

Discontinuous Galerkin methods can be viewed as high-order finite volume methods. The aim of this section is to identify the local conservation properties associated with dG methods. Such properties are important when the diffusive flux is to be used as an advective velocity in a transport problem, e.g., in the context of coupled porous media flow and contaminant transport.

Let $T \in \mathcal{T}_h$ and let $\xi \in \mathbb{P}_d^k(T)$. Integration by parts shows that, for the exact solution u,

$$\int_T f\xi = -\int_T (\triangle u)\xi = \int_T \nabla u\cdot\nabla\xi - \int_{\partial T} (\nabla u\cdot \mathrm{n}_T)\xi.$$

Therefore, defining on each mesh face $F \in \mathcal{F}_h$ the exact flux as

$$\Phi_F(u) := -\nabla u\cdot \mathrm{n}_F, \tag{86}$$

and recalling the notation $\epsilon_{T,F} = \mathrm{n}_T\cdot\mathrm{n}_F$ introduced in §3.3.3, we infer

$$\int_T \nabla u\cdot\nabla\xi + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F \Phi_F(u)\xi = \int_T f\xi.$$

This is a local conservation property satisfied by the exact solution. Our goal is to identify a similar relation satisfied by the discrete solution u_h solving (70). Using $v_h = \xi\chi_T$ as test function in (70) (where χ_T denotes the characteristic function of T), observing that $\nabla_h(\xi\chi_T) = (\nabla\xi)\chi_T$, and recalling the definition (68) of a_h^{sip}, we obtain

$$\begin{aligned}\int_T f\xi = a_h^{\mathrm{sip}}(u_h, \xi\chi_T) = \int_T \nabla u_h\cdot\nabla\xi &- \sum_{F\in\mathcal{F}_T}\int_F \{\!\{\nabla_h u_h\}\!\}\cdot\mathrm{n}_F [\![\xi\chi_T]\!] \\ &- \sum_{F\in\mathcal{F}_T}\int_F \{\!\{(\nabla\xi)\chi_T\}\!\}\cdot\mathrm{n}_F[\![u_h]\!] + \sum_{F\in\mathcal{F}_T}\frac{\eta}{h_F}\int_F [\![u_h]\!][\![\xi\chi_T]\!].\end{aligned}$$

Let $l \in \{k-1, k\}$. The first and third terms on the right-hand side sum up to $\int_T G_h^{k-1}(u_h)\cdot\nabla\xi$ since $\nabla\xi \in [\mathbb{P}_d^{k-1}(T)]^d$ and $l \geq k-1$, while in the second and fourth terms, we observe that $[\![\xi\chi_T]\!] = \epsilon_{T,F}\xi$. As a result, for all $T \in \mathcal{T}_h$ and all $\xi \in \mathbb{P}_d^k(T)$,

$$\int_T G_h^l(u_h)\cdot\nabla\xi + \sum_{F\in\mathcal{F}_T}\epsilon_{T,F}\int_F \phi_F(u_h)\xi = \int_T f\xi, \tag{87}$$

with the numerical flux $\phi_F(u_h)$ defined as

$$\phi_F(u_h) := -\{\!\{\nabla_h u_h\}\!\}\cdot\mathrm{n}_F + \frac{\eta}{h_F}[\![u_h]\!]. \tag{88}$$

We notice that the two contributions to $\phi_F(u_h)$ in (88) respectively stem from the consistency term and the penalty term (cf. Definition 4.7). Equation (87) is the local conservation property satisfied by the dG approximation. Interestingly, the expression (88) is consistent with (86) since, for the exact solution u, $\phi_F(u) = \Phi_F(u)$. We also observe that the local conservation property (87) is richer than that encountered in finite volume methods, which can be recovered by just taking $\xi \equiv 1$, i.e.,

$$\sum_{F\in\mathcal{F}_T}\epsilon_{T,F}\int_F \phi_F(u_h) = \int_T f. \tag{89}$$

4.4 Mixed dG methods

In this section, we discuss mixed dG methods, that is, dG approximations to the mixed formulation (61) with the homogeneous Dirichlet boundary condition (57b). Such methods produce an approximation u_h for the potential u and an approximation σ_h for the diffusive flux σ.

Definition 4.15 (Discrete potential and discrete diffusive flux)**.** *The scalar-valued function* u_h *is termed the* discrete potential *and the vector-valued function* σ_h *the* discrete diffusive flux.

First, we reformulate the SIP method of §4.2 as a mixed dG method and show how the discrete diffusive flux can be eliminated locally. Then, we formulate more general mixed dG methods in terms of local problems using numerical fluxes for the discrete potential and the diffusive flux following Bassi, Rebay and coworkers [11, 10]. This leads, in particular, to the LDG methods introduced by Cockburn and Shu [43]. In these methods, the discrete diffusive flux can also be eliminated locally. Finally, we discuss hybrid mixed dG methods methods where additional degrees of freedom are introduced at interfaces, thereby allowing one to eliminate locally both the discrete potential and the discrete diffusive flux.

4.4.1 The SIP method as a mixed dG method

One possible weak formulation of the mixed formulation (61) with the homogeneous Dirichlet boundary condition (57b) consists in finding $(\sigma, u) \in X := [L^2(\Omega)]^d \times H^1_0(\Omega)$ such that

$$\begin{cases} m(\sigma,\tau) + b(\tau,u) = 0 & \forall \tau \in [L^2(\Omega)]^d, \\ -b(\sigma,v) = \int_\Omega fv & \forall v \in H^1_0(\Omega), \end{cases} \tag{90}$$

where, for all σ, $\tau \in [L^2(\Omega)]^d$ and for all $v \in H^1_0(\Omega)$, we have defined the bilinear forms

$$m(\sigma,\tau) := \int_\Omega \sigma{\cdot}\tau, \qquad b(\tau,v) := \int_\Omega \tau{\cdot}\nabla v.$$

It is easily seen that $(\sigma, u) \in X$ solves (90) if and only if $\sigma = -\nabla u$ and u solves the weak problem (58).

At the discrete level, a mixed dG approximation can be designed as follows. We consider a polynomial degree $k \geq 1$ for the approximation of the potential and choose the polynomial degree for the approximation of the diffusive flux, say l, such that $l \in \{k-1, k\}$. The relevant discrete spaces are

$$\Sigma_h := [\mathbb{P}^l_d(\mathcal{T}_h)]^d, \qquad U_h := \mathbb{P}^k_d(\mathcal{T}_h), \qquad X_h := \Sigma_h \times U_h.$$

The discrete problem consists in finding $(\sigma_h, u_h) \in X_h$ such that

$$\begin{cases} m(\sigma_h,\tau_h) + b_h(\tau_h,u_h) = 0 & \forall \tau_h \in \Sigma_h, \\ -b_h(\sigma_h,v_h) + \hat{s}^{\text{sip}}_h(u_h,v_h) = \int_\Omega fv_h & \forall v_h \in U_h, \end{cases} \tag{91}$$

with discrete bilinear form

$$b_h(\tau_h, v_h) := \int_\Omega \tau_h{\cdot}G^l_h(v_h),$$

where the discrete gradient operator G^l_h is defined by (82) and the stabilization bilinear form $\hat{s}^{\text{sip}}_h$ by (85).

Proposition 4.16 (Elimination of discrete diffusive flux)**.** *The pair* $(\sigma_h, u_h) \in X_h$ *solves* (91) *if and only if*

$$\sigma_h = -G^l_h(u_h), \tag{92}$$

and $u_h \in U_h$ *is such that*

$$\int_\Omega G^l_h(u_h){\cdot}G^l_h(v_h) + \hat{s}^{\text{sip}}_h(u_h,v_h) = \int_\Omega fv_h \qquad \forall v_h \in U_h. \tag{93}$$

Proposition 4.16 shows that the mixed dG method (91) is in fact equivalent to a problem in the sole unknown u_h. In particular, the above choice for b_h and $\hat{s}^{\text{sip}}_h$ yields the SIP method of §4.2.

4.4.2 Numerical fluxes

We focus for simplicity on equal-order approximations for the potential and the diffusive flux, that is, we set $l = k$ so that $\Sigma_h := [\mathbb{P}^k_d(\mathcal{T}_h)]^d$, while, as before, $U_h := \mathbb{P}^k_d(\mathcal{T}_h)$. Similarly to §4.3.3, we can derive a local formulation by localizing test functions to a single mesh element. Let $T \in \mathcal{T}_h$, let $\zeta \in [\mathbb{P}^k_d(T)]^d$, and let $\xi \in \mathbb{P}^k_d(T)$. Integrating by parts in T, splitting the boundary integral on ∂T as a sum over the mesh faces $F \in \mathcal{F}_T$, and setting $\epsilon_{T,F} = \mathrm{n}_T{\cdot}\mathrm{n}_F$, we infer for the exact solution that

$$\int_T \sigma{\cdot}\zeta - \int_T u\nabla{\cdot}\zeta + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F u_F(\zeta{\cdot}\mathrm{n}_F) = 0,$$
$$-\int_T \sigma{\cdot}\nabla\xi + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F (\sigma_F{\cdot}\mathrm{n}_F)\xi = \int_T f\xi,$$

since $\sigma = -\nabla u$ and $\nabla{\cdot}\sigma = f$. The traces u_F and $\sigma_F{\cdot}\mathrm{n}_F$ are single-valued on each interface; cf. Lemma 4.4. At the discrete level, the general form of the mixed dG approximation is derived by introducing numerical fluxes for the discrete potential and for the discrete diffusive flux. These two numerical fluxes, which are denoted by $\hat{u}_F$ and $\hat{\sigma}_F$ for all $F \in \mathcal{F}_h$, are single-valued on each $F \in \mathcal{F}_h$. The numerical flux $\hat{u}_F$ is scalar-valued and the numerical flux $\hat{\sigma}_F$ is vector-valued. We obtain, for all $T \in \mathcal{T}_h$, all $\zeta \in [\mathbb{P}^k_d(T)]^d$, and all $\xi \in \mathbb{P}^k_d(T)$,

$$\int_T \sigma_h{\cdot}\zeta - \int_T u_h\nabla{\cdot}\zeta + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F \hat{u}_F(\zeta{\cdot}\mathrm{n}_F) = 0, \tag{94a}$$
$$-\int_T \sigma_h{\cdot}\nabla\xi + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F (\hat{\sigma}_F{\cdot}\mathrm{n}_F)\xi = \int_T f\xi. \tag{94b}$$

For the SIP method, the numerical fluxes are given by

$$\hat{u}_F = \begin{cases} \{\!\{u_h\}\!\} & \forall F \in \mathcal{F}^i_h, \\ 0 & \forall F \in \mathcal{F}^b_h, \end{cases} \tag{95a}$$
$$\hat{\sigma}_F = -\{\!\{\nabla_h u_h\}\!\} + \eta h_F^{-1}[\![u_h]\!]\mathrm{n}_F \quad \forall F \in \mathcal{F}_h. \tag{95b}$$

A first possible variant of the SIP method consists in keeping the definition (95a) for the numerical flux $\hat{u}_F$ and defining the numerical flux $\hat{\sigma}_F$ as

$$\hat{\sigma}_F = \{\!\{\sigma_h\}\!\} + \eta h_F^{-1}[\![u_h]\!]\mathrm{n}_F.$$

The resulting dG method belongs to the class of LDG methods. The discrete diffusive flux σ_h can still be eliminated locally (since the numerical flux $\hat{u}_F$ only depends on u_h), and the discrete potential $u_h \in U_h$ is such that

$$a_h^{\mathrm{ldg}}(u_h, v_h) = \int_\Omega f v_h \qquad \forall v_h \in U_h,$$

with the discrete bilinear form

$$\begin{aligned} a_h^{\mathrm{ldg}}(u_h, v_h) &= \int_\Omega \nabla_h u_h{\cdot}\nabla_h v_h - \sum_{F\in\mathcal{F}_h} \int_F (\{\!\{\nabla_h u_h\}\!\}{\cdot}\mathrm{n}_F[\![v_h]\!] + \{\!\{\nabla_h v_h\}\!\}{\cdot}\mathrm{n}_F[\![u_h]\!]) \\ &\quad + \int_\Omega \mathrm{R}^k_h([\![u_h]\!]){\cdot}\,\mathrm{R}^k_h([\![v_h]\!]) + \sum_{F\in\mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![u_h]\!][\![v_h]\!] \\ &= \int_\Omega G^k_h(u_h){\cdot}G^k_h(v_h) + \sum_{F\in\mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![u_h]\!][\![v_h]\!]. \end{aligned}$$

A nice feature of the discrete bilinear form a_h^{ldg} is that discrete coercivity holds on U_h with respect to the $|\!|\!|{\cdot}|\!|\!|_{\mathrm{sip}}$-norm for any $\eta > 0$ (a simple choice is $\eta = 1$). The drawback is that the elementary *stencil* associated with the term $\int_\Omega \mathrm{R}_h^k([\![u_h]\!])\cdot \mathrm{R}_h^k([\![v_h]\!])$ consists of a given mesh element, its neighbors, and the neighbors of its neighbors in the sense of faces; cf. Figure 8. Such a stencil is considerably larger than that associated with the SIP method (compare with Figure 7).

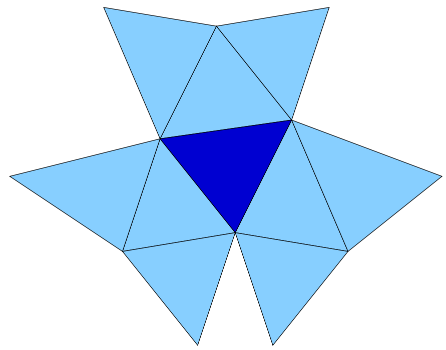

Figure 8: Example of LDG stencil of an element $T \in \mathcal{T}_h$ when $\mathcal{T}_h$ is a matching triangular mesh; the mesh element is highlighted in dark, and all the nine other elements, highlighted in light, also belong to the stencil

More general forms of the LDG method can be designed with the numerical fluxes

$$\hat{u}_F = \begin{cases} \{\!\{u_h\}\!\} + \Upsilon\cdot \mathrm{n}_F [\![u_h]\!] & \forall F \in \mathcal{F}_h^i, \\ 0 & \forall F \in \mathcal{F}_h^b, \end{cases}$$

$$\hat{\sigma}_F = \begin{cases} \{\!\{\sigma_h\}\!\} - \Upsilon [\![\sigma_h]\!]\cdot \mathrm{n}_F + \eta h_F^{-1}[\![u_h]\!]\mathrm{n}_F & \forall F \in \mathcal{F}_h^i, \\ \sigma_h + \eta h_F^{-1} u_h \mathrm{n} & \forall F \in \mathcal{F}_h^b, \end{cases}$$

where Υ is vector-valued and $\eta > 0$ is scalar-valued (in LDG methods, ηh_F^{-1} is often denoted by C_{11} and Υ by C_{12}). Since the numerical flux $\hat{u}_F$ only depends on u_h, the discrete diffusive flux σ_h can be eliminated locally. The above form of the diffusive fluxes ensures symmetry and discrete stability for the resulting dG method. A simple choice for the penalty parameter is again $\eta = 1$, while the auxiliary vector-parameter Υ can be freely chosen. LDG methods for the Poisson problem have been extensively analyzed by Castillo, Cockburn, Perugia, and Schötzau [24]. Variants of the LDG method aiming at reducing the stencil have been discussed by Sherwin, Kirby, Peiró, Taylor, and Zienkiewicz [76], Peraire and Persson [72], and Castillo [25].

A further variant of the SIP and LDG methods consists in considering the numerical fluxes

$$\hat{u}_F = \begin{cases} \{\!\{u_h\}\!\} + \eta_\sigma [\![\sigma_h]\!]\cdot \mathrm{n}_F & \forall F \in \mathcal{F}_h^i, \\ 0 & \forall F \in \mathcal{F}_h^b, \end{cases}$$

$$\hat{\sigma}_F = \{\!\{\sigma_h\}\!\} + \eta_u [\![u_h]\!]\mathrm{n}_F \quad \forall F \in \mathcal{F}_h.$$

Here, the penalty parameters η_u and η_σ are positive user-dependent real numbers, and a simple choice is to set $\eta_u = \eta_\sigma = 1$. This method can be analyzed in the more general context of Friedrichs' systems (see Ern and Guermond [50]). Because the numerical flux $\hat{u}_F$ depends on σ_h, (94a) can no longer be used to express locally the discrete diffusive flux σ_h in terms of the discrete potential u_h. This precludes the local elimination of σ_h and, therefore, enhances the computational cost of the approximation method. The approach presents, however, some advantages since it can be used with polynomial degree $k = 0$ and there is no minimal threshold on the penalty parameters (apart from being positive). Moreover, the approximation on the diffusive flux is more accurate yielding convergence rates in the L^2-norm of order $h^{k+1/2}$ for smooth solutions, as opposed to the convergence rates of order h^k delivered by the SIP method.

Finally, we mention that an even more general presentation can allow for two-valued numerical fluxes at interfaces; see Arnold, Brezzi, Cockburn, and Marini [3] for a unified analysis of dG methods.

4.4.3 Hybrid mixed dG methods

The key idea in hybrid mixed dG methods is to introduce additional degrees of freedom at interfaces, thereby allowing one to eliminate locally both the discrete potential and the discrete diffusive flux. Herein, we focus on the HDG methods introduced by Cockburn, Gopalakrishnan, and Lazarov [31]; see also Causin and Sacco [27] for a different approach based on a discontinuous Petrov–Galerkin formulation, Droniou and Eymard [48] for similar ideas in the context of hybrid mixed finite volume schemes, and Ewing, Wang, and Yang for hybrid primal dG methods [54].

In the HDG method, the additional degrees of freedom are used to enforce the continuity of the normal component of the discrete diffusive flux. These additional degrees of freedom act as Lagrange multipliers in the discrete problem and can be interpreted as single-valued traces of the discrete potential on interfaces. We introduce the discrete space

$$\Lambda_h := \bigoplus_{F\in\mathcal{F}_h^i} \mathbb{P}_{d-1}^k(F).$$

A function $\mu_h \in \Lambda_h$ is such that, for all $F \in \mathcal{F}_h^i$, $\mu_h|_F \in \mathbb{P}_{d-1}^k(F)$. The discrete unknowns $(\sigma_h, u_h, \lambda_h) \in \Sigma_h \times U_h \times \Lambda_h$ satisfy the following local problems: For all $T \in \mathcal{T}_h$, all $\zeta \in [\mathbb{P}_d^k(T)]^d$, and all $\xi \in \mathbb{P}_d^k(T)$,

$$\int_T \sigma_h\cdot\zeta - \int_T u_h \nabla\cdot\zeta + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F \hat{u}_F(\zeta\cdot\mathrm{n}_F) = 0, \tag{96a}$$

$$-\int_T \sigma_h\cdot\nabla\xi + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F (\hat{\sigma}_{T,F}\cdot\mathrm{n}_F)\xi = \int_T f\xi, \tag{96b}$$

while normal diffusive flux continuity is enforced by setting, for all $F \in \mathcal{F}_T \cap \mathcal{F}_h^i$ and all $\mu \in \mathbb{P}_{d-1}^k(F)$,

$$\int_F [\![\hat{\sigma}_{T,F}]\!]\cdot\mathrm{n}_F \mu = 0. \tag{97}$$

Here, the numerical fluxes are such that

$$\hat{u}_F = \begin{cases} \lambda_h & \forall F \in \mathcal{F}_h^i, \\ 0 & \forall F \in \mathcal{F}_h^b, \end{cases} \tag{98a}$$

$$\hat{\sigma}_{T,F} = \sigma_h|_T + \tau_T(u_h|_T - \hat{u}_F)\mathrm{n}_T \quad \forall F \in \mathcal{F}_h, \tag{98b}$$

with penalty parameter τ_T defined elementwise. We observe that (97) indeed enforces $[\![\hat{\sigma}_{T,F}]\!]\cdot\mathrm{n}_F = 0$ for all $F \in \mathcal{F}_h^i$ since $[\![\hat{\sigma}_{T,F}]\!]\cdot\mathrm{n}_F \in \mathbb{P}_{d-1}^k(F)$. As a result, the quantity $(\hat{\sigma}_{T,F}\cdot\mathrm{n}_F)$ in (96b) is indeed single-valued.

Lemma 4.17 (HDG as mixed dG method)**.** *Let $(\sigma_h, u_h, \lambda_h) \in \Sigma_h \times U_h \times \Lambda_h$ solve* (96)–(97)*. Then, the pair $(\sigma_h, u_h) \in \Sigma_h \times U_h$ solves the local problems of the mixed dG formulation* (94) *with numerical fluxes such that, for all $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$,*

$$\hat{u}_F = \{\!\{u_h\}\!\} + C_{12}\cdot[\![u_h]\!]\mathrm{n}_F + C_{22}[\![\sigma_h]\!]\cdot\mathrm{n}_F, \tag{99a}$$

$$\hat{\sigma}_F = \{\!\{\sigma_h\}\!\} + C_{11}[\![u_h]\!]\mathrm{n}_F - C_{12}[\![\sigma_h]\!]\cdot\mathrm{n}_F, \tag{99b}$$

with the parameters

$$C_{11} = \frac{\tau_1\tau_2}{\tau_1+\tau_2}, \qquad C_{12} = \frac{\tau_1-\tau_2}{2(\tau_1+\tau_2)}\mathrm{n}_F, \qquad C_{22} = \frac{1}{\tau_1+\tau_2},$$

where $\tau_i := \tau_{T_i}$, $i \in \{1,2\}$. Moreover, for all $F \in \mathcal{F}_h^b$ with $F = \partial T \cap \partial\Omega$, $\hat{u}_F = 0$ and $\hat{\sigma}_F = \sigma_h + \tau_T u_h$.

We observe that the numerical flux $\hat{u}_F$ in (99a) depends on σ_h since $C_{22} \neq 0$. As a result, the discrete diffusive flux cannot be eliminated locally to derive a discrete problem for the sole discrete potential. Instead, a computationally efficient implementation of HDG methods consists in using (96) to eliminate locally both the discrete potential and the discrete diffusive flux, so as to obtain, using (97), a discrete problem where the sole unknown is $\lambda_h \in \Lambda_h$. For a given interface $F \in \mathcal{F}_h^i$ with $F = \partial T_1 \cap \partial T_2$, the stencil associated with this interface is

$$\mathcal{S}(F) = \{F' \in \mathcal{F}_h^i \mid F' \in \mathcal{F}_{T_1} \cup \mathcal{F}_{T_2}\}.$$

For matching simplicial meshes, the set $\mathcal{S}(F)$ generally contains 5 interfaces for $d = 2$ and 7 interfaces for $d = 3$.

HDG methods for elliptic problems have been analyzed by Cockburn, Dong, and Guzmán [30] and Cockburn, Guzmán, and Wang [32] where error estimates in various norms are derived for various choices of the penalty parameter τ. In particular, L^2-norm error estimates of order h^{k+1} can be derived both for the potential and the diffusive flux for smooth solutions and polynomial order $k \geq 0$. Moreover, for $k \geq 1$, a postprocessed potential converging at order h^{k+2} can be derived, similarly to classical mixed finite element methods.

5 Incompressible flows

The equations governing fluid motion are the Navier–Stokes equations, which express the fundamental laws of mass and momentum conservation. In their general form, these equations were first derived by Navier (1827) and Poisson (1831), while a more specific derivation was found by Saint-Venant (1843) and Stokes (1845) based on the assumption that the stresses are linear functions of the strain rates (or deformation velocities), that is, for Newtonian fluids. In this chapter, we are concerned with the special case of incompressible (that is, constant density) Newtonian flows, thereby leading to the so-called Incompressible Navier–Stokes (INS) equations. In these equations, the dependent variables are the velocity and the pressure. The mass conservation equation enforces zero divergence on the velocity field (because of incompressibility), while the momentum conservation equation expresses the balance between diffusion (due to viscosity), nonlinear convection, pressure gradient, and external forcings.

The main difficulties in the discretization of the steady INS equations are (i) the zero-divergence constraint on the velocity and (ii) the contribution of the nonlinear convection term to the kinetic energy balance. The first issue is addressed in §5.1 in the simpler context of the steady Stokes equations. In §5.2, we turn to the steady INS equations. The central issue is now the discretization of the nonlinear convection term. An important ingredient is to mimic the fact that, at the continuous level, this term does not contribute to the kinetic energy balance.

5.1 Steady Stokes flows

In this section, we consider the steady Stokes equations. These equations describe incompressible viscous flows under the assumption that the fluid motion is sufficiently slow so that diffusion dominates over convection in the transport of momentum.

5.1.1 The continuous setting

Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a polyhedron. The steady Stokes equations can be expressed in the form

$$-\triangle u + \nabla p = f \quad \text{in } \Omega, \tag{100a}$$

$$\nabla \cdot u = 0 \quad \text{in } \Omega, \tag{100b}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{100c}$$

$$\langle p \rangle_\Omega = 0, \tag{100d}$$

where $u : \Omega \to \mathbb{R}^d$ with Cartesian components $(u_i)_{1 \leq i \leq d}$ is the velocity field, $p : \Omega \to \mathbb{R}$ the pressure, and $f : \Omega \to \mathbb{R}^d$ with Cartesian components $(f_i)_{1 \leq i \leq d}$ the forcing term. Equation (100a) expresses the conservation of momentum. Equation (100b) expresses the conservation of mass, thereby enforcing the divergence-free constraint on the velocity. Equation (100c) enforces a homogeneous Dirichlet boundary condition on the velocity; other boundary conditions can be considered, as discussed, e.g., by Ern and Guermond [49, p. 179]. Finally, condition (100d), where $\langle\cdot\rangle_\Omega$ denotes the mean value over Ω, is added to avoid leaving the pressure undetermined up to an additive constant.

Remark 5.1 (Stress and strain tensors, viscosity). A more general form of the momentum conservation equation (100a) takes the form

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

where $\sigma : \Omega \to \mathbb{R}^{d,d}$ is the *stress tensor*. In Newtonian flows, stresses are proportional to strain rates. More specifically, introducing for a given velocity field u the (linearized) *strain tensor* $\varepsilon : \Omega \to \mathbb{R}^{d,d}$ such that $\varepsilon = \frac{1}{2}(\nabla u + \nabla u^t)$, there holds

$$\sigma = 2\nu\varepsilon,$$

where $\nu > 0$ is the (kinematic) *viscosity*. Taking the viscosity constant for simplicity, we obtain

$$-\nu\nabla\cdot(\nabla u + \nabla u^t) + \nabla p = f, \tag{101}$$

and up to rescaling of the pressure and the source term, we can assume that $\nu = 1$. Then, observing that $\nabla\cdot(\nabla u) = \triangle u$ and $\nabla\cdot(\nabla u)^t = \nabla(\nabla\cdot u) = 0$ because of incompressibility, we recover (100a). Considering the form (101) of the momentum conservation equation is appropriate when dealing with other boundary conditions than (100c), e.g., when weakly enforcing the Navier slip boundary condition $(\sigma\cdot\mathrm{n} + \lambda u)\cdot\mathrm{t} = 0$ where t is a tangent vector to the boundary $\partial\Omega$ and $\lambda \geq 0$ a given parameter.

We assume that the forcing term f is in $[L^2(\Omega)]^d$. Owing to (100c), the natural space for the velocity is $[H_0^1(\Omega)]^d$, while owing to (100d), the natural space for the pressure is $L_0^2(\Omega) \subset L^2(\Omega)$ where

$$L_0^2(\Omega) := \left\{ q \in L^2(\Omega) \mid \langle q\rangle_\Omega = 0 \right\}.$$

We set

$$U := [H_0^1(\Omega)]^d, \qquad P := L_0^2(\Omega), \qquad X := U \times P. \tag{102}$$

The spaces U, P, and X are Hilbert spaces when equipped with the inner products inducing the norms

$$\|v\|_U := \|v\|_{[H^1(\Omega)]^d} = \left(\sum_{i=1}^d \|v_i\|_{H^1(\Omega)}^2 \right)^{1/2},$$
$$\|q\|_P := \|q\|_{L^2(\Omega)}, \quad \|(v,q)\|_X := \left(\|v\|_U^2 + \|q\|_P^2 \right)^{1/2}.$$

We define, for all $u, v \in U$ and for all $q \in P$, the bilinear forms

$$a(u,v) := \int_\Omega \nabla u{:}\nabla v = \sum_{i,j=1}^d \int_\Omega \partial_j u_i \, \partial_j v_i = (\nabla u, \nabla v)_{[L^2(\Omega)]^{d,d}}, \tag{103a}$$

$$b(v,q) := -\int_\Omega q\nabla\cdot v = -(\nabla\cdot v, q)_P. \tag{103b}$$

The weak formulation of problem (100) reads: Find $(u,p) \in X$ such that

$$a(u,v) + b(v,p) = \int_\Omega f\cdot v \quad \forall v \in U, \tag{104a}$$

$$-b(u,q) = 0 \qquad \forall q \in P, \tag{104b}$$

or, equivalently,

$$\text{Find } (u,p) \in X \text{ s.t. } c((u,p),(v,q)) = \int_\Omega f \cdot v \text{ for all } (v,q) \in X,$$

with

$$c((u,p),(v,q)) := a(u,v) + b(v,p) - b(u,q).$$

While the bilinear form c is clearly not coercive on X, we observe that the bilinear form a is coercive on U. Indeed, applying the continuous Poincaré inequality (60) to each velocity component, we infer that there exists $\alpha_\Omega > 0$, only depending on Ω, such that

$$\forall v \in U, \qquad a(v,v) = \|\nabla v\|^2_{[L^2(\Omega)]^{d,d}} \geq \alpha_\Omega \|v\|^2_U. \tag{105}$$

This yields a so-called *partial coercivity* for the bilinear form c in the form

$$\forall (v,q) \in X, \qquad c((v,q),(v,q)) = a(v,v) \geq \alpha_\Omega \|v\|^2_U. \tag{106}$$

Remark 5.2 (Saddle-point problem). A problem of the form (104) is said to have a *saddle-point* structure since $(u,p) \in X$ solves (104) if and only if (u,p) is a saddle-point of the Lagrangian $\mathcal{L} : X \to \mathbb{R}$ such that, for all $(v,q) \in X$,

$$\mathcal{L}(v,q) = \frac{1}{2} a(v,v) + b(v,q).$$

In this context, the pressure plays the role of the Lagrange multiplier associated with the incompressibility constraint.

We introduce the divergence operator $B \in \mathcal{L}(U,P)$ such that

$$B : U \ni v \longmapsto Bv := -\nabla \cdot v \in P. \tag{107}$$

(The fact that Bv has zero mean is a consequence of the divergence theorem since $\int_\Omega Bv = -\int_\Omega \nabla \cdot v = -\int_{\partial\Omega} (v \cdot \mathrm{n}) = 0$.) The operator B is readily linked to the bilinear form b since there holds

$$(Bv, q)_P = b(v,q) \qquad \forall (v,q) \in X.$$

The well-posedness of the Stokes problem (104) hinges on the surjectivity of the operator B or, equivalently, on an inf-sup condition on the bilinear form b (se, e.g., Girault and Raviart [55, §2.2]).

Theorem 5.3 (Surjectivity of divergence operator, inf-sup condition on b)**.** *Let $\Omega \in \mathbb{R}^d$, $d \geq 1$, be a connected domain. Then, the operator B is* surjective*. Equivalently, there exists a real number $\beta_\Omega > 0$, only depending on Ω, such that, for all $q \in P$, there is $v_q \in U$ satisfying*

$$q = -Bv_q = \nabla \cdot v_q \qquad \text{and} \qquad \beta_\Omega \|v_q\|_U \leq \|q\|_P. \tag{108}$$

Moreover, property (108) *is equivalent to the following* inf-sup condition *on the bilinear form b:*

$$\forall q \in P, \qquad \beta_\Omega \|q\|_P \leq \sup_{w \in U \setminus \{0\}} \frac{b(w,q)}{\|w\|_U}. \tag{109}$$

For all $q \in P$, a field $v_q \in U$ satisfying (108) *is called a* velocity lifting *of q.*

Theorem 5.4 (Well-posedness)**.** *Problem* (104) *is well-posed.*

5.1.2 Equal-order discontinuous velocities and pressures

In this section, we consider one possible dG discretization of the steady Stokes equations based on equal-order discontinuous velocities and pressures. Other approaches are discussed in §5.1.3. DG methods based on equal-order discontinuous velocities and pressures have been introduced by Cockburn, Kanschat, Schötzau, and Schwab [38] for the Stokes equations and extended to the Oseen equations in [34] and to the INS equations in [37].

Let $\mathcal{T}_h$ be a mesh of Ω belonging to an admissible mesh sequence with mesh regularity parameters denoted by ϱ. Recalling the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined by (6) with polynomial degree $k \geq 1$, we define the discrete spaces

$$U_h := [\mathbb{P}_d^k(\mathcal{T}_h)]^d, \qquad P_h := \mathbb{P}_{d,0}^k(\mathcal{T}_h), \qquad X_h := U_h \times P_h, \tag{110}$$

where $\mathbb{P}_{d,0}^k(\mathcal{T}_h)$ denotes the subspace of $\mathbb{P}_d^k(\mathcal{T}_h)$ spanned by functions having zero mean-value over Ω. The discrete solution is sought in the space X_h.

To discretize the diffusion term, we use, for each velocity component, the SIP bilinear form (cf. §4.2). We define on $U_h \times U_h$ the bilinear form

$$a_h(v_h, w_h) := \sum_{i=1}^{d} a_h^{\text{sip}}(v_{h,i}, w_{h,i}), \tag{111}$$

where $(v_{h,i})_{1 \leq i \leq d}$ and $(w_{h,i})_{1 \leq i \leq d}$ denote the Cartesian components of v_h and w_h, respectively, and where a_h^{sip} is defined by (68). It is natural to equip the discrete velocity space U_h with the $|||\cdot|||_{\text{sip}}$-norm defined by (71) for each Cartesian component, so that we set

$$|||v_h|||_{\text{vel}} := \left(\sum_{i=1}^{d} |||v_{h,i}|||_{\text{sip}}^2 \right)^{1/2} = \left(\|\nabla_h v_h\|_{[L^2(\Omega)]^{d,d}}^2 + |v_h|_{\text{J}}^2 \right)^{1/2}, \tag{112}$$

with the $|\cdot|_{\text{J}}$-seminorm acting now on vector-valued arguments as

$$|v_h|_{\text{J}} = \left(\sum_{F \in \mathcal{F}_h} h_F^{-1} \|[\![v_h]\!]\|_{[L^2(F)]^d}^2 \right)^{1/2}.$$

We assume that the penalty parameter η is such that $\eta > \underline{\eta}$ so that

$$\forall v_h \in U_h, \qquad a_h(v_h, v_h) \geq \alpha |||v_h|||_{\text{vel}}^2, \tag{113}$$

where $\alpha = C_\eta$ as defined in Lemma 4.10.

To discretize the pressure-velocity coupling, we need a discrete counterpart of the bilinear form b defined on $U \times P$ by (103). We define on $U_h \times P_h$ the discrete bilinear form

$$b_h(v_h, q_h) = -\int_\Omega q_h \nabla_h \cdot v_h + \sum_{F \in \mathcal{F}_h} \int_F [\![v_h]\!] \cdot \mathrm{n}_F \{\!\{q_h\}\!\}, \tag{114}$$

where the broken divergence operator $\nabla_h \cdot$ acts elementwise, like the broken gradient operator ∇_h defined by (23). We observe that elementwise integration by parts yields

$$b_h(v_h, q_h) = \int_\Omega v_h \cdot \nabla_h q_h - \sum_{F \in \mathcal{F}_h^i} \int_F \{\!\{v_h\}\!\} \cdot \mathrm{n}_F [\![q_h]\!]. \tag{115}$$

Similarly to the operator B at the continuous level, we introduce the discrete operator B_h : $U_h \to P_h$ such that, for all $(v_h, q_h) \in X_h$,

$$(B_h v_h, q_h)_P = b_h(v_h, q_h).$$

It turns out that, contrary to the exact operator B, the discrete operator B_h is not surjective. As a result, the L^2-norm of a function in P_h cannot be controlled uniquely in terms of b_h. To recover control, it is necessary to add the following pressure seminorm defined on $H^1(\mathcal{T}_h)$:

$$|q|_p := \left(\sum_{F\in\mathcal{F}_h^i} h_F \|[\![q]\!]\|_{L^2(F)}^2 \right)^{1/2}.$$

Lemma 5.5 (Stability for b_h)**.** *There exists $\beta > 0$, independent of h, such that*

$$\forall q_h \in P_h, \qquad \beta\|q_h\|_P \leq \sup_{w_h\in U_h\setminus\{0\}} \frac{b_h(w_h,q_h)}{|\!|\!| w_h |\!|\!|_{\mathrm{vel}}} + |q_h|_p. \tag{116}$$

Remark 5.6 (Ladyzhenskaya–Babuška–Brezzi (LBB) condition). In the setting of conforming mixed finite element approximations, the stability of the discrete bilinear form coupling velocity and pressure takes the form of an inf-sup condition without stabilization term, the so-called Ladyzhenskaya–Babuška–Brezzi (LBB) condition (see Babuška [5] and Brezzi [17]). Condition (116) can be viewed as an extended LBB condition owing to the additional presence of the pressure seminorm on the right-hand side.

We consider the following discretization of problem (104): Find $(u_h, p_h) \in X_h$ such that

$$a_h(u_h, v_h) + b_h(v_h, p_h) = \int_\Omega f\cdot v_h \qquad \forall v_h \in U_h, \tag{117a}$$

$$-b_h(u_h, q_h) + s_h(p_h, q_h) = 0 \qquad \forall q_h \in P_h, \tag{117b}$$

where the discrete bilinear form a_h is defined by (111), the discrete bilinear form b_h by (114) (or, equivalently, by (115)), and where

$$s_h(q_h, r_h) := \sum_{F\in\mathcal{F}_h^i} h_F \int_F [\![q_h]\!][\![r_h]\!]. \tag{118}$$

The stabilization bilinear form s_h is meant to control pressure jumps across interfaces, thereby allowing to control the L^2-norm of the discrete pressure by virtue of Lemma 5.5. The following formulation, equivalent to (117), is obtained by summing equations (117a) and (117b): Find $(u_h, p_h) \in X_h$ such that

$$c_h((u_h, p_h),(v_h, q_h)) = \int_\Omega f\cdot v_h \text{ for all } (v_h, q_h) \in X_h, \tag{119}$$

where

$$c_h((u_h, p_h),(v_h, q_h)) := a_h(u_h, v_h) + b_h(v_h, p_h) - b_h(u_h, q_h) + s_h(p_h, q_h). \tag{120}$$

Owing to (113), we infer partial coercivity for c_h in the form

$$\begin{aligned}\forall (v_h, q_h) \in X_h, \qquad c_h((v_h, q_h),(v_h, q_h)) &= a_h(v_h, v_h) + s_h(q_h, q_h)\\ &\geq \alpha|\!|\!| v_h |\!|\!|_{\mathrm{vel}}^2 + |q_h|_p^2.\end{aligned} \tag{121}$$

To prove discrete well-posedness, we establish first the discrete inf-sup stability of the bilinear form c_h when the discrete space X_h is equipped with the norm

$$|\!|\!|(v_h, q_h)|\!|\!|_{\mathrm{sto}} := \left(|\!|\!| v_h |\!|\!|_{\mathrm{vel}}^2 + \|q_h\|_P^2 + |q_h|_p^2\right)^{1/2}. \tag{122}$$

Lemma 5.7 (Discrete inf-sup stability)**.** *Assume that the penalty parameter η in the SIP method is such that $\eta > \underline{\eta}$ with $\underline{\eta}$ defined in Lemma* 4.10*. Then, there is $\gamma > 0$, independent of h, such that, for all $(v_h, q_h) \in X_h$,*

$$\gamma|\!|\!|(v_h, q_h)|\!|\!|_{\mathrm{sto}} \leq \sup_{(w_h,r_h)\in X_h\setminus\{0\}} \frac{c_h((v_h, q_h),(w_h, r_h))}{|\!|\!|(w_h, r_h)|\!|\!|_{\mathrm{sto}}}. \tag{123}$$

As a consequence of Lemma 2.16, the discrete problem (117) or, equivalently, (119) is well-posed.

Convergence to smooth solutions To analyze the convergence of the solution of the discrete Stokes problem (117) or, equivalently, (119) in the case of smooth exact solutions. We proceed in the spirit of Theorem 2.20 and derive an error estimate in the $|||\cdot|||_{\mathrm{sto}}$-norm. Some additional regularity of the exact solution $(u, p) \in X$ is needed to assert consistency by plugging the pair (u, p) into the discrete bilinear form c_h. Concerning the velocity, we hinge for simplicity on Assumption 4.3 for all the velocity components. Concerning the pressure, we need traces on all interfaces and that the resulting jumps vanish; again for simplicity, this requirement is matched by assuming $H^1(\Omega)$-regularity for the pressure.

Assumption 5.8 (Regularity of the exact solution and space X_*)**.** *We assume that the exact solution (u, p) is in $X_* := U_* \times P_*$ where*

$$U_* := U \cap [H^2(\Omega)]^d, \qquad P_* := P \cap H^1(\Omega).$$

In the spirit of §2.5, we set

$$U_{*h} := U_* + U_h, \qquad P_{*h} := P_* + P_h, \qquad X_{*h} := X_* + X_h.$$

We extend the discrete bilinear form a_h defined by (111) to $U_{*h} \times U_h$ and the $|||\cdot|||_{\mathrm{vel}}$-norm to U_{*h}. The discrete bilinear form b_h can be extended to $[H^1(\mathcal{T}_h)]^d \times H^1(\mathcal{T}_h)$. Finally, we extend the discrete bilinear form c_h defined by (120) to $X_{*h} \times X_h$ and we extend the $|||\cdot|||_{\mathrm{sto}}$-norm defined by (122) to X_{*h}.

Lemma 5.9 (Jumps of ∇u and p across interfaces)**.** *Assume $(u, p) \in X_*$. Then,*

$$[\![\nabla u]\!]\cdot \mathrm{n}_F = 0 \quad \text{and} \quad [\![p]\!] = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{124}$$

Lemma 5.10 (Consistency)**.** *Assume that $(u, p) \in X_*$. Then,*

$$c_h((u, p), (v_h, q_h)) = \int_\Omega f \cdot v_h \qquad \forall (v_h, q_h) \in X_h.$$

Owing to Theorem 2.20 and recalling that discrete inf-sup stability holds true using the $|||\cdot|||_{\mathrm{sto}}$-norm, it remains to investigate the boundedness of the discrete bilinear form c_h. To this purpose, we define on X_{*h} the norm

$$|||(v, q)|||_{\mathrm{sto},*}^2 := |||(v, q)|||_{\mathrm{sto}}^2 + \sum_{T \in \mathcal{T}_h} h_T \|\nabla v|_T \cdot \mathrm{n}_T\|_{L^2(\partial T)}^2 + \sum_{T \in \mathcal{T}_h} h_T \|q\|_{L^2(\partial T)}^2.$$

There exists C_{bnd}, independent of h, such that, for all $(v, q) \in X_{*h}$ and all $(w_h, r_h) \in X_h$,

$$c_h((v, q), (w_h, r_h)) \le C_{\mathrm{bnd}} |||(v, q)|||_{\mathrm{sto},*} |||(w_h, r_h)|||_{\mathrm{sto}}.$$

Theorem 5.11 ($|||\cdot|||_{\mathrm{sto}}$-norm error estimate and convergence rate)**.** *Let $(u, p) \in X_*$ denote the unique solution of problem* (104)*. Let $(u_h, p_h) \in X_h$ solve* (119) *with c_h defined by* (120)*. Then, there is C, independent of h, such that*

$$|||(u - u_h, p - p_h)|||_{\mathrm{sto}} \le C \inf_{(v_h, q_h) \in X_h} |||(u - v_h, p - q_h)|||_{\mathrm{sto},*}. \tag{125}$$

Moreover, if $(u, p) \in [H^{k+1}(\Omega)]^d \times H^k(\Omega)$,

$$|||(u - u_h, p - p_h)|||_{\mathrm{sto}} \le C_{u,p} h^k,$$

with $C_{u,p} = C\left(\|u\|_{[H^{k+1}(\Omega)]^d} + \|p\|_{H^k(\Omega)}\right)$.

Remark 5.12 (Regularity assumption on the pressure). The regularity assumption $p \in H^k(\Omega)$ is just what is needed to achieve the overall convergence rate in the $|||\cdot|||_{\mathrm{sto}}$-norm of order h^k. Since polynomials of degree $\le k$ are used for the pressure, the contribution of the pressure terms to the error upper bound would be of order h^{k+1} if $p \in H^{k+1}(\Omega)$. In this case, the overall error would be dominated by the velocity error which is still of order h^k.

Remark 5.13 (L^2-norm error estimate on the velocity). An optimal L^2-error estimate on the velocity can be obtained using a duality argument in the same spirit as in §4.2.4 for the Poisson problem. To apply the Aubin–Nitsche argument [4], we need additional regularity for the solution of the Stokes problem (see Cattabriga [26] and Amrouche and Girault [1]).

Numerical fluxes We consider test functions having support localized to a single mesh element. We define the numerical fluxes

$$\phi_F^{\mathrm{grad}}(p_h) := \begin{cases} \{\!\{p_h\}\!\}\mathrm{n}_F & \text{if } F \in \mathcal{F}_h^i, \\ p_h\mathrm{n} & \text{if } F \in \mathcal{F}_h^b, \end{cases} \tag{126}$$

$$\phi_F^{\mathrm{div}}(u_h, p_h) := \begin{cases} \{\!\{u_h\}\!\}\cdot\mathrm{n}_F + h_F[\![p_h]\!] & \text{if } F \in \mathcal{F}_h^i, \\ 0 & \text{if } F \in \mathcal{F}_h^b, \end{cases} \tag{127}$$

and observe that $\phi_F^{\mathrm{grad}}(p_h)$ is vector-valued whereas $\phi_F^{\mathrm{div}}(u_h, p_h)$ is scalar-valued. Moreover, referring to §4.3.3 and, in particular, to (88) for the numerical fluxes associated with the SIP method, we consider here the vector-valued numerical fluxes

$$\phi_F^{\mathrm{diff}}(u_h) = -\{\!\{\nabla_h u_h\}\!\}\cdot\mathrm{n}_F + \frac{\eta}{h_F}[\![u_h]\!]. \tag{128}$$

Let $T \in \mathcal{T}_h$ and let $\xi \in [\mathbb{P}_d^k(T)]^d$ with Cartesian components $(\xi_i)_{1\le i\le d}$. Using $v_h = \xi\chi_T$ as a test function in the discrete momentum conservation equation (117a) (where χ_T denotes the characteristic function of T), we obtain

$$\int_T \sum_{i=1}^d G_h^l(u_{h,i})\cdot\nabla\xi_i - \int_T p_h \nabla\cdot\xi + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F \left[\phi_F^{\mathrm{diff}}(u_h) + \phi_F^{\mathrm{grad}}(p_h)\right]\cdot\xi = \int_T f\cdot\xi, \tag{129}$$

where $l \in \{k-1, k\}$, G_h^l is the discrete gradient operator, and $\epsilon_{T,F} = \mathrm{n}_T\cdot\mathrm{n}_F$.

Similarly, let $\zeta \in \mathbb{P}_d^k(T)$. Using $q_h = \zeta\chi_T - \langle\zeta\chi_T\rangle_\Omega$ as a test function in the discrete mass conservation equation (117b) and using the expression (115) of the discrete bilinear form b_h, we obtain

$$-\int_T u_h\cdot\nabla\zeta + \sum_{F\in\mathcal{F}_T} \epsilon_{T,F} \int_F \phi_F^{\mathrm{div}}(u_h, p_h)\zeta = 0. \tag{130}$$

Equations (129) and (130) express the local conservation properties satisfied by the dG approximation. We observe that, in the numerical fluxes $\phi_F^{\mathrm{grad}}(p_h)$ and $\phi_F^{\mathrm{div}}(u_h, p_h)$, the centered part results from the discrete bilinear form b_h, while the presence of the pressure jump in the flux $\phi_F^{\mathrm{div}}(u_h, p_h)$ stems from stabilizing the pressure jumps across interfaces.

Convergence to minimal regularity solutions In this section, we study the convergence of the sequence

$$(u_{\mathcal{H}}, p_{\mathcal{H}}) := ((u_h, p_h))_{h\in\mathcal{H}},$$

where, for all $h \in \mathcal{H}$, (u_h, p_h) solves the discrete problem (119), to the unique solution (u, p) of the steady Stokes problem (104) using mimimal regularity on (u, p), that is to say, $(u, p) \in X$. This result is an important building block in the convergence study of the dG discretization of the INS equations undertaken in §5.2.3. For conciseness of notation, subsequences are not renumbered in what follows.

To analyze the convergence of the diffusion term, we formulate the discrete bilinear form a_h using discrete gradients, namely, for all $v_h, w_h \in U_h$,

$$a_h(v_h, w_h) = \int_\Omega \sum_{i=1}^d G_h^l(v_{h,i})\cdot G_h^l(w_{h,i}) + \hat{s}_h(v_h, w_h), \tag{131}$$

with $l \in \{k-1, k\}$ and

$$\hat{s}_h(v_h, w_h) := \sum_{F\in\mathcal{F}_h} \frac{\eta}{h_F} \int_F [\![v_h]\!]\cdot[\![w_h]\!] - \int_\Omega \sum_{i=1}^d \mathrm{R}_h^l([\![v_{h,i}]\!])\cdot \mathrm{R}_h^l([\![w_{h,i}]\!]).$$

The expression (131) is equivalent to (111) on $U_h \times U_h$.

For any integer $l \geq 0$, we define the discrete divergence operator $D_h^l : [H^1(\mathcal{T}_h)]^d \to L^2(\Omega)$ such that, for all $v \in [H^1(\mathcal{T}_h)]^d$ with Cartesian components $(v_i)_{1\leq i\leq d}$,

$$D_h^l(v) := \sum_{i=1}^{d} G_h^l(v_i)\cdot e_i,$$

where e_i denotes the ith vector of the Cartesian basis of $\mathbb{R}^d$. Then, using the expression (114) for b_h, we observe that, for all $(v_h, q_h) \in X_h$,

$$b_h(v_h, q_h) = -\int_\Omega q_h D_h^k(v_h). \tag{132}$$

We can also introduce a new discrete gradient operator $\mathcal{G}_h^l : H^1(\mathcal{T}_h) \to L^2(\Omega)$ such that, for all $q \in H^1(\mathcal{T}_h)$,

$$\mathcal{G}_h^l(q) := \nabla_h q - \sum_{F\in\mathcal{F}_h^i} \mathrm{r}_F^l([\![q]\!]). \tag{133}$$

The only difference with respect to the discrete gradient operator G_h^l defined by (82) is that boundary faces are not included in the summation on the right-hand side of (133). A motivation for this modification is that there holds

$$\forall (v_h, q_h) \in X_h, \qquad \int_\Omega v_h\cdot\mathcal{G}_h^k(q_h) = -\int_\Omega q_h D_h^k(v_h),$$

so that an alternative expression for b_h on X_h is

$$b_h(v_h, q_h) = \int_\Omega v_h\cdot\mathcal{G}_h^k(q_h).$$

Theorem 5.14 (Convergence to minimal regularity solutions)**.** *Let $k \geq 1$. Let $(u_\mathcal{H}, p_\mathcal{H})$ be the sequence of approximate solutions generated by solving the discrete problems* (119) *on the admissible mesh sequence $\mathcal{T}_\mathcal{H}$. Then, as $h \to 0$,*

$$\begin{aligned}
u_h &\to u && \text{in } [L^2(\Omega)]^d,\\
\nabla_h u_h &\to \nabla u && \text{in } [L^2(\Omega)]^{d,d},\\
|u_{h,i}|_\mathrm{J} &\to 0 && \text{for all } i \in \{1,\dots,d\},\\
p_h &\to p && \text{in } L^2(\Omega),\\
|p_h|_p &\to 0,
\end{aligned}$$

where $(u,p) \in X$ denotes the unique solution to (104)*.*

5.1.3 Formulations without pressure stabilization

Fully discontinuous formulations, such as the one presented in §5.1.2, are appealing in problems where corner singularities are present (e.g., the well-known lid-driven cavity problem), since, in this context, discontinuous pressures are generally less prone to spurious oscillations. Using equal-order velocity and pressure spaces, however, requires penalizing pressure jumps across interfaces to achieve discrete stability. Such a term introduces a tighter coupling between the discrete momentum and mass conservation equations, since the pressure is also explicitly present in the mass conservation equation. In practice, this can be a drawback when using classical solution methods (such as the Uzawa method) for saddle-point problems in the steady case or projection methods in the unsteady case.

It turns out that the pressure penalty term can be omitted in various cases which, however, do not accommodate the same level of mesh generality as in §5.1.2. On matching affine quadrilateral or hexahedral meshes, formulations without pressure stabilization have been analyzed by Toselli [81] for different couples of polynomial degrees for velocity and pressure.

On matching simplicial meshes with polynomials for the pressure one degree less than for the velocity, inf-sup stability has been proven by Hansbo and Larson [58] in the incompressible limit of two-dimensional linear elasticity and by Girault, Rivière, and Wheeler [56] for the two- and three-dimensional Stokes equations in the context of domain decomposition methods (with polynomial degree for the velocity between 1 and 3). Still on matching simplicial meshes for $d \in \{2,3\}$, a fully parameter-free dG approximation using piecewise affine discrete velocities supplemented by element bubble functions coupled with continuous piecewise affine and/or piecewise constant discrete pressures has been analyzed by Burman and Stamm [23].

A means to achieve discrete inf-sup stability on matching simplicial meshes is to consider a discontinuous approximation of the velocity together with a continuous approximation of the pressure. This approach constitutes the basis for the projection method derived by Botti and Di Pietro [13] for the unsteady INS equations.

5.2 Steady Navier–Stokes flows

In this section, we consider steady Navier–Stokes flows. The main difference with respect to §5.1 is the inclusion of a nonlinear term modeling the convective transport of momentum. The discretization with dG methods of this nonlinear term is the main focus of this section. We also account for the viscosity ν in the momentum conservation equation. For steady Navier–Stokes flows, the viscosity ν is important since it quantifies the relative importance of convective and diffusive momentum transport.

5.2.1 The continuous setting

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3,4\}$, be a polyhedron, let $f \in [L^2(\Omega)]^d$ be the forcing term, and let $\nu > 0$ be the viscosity. The discussion of this section is confined to space dimensions up to 4 since the nonlinear term requires embeddings of functional spaces valid for $d \leq 4$. The steady INS problem reads

$$-\nu \triangle u + (u \cdot \nabla) u + \nabla p = f \quad \text{in } \Omega, \tag{134a}$$
$$\nabla \cdot u = 0 \quad \text{in } \Omega, \tag{134b}$$
$$u = 0 \quad \text{on } \partial\Omega, \tag{134c}$$
$$\langle p \rangle_\Omega = 0. \tag{134d}$$

Remark 5.15 (Conservative formulation). Since $(u \cdot \nabla) u = \nabla \cdot (u \otimes u)$ because $\nabla \cdot u = 0$, the momentum conservation equation (134a) can be rewritten in the *conservative form*

$$-\nu \triangle u + \nabla \cdot (u \otimes u) + \nabla p = f.$$

In contrast, equation (134a) is said to be in *nonconservative form*.

The weak formulation of system (134) reads: Find $(u,p) \in X$ such that

$$c((u,p),(v,q)) + t(u,u,v) = \int_\Omega f \cdot v \text{ for all } (v,q) \in X, \tag{135}$$

where $X = U \times P$ is defined by (102), the bilinear form $c \in \mathcal{L}(X \times X, \mathbb{R})$ now accounts for the viscosity and is given by

$$c((u,p),(v,q)) = \nu a(u,v) + b(v,p) - b(u,q),$$

with a and b still defined by (103), and the trilinear form $t \in \mathcal{L}(U \times U \times U, \mathbb{R})$ is such that

$$t(w,u,v) := \int_\Omega (w \cdot \nabla u) \cdot v = \int_\Omega \sum_{i,j=1}^{d} w_j (\partial_j u_i) v_i. \tag{136}$$

The trilinear form is indeed bounded on $U \times U \times U$: There is τ_Ω, only depending on Ω, such that, for all $w, u, v \in U$,

$$t(w,u,v) \leq \tau_\Omega \|w\|_U \|u\|_U \|v\|_U. \tag{137}$$

A further important property of the trilinear form t defined by (136) is skew-symmetry with respect to the last two arguments whenever the first argument is divergence-free and has zero normal component on the boundary. For simplicity, we consider that the three arguments of the trilinear form are in U.

Lemma 5.16 (Skew-symmetry of trilinear form)**.** *For all* $w \in U$*, there holds*

$$\forall v \in U, \qquad t(w,v,v) = -\frac{1}{2}\int_\Omega (\nabla\cdot w)|v|^2. \tag{138}$$

Moreover, if $w \in V := \{v \in U \mid \nabla\cdot v = 0\}$*,*

$$\forall v \in U, \qquad t(w,v,v) = 0. \tag{139}$$

A crucial consequence of Lemma 5.16 is that, using $(v,q) = (u,p)$ as a test function in (135) and since u is divergence-free, we obtain, up to the viscosity scaling, the same *energy balance* as for steady Stokes flows, namely

$$\nu\|\nabla u\|^2_{[L^2(\Omega)]^{d,d}} = \int_\Omega f\cdot u.$$

In other words, convection does not influence energy balance.

Theorem 5.17 (Existence and uniqueness)**.** *There exists at least one* $(u,p) \in X$ *solving* (135)*. Moreover, under the* smallness condition *on the data*

$$\tau_\Omega \|f\|_{U'} < (\nu\alpha_\Omega)^2, \tag{140}$$

the solution is unique.

Remark 5.18 (Interpretation of condition (140)). At fixed viscosity ν, condition (140) means that the forcing term f must be small enough. Alternatively, at fixed f, condition (140) means that the viscosity ν must be large enough (so that sufficiently energy is dissipated by the flow).

5.2.2 The discrete setting

In this section, we derive a dG discretization of the INS equations (135). For the Stokes part (resulting from the bilinear form c), we follow the approach of §5.1.2 and consider equal-order discontinuous velocities and pressures. Alternative dG methods to approximate the INS equations have been explored by Karakashian and Jureidini [62], Girault, Rivière, and Wheeler [56], and Cockburn, Kanschat, and Schötzau [35, 36, 37].

Let $\mathcal{T}_\mathcal{H}$ denote an admissible mesh sequence and let $k \geq 1$ be an integer. We consider the discrete spaces (cf. (110))

$$U_h := [\mathbb{P}^k_d(\mathcal{T}_h)]^d, \qquad P_h := \mathbb{P}^k_{d,0}(\mathcal{T}_h), \qquad X_h := U_h \times P_h.$$

The material in this section is restricted to $d \leq 3$.

When working with dG approximations, the convective velocity is generally not divergence-free (but only weakly divergence-free), so that the important property (139) is generally not satisfied. Following Temam [78, 79], a possible way to circumvent this difficulty is to modify the trilinear form t and to consider instead, for all $w, u, v \in U$,

$$\begin{aligned} t'(w,u,v) &= t(w,u,v) + \frac{1}{2}\int_\Omega (\nabla\cdot w) u\cdot v \\ &= \int_\Omega (w\cdot\nabla u)\cdot v + \frac{1}{2}\int_\Omega (\nabla\cdot w) u\cdot v. \end{aligned} \tag{141}$$

The following result is then a straightforward consequence of (138): For all $w \in U$, there holds

$$\forall v \in U, \qquad t'(w, v, v) = 0. \tag{142}$$

Moreover, $(u, p) \in X$ solves (135) if and only if $(u, p) \in X$ is such that

$$c((u, p), (v, q)) + t'(u, u, v) = \int_\Omega f{\cdot}v \text{ for all } (v, q) \in X.$$

We start with Temam's modification of the trilinear form t. Specifically, we consider broken differential operators in the trilinear form t' defined by (141) and set, for all $w_h, u_h, v_h \in U_h$,

$$t_h^{(0)}(w_h, u_h, v_h) := \int_\Omega (w_h{\cdot}\nabla_h u_h){\cdot}v_h + \frac{1}{2}\int_\Omega (\nabla_h{\cdot}w_h) u_h{\cdot}v_h.$$

Our first goal is to derive a discrete counterpart of (142). For all $w_h, v_h \in U_h$, integrating by parts elementwise and proceeding as usual, we obtain

$$t_h^{(0)}(w_h, v_h, v_h) = \frac{1}{2}\sum_{F\in\mathcal{F}_h}\int_F [\![w_h]\!]{\cdot}\mathrm{n}_F \{\!\{v_h{\cdot}v_h\}\!\} + \sum_{F\in\mathcal{F}_h^i}\int_F \{\!\{w_h\}\!\}{\cdot}\mathrm{n}_F [\![v_h]\!]{\cdot}\{\!\{v_h\}\!\}.$$

Since the right-hand side of the above equation is nonzero, we modify $t_h^{(0)}$ as

$$\begin{aligned} t_h(w_h, u_h, v_h) := \int_\Omega (w_h{\cdot}\nabla_h u_h){\cdot}v_h - \sum_{F\in\mathcal{F}_h^i}\int_F \{\!\{w_h\}\!\}{\cdot}\mathrm{n}_F [\![u_h]\!]{\cdot}\{\!\{v_h\}\!\} \\ + \frac{1}{2}\int_\Omega (\nabla_h{\cdot}w_h)(u_h{\cdot}v_h) - \frac{1}{2}\sum_{F\in\mathcal{F}_h}\int_F [\![w_h]\!]{\cdot}\mathrm{n}_F \{\!\{u_h{\cdot}v_h\}\!\}. \end{aligned} \tag{143}$$

This choice, which incorporates Temam's modification *at the discrete level*, possesses the following important property which is the discrete counterpart of Lemma 5.16.

Lemma 5.19 (Skew-symmetry of discrete trilinear form)**.** *For all $w_h \in U_h$, there holds*

$$\forall v_h \in U_h, \qquad t_h(w_h, v_h, v_h) = 0. \tag{144}$$

We now address the boundedness of the discrete trilinear form t_h on $U_h \times U_h \times U_h$. Recall that the discrete velocity space U_h is equipped with the $|\!|\!|{\cdot}|\!|\!|_{\mathrm{vel}}$-norm defined by (112). Then, there is τ, independent of h, such that, for all $w_h, u_h, v_h \in U_h$, there holds

$$t_h(w_h, u_h, v_h) \le \tau |\!|\!|w_h|\!|\!|_{\mathrm{vel}} |\!|\!|u_h|\!|\!|_{\mathrm{vel}} |\!|\!|v_h|\!|\!|_{\mathrm{vel}}.$$

Let a_h and b_h be the discrete bilinear forms considered for the linear Stokes equations, cf. (111) for a_h and (114) or, equivalently, (115) for b_h. Let t_h be the discrete trilinear form defined by (143). The discrete INS problem reads: Find $(u_h, p_h) \in X_h$ such that

$$\nu a_h(u_h, v_h) + t_h(u_h, u_h, v_h) + b_h(v_h, p_h) = \int_\Omega f{\cdot}v_h \qquad \forall v_h \in U_h, \tag{145a}$$

$$-b_h(u_h, q_h) + \nu^{-1} s_h(p_h, q_h) = 0 \qquad \forall q_h \in P_h, \tag{145b}$$

or, equivalently, such that

$$c_h((u_h, p_h), (v_h, q_h)) + t_h(u_h, u_h, v_h) = \int_\Omega f{\cdot}v_h \qquad \forall (v_h, q_h) \in X_h, \tag{146}$$

with

$$c_h((u_h, p_h), (v_h, q_h)) := \nu a_h(u_h, v_h) + b_h(v_h, p_h) - b_h(u_h, q_h) + \nu^{-1} s_h(p_h, q_h).$$

We observe that both the diffusion and pressure stabilization terms differ from the case of the linear Stokes equations, cf. (120), since the former is scaled by the viscosity and the latter by the reciprocal of the viscosity.

Recalling (113), let $\alpha > 0$ denote the coercivity parameter of the discrete bilinear form a_h such that

$$\forall v_h \in U_h, \qquad a_h(v_h, v_h) \geq \alpha \|v_h\|_{\mathrm{vel}}^2.$$

This leads to partial coercivity for the discrete bilinear form c_h in the form

$$\forall (v_h, q_h) \in X_h, \qquad c_h((v_h, q_h), (v_h, q_h)) \geq \nu\alpha \|v_h\|_{\mathrm{vel}}^2 + \nu^{-1}|q_h|_p^2. \tag{147}$$

Moreover, we redefine the $\|\cdot\|_{\mathrm{sto}}$-norm as

$$\|(v_h, q_h)\|_{\mathrm{sto}} := \left(\nu \|v_h\|_{\mathrm{vel}}^2 + \|q_h\|_P^2 + \nu^{-1}|q_h|_p^2\right)^{1/2}.$$

It is straightforward to verify, as in the proof of Lemma 5.7, the following discrete inf-sup condition: There is $\gamma > 0$, independent of h and of the viscosity ν, such that, for all $(v_h, q_h) \in X_h$,

$$\gamma \|(v_h, q_h)\|_{\mathrm{sto}} \leq \sup_{(w_h, r_h) \in X_h \setminus \{0\}} \frac{c_h((v_h, q_h), (w_h, r_h))}{\|(w_h, r_h)\|_{\mathrm{sto}}}. \tag{148}$$

We observe that the fact that γ is independent of ν results from the scaling used in the pressure stabilization.

Theorem 5.20 (Existence and uniqueness)**.** *There exists at least one* $(u_h, p_h) \in X_h$ *solving* (146)*. Moreover, under the smallness condition*

$$\tau \|f\|_{[L^2(\Omega)]^d} < (\nu\alpha)^2, \tag{149}$$

the solution is unique.

5.2.3 Convergence analysis

In this section, we address the convergence of the sequence $(u_{\mathcal{H}}, p_{\mathcal{H}})$ of solutions to the discrete problem (146) on the admissible mesh sequence $\mathcal{T}_{\mathcal{H}}$ to a solution (u, p) of the INS equations (135).

Theorem 5.21 (Convergence)**.** *Let* $(u_{\mathcal{H}}, p_{\mathcal{H}})$ *be a sequence of approximate solutions generated by solving the discrete problems* (146) *on the admissible mesh sequence* $\mathcal{T}_{\mathcal{H}}$*. Then, as* $h \to 0$*, up to a subsequence,*

$$\begin{aligned} u_h &\to u && \text{in } [L^2(\Omega)]^d, \\ \nabla_h u_h &\to \nabla u && \text{in } [L^2(\Omega)]^{d,d}, \\ |u_h|_{\mathrm{J}} &\to 0, && \\ p_h &\to p && \text{in } L^2(\Omega), \\ |p_h|_p &\to 0, && \end{aligned}$$

where $(u, p) \in X$ *is a solution of* (135)*. Moreover, under the smallness condition* (149)*, the whole sequence converges to the unique solution of* (135)*.*

Remark 5.22 (Reformulation of discrete trilinear form)*.* In the convergence analysis, the following equivalent expression of t_h in terms of discrete gradients and discrete divergence is important: For all $w_h, u_h, v_h \in U_h$,

$$\begin{aligned} t_h(w_h, u_h, v_h) = &\int_\Omega \sum_{i=1}^{d} w_h \cdot \mathcal{G}_h^{2k}(u_{h,i}) v_{h,i} + \frac{1}{2}\int_\Omega D_h^{2k}(w_h) u_h \cdot v_h \\ &+ \frac{1}{4} \sum_{F \in \mathcal{F}_h^i} \int_F [\![w_h]\!] \cdot \mathrm{n}_F [\![u_h]\!] \cdot [\![v_h]\!]. \end{aligned} \tag{150}$$

We observe that the polynomial degree used for the discrete gradients and divergence is $2k$ owing to the nonlinearities.

References

[1] C. Amrouche and V. Girault. On the existence and regularity of the solution of Stokes problem in arbitrary dimension. *Proc. Japan. Acad.*, 67:171–175, 1991.

[2] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19:742–760, 1982.

[3] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2002.

[4] J.-P. Aubin. *Analyse fonctionnelle appliquée.* Presses Universitaires de France, Paris, 1987.

[5] I. Babuška. The finite element method with lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973.

[6] I. Babuška. The finite element method with penalty. *Math. Comp.*, 27:221–228, 1973.

[7] I. Babuška and M. Zlámal. Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.*, 10(5):863–875, 1973.

[8] G. A. Baker. Finite element methods for elliptic equations using nonconforming elements. *Math. Comp.*, 31(137):45–49, 1977.

[9] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` — a general-purpose object-oriented finite element library. *ACM Trans. Math. Softw.*, 33(4), 2007.

[10] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131(2):267–279, 1997.

[11] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini. A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. In R. Decuypere and G. Dibelius, editors, *Proceedings of the* 2^{nd} *European Conference on Turbomachinery Fluid Dynamics and Thermodynamics*, pages 99–109, 1997.

[12] P. Bastian, M. Droske, C. Engwer, R. Klöfkorn, T. Neubauer, M. Ohlberger, and M. Rumpf. Towards a unified framework for scientific computing. In *Proc. of the 15th international conference on domain decomposition methods*, 2005.

[13] L. Botti and D. A. Di Pietro. A pressure-correction scheme for convection-dominated incompressible flows with discontinuous velocity and continuous pressure. Submitted. Preprint available at `http://hal.archives-ouvertes.fr/hal-00458293/en/`, 2010.

[14] D. Braess. *Finite elements.* Cambridge University Press, Cambridge, third edition, 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[15] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics.* Springer, New York, third edition, 2008.

[16] H. Brézis. *Analyse fonctionnelle. Théorie et applications. [Functional Analysis. Theory and applications].* Applied Mathematics Series for the Master's Degree. Masson, Paris, France, 1983.

[17] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers. *RAIRO Anal. Numér.*, pages 129–151, 1974.

[18] F. Brezzi, G. Manzini, L. D. Marini, P. Pietra, and A. Russo. Discontinuous finite elements for diffusion problems. In *Atti del convegno in onore di F. Brioschi (Milano 1997)*, pages 197–217. Istituto Lombardo, Accademia di Scienze e di Lettere, 1999.

[19] F. Brezzi, G. Manzini, L. D. Marini, P. Pietra, and A. Russo. Discontinuous Galerkin approximations for elliptic problems. *Numer. Methods Partial Differential Equations*, 16(4):365–378, 2000.

[20] F. Brezzi, L. D. Marini, and E. Süli. Discontinuous Galerkin methods for first-order hyperbolic problems. *Math. Models Methods Appl. Sci.*, 14(12):1893–1903, 2004.

[21] A. Buffa and C. Ortner. Compact embeddings of broken Sobolev spaces and applications. *IMA J. Numer. Anal.*, 4(29):827–855, 2009.

[22] E. Burman and A. Ern. Discontinuous Galerkin approximation with discrete variational principle for the nonlinear Laplacian. *C. R. Math. Acad. Sci. Paris*, 346(17–18):1013–1016, 2008.

[23] E. Burman and B. Stamm. Bubble stabilized discontinuous Galerkin method for Stokes' problem. *Math. Models Methods Appl. Sci.*, 20(2):297–313, 2010.

[24] 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(5):1676–1706 (electronic), 2000.

[25] P. E. Castillo. Stencil reduction algorithms for the local discontinuous Galerkin method. *Int. J. Numer. Meth. Engng.*, 81:1475–1491, 2010.

[26] L. Cattabriga. Su un problema al contorno relativo al sistema di equazioni di Stokes. *Rend. Sem. Mat. Univ. Padova*, 31:308–340, 1961.

[27] P. Causin and R. Sacco. A discontinuous Petrov-Galerkin method with Lagrangian multipliers for second order elliptic problems. *SIAM J. Numer. Anal.*, 43(1):280–302 (electronic), 2005.

[28] G. Chavent and B. Cockburn. The local projection P^0P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO Modél. Math. Anal. Numér.*, 23(4):565–592, 1989.

[29] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[30] B. Cockburn, B. Dong, and J. Guzmán. A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Math. Comp.*, 77(264):1887–1916, 2008.

[31] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.

[32] B. Cockburn, J. Guzmán, and H. Wang. Superconvergent discontinuous Galerkin methods for second-order elliptic problems. *Math. Comp.*, 78(265):1–24, 2009.

[33] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Math. Comp.*, 54(190):545–581, 1990.

[34] B. Cockburn, G. Kanschat, and D. Schötzau. The local discontinuous Galerkin method for the Oseen equations. *Math. Comp.*, 73:569–593, 2004.

[35] B. Cockburn, G. Kanschat, and D. Schötzau. A locally conservative LDG method for the incompressible Navier-Stokes equations. *Math. Comp.*, 74:1067–1095, 2005.

[36] B. Cockburn, G. Kanschat, and D. Schötzau. A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations. *J. Sci. Comput.*, 31(1-2):61–73, 2007.

[37] B. Cockburn, G. Kanschat, and D. Schötzau. An equal-order DG method for the incompressible Navier-Stokes equations. *J. Sci. Comput.*, 40(1-3):188–210, 2009.

[38] B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. Local Discontinuous Galerkin methods for the Stokes system. *SIAM J. Numer. Anal.*, 40(1):319–343 (electronic), 2002.

[39] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. *Discontinuous Galerkin Methods - Theory, Computation and Applications*, volume 11 of *Lecture Notes in Computer Science and Engineering*. Springer, 2000.

[40] B. Cockburn, S. Lin, and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems. *J. Comput. Phys.*, 84(1):90–113, 1989.

[41] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comp.*, 52(186):411–435, 1989.

[42] B. Cockburn and C.-W. Shu. The Runge-Kutta local projection P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO Modél. Math. Anal. Numér.*, 25(3):337–361, 1991.

[43] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin finite element method for convection-diffusion systems. *SIAM J. Numer. Anal.*, 35:2440–2463, 1998.

[44] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems. *J. Comput. Phys.*, 141(2):199–224, 1998.

[45] I. Danaila, F. Hecht, and O. Pironneau. *Simulation numérique en C++*. Dunod, Paris, 2003. `http://www.freefem.org`.

[46] D. A. Di Pietro and A Ern. Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier-Stokes equations. *Math. Comp.*, 79(271):1303–1330, 2010.

[47] J. Douglas Jr. and T. Dupont. *Lecture Notes in Physics*, volume 58, chapter Interior penalty procedures for elliptic and parabolic Galerkin methods. Springer-Verlag, Berlin, 1976.

[48] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Num. Math.*, 105(1):35–71, 2006.

[49] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, NY, 2004.

[50] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. I. General theory. *SIAM J. Numer. Anal.*, 44(2):753–778, 2006.

[51] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. II. Second-order elliptic PDEs. *SIAM J. Numer. Anal.*, 44(6):2363–2388, 2006.

[52] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. III. Multi-field theories with partial coercivity. *SIAM J. Numer. Anal.*, 46(2):776–804, 2008.

[53] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[54] R. E. Ewing, Junping Wang, and Y. Yang. A stabilized discontinuous finite element method for elliptic problems. *Numer. Linear Algebra Appl.*, 10(1-2):83–104, 2003. Dedicated to the 60th birthday of Raytcho Lazarov.

[55] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.

[56] V. Girault, B. Rivière, and M. F. Wheeler. A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems. *Math. Comp.*, 74(249):53–84 (electronic), 2004.

[57] P. Grisvard. *Singularities in Boundary Value Problems*. Masson, Paris, 1992.

[58] P. Hansbo and M. G. Larson. Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method. *Comput. Methods Appl. Mech. Engrg.*, 191(17-18):1895–1908, 2002.

[59] J. S. Hesthaven and T. Warburton. *Nodal discontinuous Galerkin methods*, volume 54 of *Texts in Applied Mathematics*. Springer, New York, 2008. Algorithms, analysis, and applications.

[60] J. Jaffré, C. Johnson, and A. Szepessy. Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *M3AS*, 5(3):367–386, 1995.

[61] C. Johnson and J. Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comp.*, 46(173):1–26, 1986.

[62] O. A. Karakashian and W. N. Jureidini. A nonconforming finite element method for the stationary Navier-Stokes equations. *SIAM J. Numer. Anal.*, 35(1):93–120 (electronic), 1998.

[63] B. Kirk, J. W. Peterson, R. H. Stogner, and G. F. Carey. `libmesh`: A C++ library for parallel adaptive mesh refinement/coarsening simulation. *Engineering with Computers*, 22(3–4):237–254, 2006.

[64] P. D. Lax and A. N. Milgram. Parabolic equations. In *Contributions to the theory of partial differential equations*, Annals of Mathematics Studies, no. 33, pages 167–190. Princeton University Press, Princeton, N. J., 1954.

[65] P. Lesaint. Finite element methods for symmetric hyperbolic equations. *Numer. Math.*, 21:244–255, 1973/74.

[66] P. Lesaint. *Sur la résolution des systèmes hyperboliques du premier ordre par des méthodes d'éléments finis*. PhD thesis, University of Paris VI, 1975.

[67] P. Lesaint and P.-A. Raviart. On a finite element method for solving the neutron transport equation. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 89–123. Publication No. 33. Math. Res. Center, Univ. of Wisconsin-Madison, Academic Press, New York, 1974.

[68] A. Lew, P. Neff, D. Sulsky, and M. Ortiz. Optimal BV estimates for a discontinuous Galerkin method for linear elasticity. *AMRX Appl. Math. Res. Express*, 3:73–106, 2004.

[69] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971. Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

[70] J. Nitsche. On Dirichlet problems using subspaces with nearly zero boundary conditions. In *The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972)*, pages 603–627. Academic Press, New York, 1972.

[71] K. B. Ølgaard, A. Logg, and G. N. Wells. Automated code generation for discontinuous Galerkin methods. *SIAM J. Sci. Comput.*, 31(2):849–864, 2008.

[72] J. Peraire and P.-O. Persson. The compact discontinuous Galerkin (CDG) method for elliptic problems. *SIAM J. Sci. Comput.*, 30(4):1806–1824, 2008.

[73] I. Perugia and D. Schötzau. The *hp*-local discontinuous Galerkin method for low-frequency time-harmonic Maxwell equations. *Math. Comp.*, 72(243):1179–1214, 2003.

[74] C. Prud'homme. A domain specific embedded language in C++ for automatic differentiation, projection, integration and variational formulations. *Scientific Programming*, 14(2):81–110, 2006.

[75] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-0479, `http://lib-www.lanl.gov/cgi-bin/getfile?00354107.pdf`, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.

[76] S. J. Sherwin, R. M. Kirby, J. Peiró, R. L. Taylor, and O. C. Zienkiewicz. On 2D elliptic discontinuous Galerkin methods. *Int. J. Numer. Meth. Engng.*, 65:752–784, 2006.

[77] G. Strang. Variational crimes in the finite element method. In A.K. Aziz, editor, *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, New York, NY, 1972. Academic Press.

[78] R. Temam. Une méthode d'approximation de la solution des équations de Navier–Stokes. *Bull. Soc. Math. France*, 98:115–152, 1968.

[79] R. Temam. *Navier-Stokes Equations*, volume 2 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, revised edition, 1979. Theory and numerical analysis, With an appendix by F. Thomasset.

[80] A. ten Eyck and A. Lew. Discontinuous Galerkin methods for non-linear elasticity. *Internat. J. Numer. Methods Engrg.*, 67(9):1204–1243, 2006.

[81] A. Toselli. *hp*-finite element discontinuous Galerkin approximations for the Stokes problem. *M3AS*, 12(11):1565–1616, 2002.

[82] M. F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.*, 15:152–161, 1978.

[83] L. Yuan and C.-W. Shu. Discontinuous Galerkin method based on non-polynomial approximation spaces. *J. Comput. Phys.*, 218(1):295–323, 2006.