DISCONTINUOUS GALERKIN METHODS FOR FRIEDRICHS' SYMMETRIC SYSTEMS. I. GENERAL THEORY*

A. ERN† AND J.-L. GUERMOND‡

Abstract. This paper presents a unified analysis of Discontinuous Galerkin methods to approximate Friedrichs' symmetric systems. An abstract set of conditions is identified at the continuous level to guarantee existence and uniqueness of the solution in a subspace of the graph of the differential operator. Then a general Discontinuous Galerkin method that weakly enforces boundary conditions and mildly penalizes interface jumps is proposed. All the design constraints of the method are fully stated, and an abstract error analysis in the spirit of Strang's Second Lemma is presented. Finally, the method is formulated locally using element fluxes, and links with other formulations are discussed. Details are given for three examples, namely advection–reaction equations, advection–diffusion–reaction equations, and the Maxwell equations in the diffusive regime.

Key words. Friedrichs' systems, Finite Elements, Partial Differential Equations, Discontinuous Galerkin Method

AMS subject classifications. 65N30, 65M60, 35F15

1. Introduction.

Discontinuous Galerkin (DG) methods have been introduced in the 1970s, and their development has since followed two somewhat parallel routes depending on whether the PDE is hyperbolic or elliptic.

For hyperbolic PDEs, the first DG method was introduced by Reed and Hill in 1973 [25] to simulate neutron transport and the first analysis of DG methods for hyperbolic equations in an already rather general and abstract form was done by Lesaint and Raviart in 1974 [22, 21]. The analysis was subsequently improved by Johnson et al. who established that the optimal order of convergence in the L^2-norm is $p+\frac{1}{2}$ if polynomials of degree p are used [19]. More recently, DG methods for hyperbolic and nearly hyperbolic equations experienced a significant development based on the ideas of numerical fluxes, approximate Riemann solvers, and slope limiters; see, e.g., Cockburn et al. [9] and references therein for a thorough review. This renewed interest in DG methods is stimulated by several factors including the flexibility offered by the use of non-matching grids and the possibility to use high-order *hp*-adaptive finite element methods; see, e.g., Süli et al. [27].

For elliptic PDEs, DG methods originated from the use of Interior Penalties (IP) to weakly enforce continuity conditions imposed on the solution or its derivatives across the interfaces between adjoining elements; see, e.g., Babuška [3], Babuška and Zlámal [4], Douglas and Dupont [13], Baker [6], Wheeler [28], and Arnold [1]. DG methods for elliptic problems in mixed form were introduced more recently. Initially, discontinuous approximation was used solely for the primal variable, the flux being still discretized in a conforming fashion; see, e.g., Dawson [11, 12]. Then, discontinuous approximation of both the primal variable and its flux has been introduced by Bassi and Rebay [7] and further extended by Cockburn and Shu [10] leading to the so-called Local Discontinuous Galerkin (LDG) method. Around the same time, Baumann and Oden [8] proposed a nonsymmetric variant of DG for elliptic problems. This method was further developed and analyzed by Oden et al. [23] and by Rivière et al. [26].

*Draft version: 9th February 2005

†CERMICS, Ecole nationale des ponts et chaussées, Champs sur Marne, 77455 Marne la Vallée Cedex 2, France. (ern@cermics.enpc.fr)

‡Dept. Math, Texas A&M, College Station, TX 77843-3368, USA (guermond@math.tamu.edu) and LIMSI (CNRS-UPR 3152), BP 133, 91403, Orsay, France.

The fact that several DG methods (including IP methods) share common features and can be tackled by similar analysis tools called for a unified analysis. A first important step in that direction has been recently accomplished by Arnold et al. [2] for elliptic equations. It is shown in [2] that it is possible to cast many DG methods for the Laplacian with homogeneous Dirichlet boundary conditions into a single framework amenable to a unified error analysis. The main idea consists of using the mixed formulation of the Laplacian to define numerical fluxes and to locally eliminate these fluxes so as to derive a method involving only the primal variable.

The goal of the present paper is to propose a unified analysis of DG methods that goes beyond the traditional hyperbolic/elliptic classification of PDEs. The key is that we make systematic use of the theory of Friedrichs' symmetric systems to formulate DG methods and to perform the convergence analysis. This paper is the first part of a more comprehensive study on DG methods for Friedrichs' symmetric systems. In this paper we concentrate on first-order PDFEs only. The forthcoming second part of this work will deal more specifically with Friedrichs' symmetric systems associated with second-order PDEs.

The paper is organized as follows. In §2 we revisit Friedrichs' theory. Friedrichs' symmetric systems [16] are systems of coupled first-order PDEs endowed with symmetry and positivity properties. Examples include advection–reaction equations, advection–diffusion-reaction equations, the wave equation, the Maxwell equations, to cite a few. The main novelty of our analysis is that we address the well-posedness of the problem in the graph space, as opposed to Friedrichs who addressed the question of the uniqueness of strong solutions and that of the existence of weak solutions in L^2. In §3 we present three important examples of Friedrichs' symmetric systems, namely advection–reaction equations, advection–diffusion–reaction equations, and a simplified version of the Maxwell equations in the diffusive regime. For completeness, we treat Dirichlet, Neumann, and Robin boundary conditions for the second example. In all cases, we show that the abstract results of §2 ensuring well-posedness hold. Drawing on earlier ideas by Lesaint and Raviart [22, 21] and Johnson et al. [19], we propose in §4 an general framework for DG methods. This section contains three main contributions. First, the generic DG method is formulated in terms of a boundary operator enforcing boundary conditions weakly and in terms of an interface operator penalizing the jumps of the solution across the mesh interfaces. Second, the convergence analysis is performed in the spirit of Strang's Second Lemma by using two different norms, namely a stability norm for which a discrete inf-sup condition holds and an approximability norm ensuring the continuity of the DG bilinear form. All the design constraints to be fulfilled by the boundary and the interface operators for the error analysis to hold are clearly stated. Finally, using integration by parts, the DG method is re-interpreted locally by introducing the concept of element fluxes and element adjoint-fluxes, thus providing a direct link with engineering practice where approximation schemes are often designed by specifying element fluxes. Finally, §5 reviews various DG approximations for the model problems investigated in §3. In all the cases, the degrees of freedom in the design of the DG method are underlined, and the full expressions of the element fluxes and element adjoint-fluxes are given.

2. Friedrichs' symmetric systems. The goal of this section is to reformulate Friedrichs' theory by giving special care to the meaning of the boundary conditions. In particular, we avoid invoking traces at the boundary. The main results of this section are Theorem 2.5 and Theorem 2.8.

2.1. The setting. Let Ω be a bounded, open, and connected Lipschitz domain in $\mathbb{R}^d$. We denote by $\mathfrak{D}(\Omega)$ the space of $\mathfrak{C}^\infty$ functions that are compactly supported in Ω.

Let m be a positive integer. Let $\mathcal{K}$ and $\{\mathcal{A}^k\}_{1\le k\le d}$ be $(d+1)$ functions on Ω with values in $\mathbb{R}^{m,m}$. Henceforth, we assume that

$$\mathcal{K} \in [L^\infty(\Omega)]^{m,m}, \tag{A1}$$

$$\forall k \in \{1,\dots,d\},\ \mathcal{A}^k \in [L^\infty(\Omega)]^{m,m} \quad \text{and} \quad \sum_{k=1}^d \partial_k \mathcal{A}^k \in [L^\infty(\Omega)]^{m,m}, \tag{A2}$$

$$\forall k \in \{1,\dots,d\},\ \mathcal{A}^k = (\mathcal{A}^k)^t \ \text{ a.e. in } \Omega. \tag{A3}$$

In the rest of this work, it is implicitly assumed that (A1)–(A3) hold. Define the $\mathbb{R}^m$-valued operator A such that for all $u \in [\mathfrak{C}^1(\Omega)]^m$,

$$Au = \sum_{k=1}^d \mathcal{A}^k \partial_k u. \tag{2.1}$$

Set $L = [L^2(\Omega)]^m$. We say that a function u in L has an A-weak derivative in L if the linear form

$$[\mathfrak{D}(\Omega)]^m \ni \varphi \longmapsto -\int_\Omega \sum_{k=1}^d u^t \partial_k(\mathcal{A}^k\varphi) \in \mathbb{R}, \tag{2.2}$$

is bounded on L, and we denote by Au the function in L that can be associated with the above linear form by means of the Riesz representation theorem. Accordingly define the graph space

$$W = \{w \in L;\, Aw \in L\}, \tag{2.3}$$

and equip W with the graph norm

$$\|w\|_W = \|Aw\|_L + \|w\|_L, \tag{2.4}$$

and the associated scalar product. W is a Hilbert space. Indeed, let v_n be a Cauchy sequence in W; i.e., v_n and Av_n are Cauchy sequences in L. Let v and w be the corresponding limits in L. Let $\varphi \in [\mathfrak{D}(\Omega)]^m$. Then, using the symmetry of $\mathcal{A}^k$ and an integration by parts yields

$$\int_\Omega \sum_{k=1}^d v^t \partial_k(\mathcal{A}^k\varphi) \leftarrow \int_\Omega \sum_{k=1}^d v_n^t \partial_k(\mathcal{A}^k\varphi) = -\int_\Omega \varphi^t A v_n \to -\int_\Omega \varphi^t w,$$

which means that v has an A-weak derivative in L and $Av = w$. Since $[\mathcal{D}(\Omega)]^m \subset W$, W is dense in L; as result, we shall henceforth use L as a pivot space, i.e., $W \subset L \equiv L' \subset W'$. Note that, owing to (A2)–(A3), $[H^1(\Omega)]^m$ is a subspace of W.

Let $K \in \mathcal{L}(L;L)$ be defined such that $K : L \ni v \mapsto \mathcal{K}v \in L$ and set

$$T = A + K. \tag{2.5}$$

Then, $T \in \mathcal{L}(W;L)$. Let $K^* \in \mathcal{L}(L;L)$ be the adjoint operator of K, i.e., for all $v \in L$, $K^*v = \mathcal{K}^t v$. Let $T^* \in \mathcal{L}(W;L)$ be the formal adjoint of T,

$$T^*w = -\sum_{k=1}^d \partial_k(\mathcal{A}^k w) + K^*w, \quad \forall w \in W. \tag{2.6}$$

In this definition $\sum_{k=1}^d \partial_k(\mathcal{A}^k w)$ is understood in the weak sense. It can easily be verified that this weak derivative exists in L whenever w is in W. Moreover, the usual rule for differentiating products applies. In particular, upon introducing the operator $\nabla{\cdot}A \in \mathcal{L}(L;L)$ such that $(\nabla{\cdot}A)w = (\sum_{k=1}^d \partial_k \mathcal{A}^k)w$ for all $w \in L$, the following holds

$$\forall w \in W, \quad Tw + T^* w = (K + K^* - \nabla{\cdot}A)\, w. \tag{2.7}$$

DEFINITION 2.1. *Let $D \in \mathcal{L}(W;W')$ be the operator such that*

$$\forall (u,v) \in W \times W, \quad \langle Du, v\rangle_{W',W} = (Tu, v)_L - (u, T^* v)_L. \tag{2.8}$$

This definition makes sense since both T and its formal adjoint T^* are in $\mathcal{L}(W;L)$. Note that D is a boundary operator in the sense that $[\mathfrak{D}(\Omega)]^m \subset \mathrm{Ker}(D)$; see also Remark 2.1.

LEMMA 2.2. *The operator D is self-adjoint.*

Proof. Let $(u,v) \in W \times W$. A straightforward calculation yields

$$\begin{aligned}\langle Du, v\rangle_{W',W} - \langle Dv, u\rangle_{W',W} &= (Tu, v)_L - (u, T^* v)_L - (Tv, u)_L + (v, T^* u)_L \\ &= (\mathcal{Z}u, v)_L - (u, \mathcal{Z}v)_L = 0,\end{aligned}$$

since $\mathcal{Z} = K + K^* - \nabla{\cdot}A$ is self-adjoint. ∎

Remark 2.1. Let $n = (n_1, \ldots, n_d)^t$ be the unit outward normal to $\partial\Omega$. The usual way of presenting Friedrichs' symmetric systems consists of assuming that the fields $\{\mathcal{A}^k\}_{1\le k\le d}$ are smooth enough so that the matrix $\mathcal{D} = \sum_{k=1}^d n_k \mathcal{A}^k$ is meaningful at the boundary. Then, owing to (A3), the operator D can be represented as follows

$$\langle Du, v\rangle_{W',W} = \int_{\partial\Omega} \sum_{k=1}^d v^t n_k \mathcal{A}^k u = \int_{\partial\Omega} v^t \mathcal{D} u,$$

whenever u and v and smooth functions. Provided $[\mathfrak{C}^1(\overline{\Omega})]^m$ is dense in $[H^1(\Omega)]^m$ and in W, it can be shown that $\mathcal{D}u \in [H^{-\frac{1}{2}}(\partial\Omega)]^m$. Further characterization and regularity results on $\mathcal{D}u$ can be found in [24].

2.2. The well-posedness result. Let V be a subspace of W. Our goal is to analyze the well-posedness of the following problem: For f in L,

$$\begin{cases} \text{Seek } u \in V \text{ such that} \\ Tu = f. \end{cases} \tag{2.9}$$

To guarantee that $T : V \to L$ is an isomorphism, we make additional (sufficient) hypotheses. One additional hypothesis is made on the operator T, namely

$$\exists \mu_0 > 0, \quad \forall w \in W, \quad (Tw, w)_L + (w, T^* w)_L \ge 2\mu_0 \|w\|_L^2. \tag{A4}$$

Hypothesis (A4) is a positivity assumption introduced by Friedrichs [16]. This hypothesis can be reformulated as follows:

$$\mathcal{K} + \mathcal{K}^t - \textstyle\sum_{k=1}^d \partial_k \mathcal{A}^k \ge 2\mu_0 \mathcal{I}_m \ \text{ a.e. on } \Omega, \tag{2.10}$$

where $\mathcal{I}_m$ is the identity matrix in $\mathbb{R}^{m,m}$. An important consequence of assumption (A4) is the following:

LEMMA 2.3. *Assume* (A4). *Then, for all* $w \in W$,

$$(Tw, w)_L \geq \mu_0 \|w\|_L^2 + \tfrac{1}{2}\langle Dw, w\rangle_{W',W}. \tag{2.11}$$

$$(T^*w, w)_L \geq \mu_0 \|w\|_L^2 - \tfrac{1}{2}\langle Dw, w\rangle_{W',W}. \tag{2.12}$$

Proof. (2.11) is derived by summing (A4) and (2.8). (2.12) is obtained by subtracting (2.8) to (A4). □

The key hypothesis introduced by Friedrichs to select boundary conditions consists of assuming that there exists a matrix-valued field at the boundary, say $\mathcal{M} : \partial\Omega \longrightarrow \mathbb{R}^{m,m}$, such that, a.e. on $\partial\Omega$,

$$\mathcal{M} \text{ is positive, i.e., } (\mathcal{M}\xi, \xi)_{\mathbb{R}^m} \geq 0 \text{ for all } \xi \text{ in } \mathbb{R}^m, \tag{2.13}$$

$$\mathbb{R}^m = \mathrm{Ker}(\mathcal{D} - \mathcal{M}) + \mathrm{Ker}(\mathcal{D} + \mathcal{M}), \tag{2.14}$$

where $\mathcal{D}$ is defined in Remark 2.1. Then, it is possible to prove uniqueness of the so-called strong solution $u \in [\mathfrak{C}^1(\overline{\Omega})]^m$ of the PDE system $Tu = f$ supplemented with the boundary condition $(\mathcal{D} - \mathcal{M})u|_{\partial\Omega} = 0$. Moreover, it is also possible to prove existence of a weak solution in L, namely of a function $u \in L$ such that the relation $(u, T^*v)_L = (f, v)_L$ holds for all $v \in [\mathfrak{C}^1(\overline{\Omega})]^m$ such that $(\mathcal{D} + \mathcal{M}^t)v|_{\partial\Omega} = 0$; see [24]. In this paper, we want to investigate the bijectivity of T in a subspace V of the graph W, and it is not possible to set $V = \{v \in W;\, (\mathcal{D} - \mathcal{M})v|_{\partial\Omega} = 0\}$ since the meaning traces should be given is not clear.

To overcome this difficulty, we modify Friedrichs' hypothesis by the following assumption: there exists an operator $M \in \mathcal{L}(W; W')$ such that

$$M \text{ is positive, i.e., } \langle Mw, w\rangle_{W',W} \geq 0 \text{ for all } w \text{ in } W, \tag{M1}$$

$$W = \mathrm{Ker}(D - M) + \mathrm{Ker}(D + M), \tag{M2}$$

$$W = \mathrm{Ker}(D - M^*) + \mathrm{Ker}(D + M^*). \tag{M3}$$

Here, $M^* \in \mathcal{L}(W; W')$ is the adjoint operator of M defined as follows: for all $(u, v) \in W \times W$, $\langle M^*u, v\rangle_{W',W} = \langle Mv, u\rangle_{W',W}$. Clearly, (M2) and (M3) imply $\mathrm{Im}(D) = \mathrm{Im}(M) = \mathrm{Im}(M^*)$; hence, $\mathrm{Ker}(D) = [\mathrm{Im}(D)]^\perp = [\mathrm{Im}(M^*)]^\perp = \mathrm{Ker}(M)$. As a result, $[\mathfrak{D}(\Omega)]^m \subset \mathrm{Ker}(M)$, i.e., M is a boundary operator. For all subsets $Z \subset W'$, $Z^\perp$ denotes the polar set of Z, i.e., the set of the continuous linear forms in $W'' \equiv W$ that are zero on Z.

Remark 2.2. Assumption (M3) does not appear as such in Friedrichs' formalism. Indeed, proceeding as in [24, p. 356], one can verify that assumptions (M1) and (M2) imply $\mathrm{Im}(D - M^*) \cap \mathrm{Im}(D + M^*) = \{0\}$. In finite dimension, this readily yields (M3). Note also that (M3) can be a consequence of (M2) in some particular cases, e.g., whenever M is self-adjoint ($M = M^*$) or whenever $M + M^* = 0$.

Set

$$V = \mathrm{Ker}(D - M) \qquad \text{and} \qquad V^* = \mathrm{Ker}(D + M^*), \tag{2.15}$$

and equip V and V^* with the graph norm (2.4). Owing to (M1), it is clear that

$$\forall w \in V, \qquad \langle Dw, w\rangle_{W',W} \geq 0, \tag{2.16}$$

$$\forall w^* \in V^*, \qquad \langle Dw^*, w^*\rangle_{W',W} \leq 0. \tag{2.17}$$

As a consequence of the above definitions and hypotheses we infer the following:

LEMMA 2.4. *Assume* (A4) *and* (M1)-(M3). *Let* V *and* V^* *be defined in* (2.15). *Then,*

(i) *T is L-coercive on V and T^* is L-coercive on V^*.*
(ii) $\mathrm{Ker}(D) \subset V \cap V^*$.
(iii) *$D(V)^\perp = V^*$ and $D(V^*)^\perp = V$.*

Proof. (i) Direct consequence of Lemma 2.3 and of (2.16)–(2.17).
(ii) Owing to (A3), (M2), and (M3), we infer $\mathrm{Ker}(D) = \mathrm{Ker}(M) = \mathrm{Ker}(M^*)$. Then it is clear that $\mathrm{Ker}(D) \subset \mathrm{Ker}(D-M) = V$ and that $\mathrm{Ker}(D) \subset \mathrm{Ker}(D+M^*) = V^*$.
(iii) Let us prove that $D(V)^\perp \subset V^*$. Let $w \in D(V)^\perp$. Let $z \in W$. Owing to (M2), set $z = z^+ + z^-$ with $z^\pm \in \mathrm{Ker}(D \pm M)$. Then,

$$\begin{aligned}\langle (D+M^*)w, z\rangle_{W',W} &= \langle (D+M^*)w, z^+\rangle_{W',W} + \langle (D+M^*)w, z^-\rangle_{W',W}\\ &= \langle (D+M)z^-, w\rangle_{W',W} = 2\langle Dz^-, w\rangle_{W',W} = 0,\end{aligned}$$

since $z^- \in V$ and $w \in D(V)^\perp$. As a result, $w \in \mathrm{Ker}(D+M^*)$. Hence, $D(V)^\perp \subset V^*$. Conversely, let $w \in V^*$. Let $v \in V$. Using the fact that $Dv = Mv$ yields

$$\langle Dv, w\rangle_{W',W} = \tfrac{1}{2}\langle (D+M)v, w\rangle_{W',W} = \tfrac{1}{2}\langle (D+M^*)w, v\rangle_{W',W} = 0,$$

i.e., $w \in D(V)^\perp$. Hence, $V^* \subset D(V)^\perp$. Proceed similarly to prove $D(V^*)^\perp = V$. ∎

THEOREM 2.5. *Assume (A4) and (M1)–(M3). Let V and V^* be defined in (2.15). Then,*
(i) *$T : V \to L$ is an isomorphism.*
(ii) *$T^* : V^* \to L$ is an isomorphism.*

Proof. We only prove (i) since the proof of (ii) is similar.
(1) Owing to (2.15), V is closed in W; hence, V is a Hilbert space. As a result, showing that $T : V \to L$ is an isomorphism amounts to proving statement (ii) in Theorem 2.6 below.
(2) Proof of (2.20). Let $u \in V$. Then, Lemma 2.4(i) implies $\sup_{v\in L\setminus\{0\}} \frac{(Tu,v)_L}{\|v\|_L} \geq \mu_0\|u\|_L$. Furthermore,

$$\begin{aligned}\sup_{v\in L\setminus\{0\}} \frac{(Tu,v)_L}{\|v\|_L} &\geq \sup_{v\in L\setminus\{0\}} \frac{(Au,v)_L}{\|v\|_L} - \|K\|_{\mathcal{L}(L;L)}\|u\|_L\\ &\geq \|Au\|_L - \frac{\|K\|_{\mathcal{L}(L;L)}}{\mu_0} \sup_{v\in L\setminus\{0\}} \frac{(Tu,v)_L}{\|v\|_L}.\end{aligned}$$

Hence,

$$\left(1 + \frac{1+\|K\|_{\mathcal{L}(L;L)}}{\mu_0}\right) \sup_{v\in L\setminus\{0\}} \frac{(Tu,v)_L}{\|v\|_L} \geq \|Au\|_L + \|u\|_L \geq \|u\|_W.$$

(3) Proof of (2.21). Assume that $v \in L$ is such that $(Tu,v)_L = 0$ for all $u \in V$. Owing to Lemma 2.4(ii) and the fact that $[\mathfrak{D}(\Omega)]^m \subset \mathrm{Ker}(D)$, it is clear that $[\mathfrak{D}(\Omega)]^m \subset V$. As a result, a standard distribution argument shows that

$$T^*v = 0, \tag{2.18}$$

in $[\mathfrak{D}'(\Omega)]^m$, where T^* is defined in (2.6). Still in the distribution sense, this means that

$$\sum_{k=1}^d \mathcal{A}^k \partial_k v = K^*v - (\nabla\cdot A)v.$$

Since the right-hand side is a bounded linear functional on L, v has a A-weak derivative in L, i.e., $v \in W$. Using (2.8), together with (2.18), yields

$$\forall u \in V, \quad \langle Du, v\rangle_{W',W} = 0, \tag{2.19}$$

i.e., $v \in D(V)^{\perp}$. Owing to Lemma 2.4(iii), $v \in V^*$. Finally, since $(T^*v, v)_L = 0$ and $v \in V^*$, Lemma 2.4(i) implies that v is zero. □

THEOREM 2.6 (Banach–Nečas–Babuška (BNB)). *Let V be a Banach space and let L be a reflexive Banach space. The following statements are equivalent:*

(i) *$T \in \mathcal{L}(V; L)$ is bijective.*

(ii) *There exists a constant $\alpha > 0$ such that*

$$\forall u \in V, \quad \sup_{v \in L\setminus\{0\}} \frac{(Tu, v)_L}{\|v\|_L} \geq \alpha \|u\|_V, \tag{2.20}$$

$$\forall v \in L, \quad ((Tu, v)_L = 0, \ \forall u \in V) \implies (v = 0). \tag{2.21}$$

Remark 2.3.

(i) As an immediate consequence of Theorem 2.5(ii), the following problem is also well-posed: For f in L,

$$\begin{cases} \text{Seek } u^* \in V^* \text{ such that} \\ T^*u^* = f. \end{cases} \tag{2.22}$$

(ii) To guarantee that $T : V \to L$ and $T^* : V^* \to L$ are isomorphisms, it is also possible to specify assumptions on the spaces V and V^* without using the boundary operator M. Introduce the cones

$$C^+ = \{w \in W;\ \langle Dw, w\rangle_{W',W} \geq 0\}, \tag{2.23}$$

$$C^- = \{w \in W;\ \langle Dw, w\rangle_{W',W} \leq 0\}. \tag{2.24}$$

Then, one can verify that under the following assumptions:

$$V \subset C^+ \text{ and } V^* \subset C^-, \tag{V1}$$

$$V^* = D(V)^{\perp} \text{ and } V = D(V^*)^{\perp}, \tag{V2}$$

$T : V \to L$ and $T^* : V^* \to L$ are isomorphisms. This way of introducing Friedrichs' symmetric systems seems to be new. We think that assumptions (V1)–(V2) are more natural than (M1)–(M3) since they do not involve the somewhat ad hoc operator M. Note that (V1)–(V2) imply that V and V^* are closed in W and that $\mathrm{Ker}(D) \subset V \cap V^*$.

2.3. Boundary conditions weakly enforced. As we have in mind to solve (2.9) by means of DG methods with the boundary conditions weakly enforced, we now propose an alternative formulation of (2.9) and of (2.22). Define the bilinear forms

$$a(u, v) = (Tu, v)_L + \tfrac{1}{2}\langle (M - D)u, v\rangle_{W',W}, \tag{2.25}$$

$$a^*(u, v) = (T^*u, v)_L + \tfrac{1}{2}\langle (M^* + D)u, v\rangle_{W',W}. \tag{2.26}$$

It is clear that a and a^* are in $\mathcal{L}(W \times W; \mathbb{R})$. A remarkable property is the following:

LEMMA 2.7. *Under assumption* (A4)*, the following holds for all* $w \in W$*,*

$$a(w,w) \geq \mu_0 \|w\|_L^2 + \tfrac{1}{2}\langle Mw, w\rangle_{W',W}, \tag{2.27}$$

$$a^*(w,w) \geq \mu_0 \|w\|_L^2 + \tfrac{1}{2}\langle Mw, w\rangle_{W',W}. \tag{2.28}$$

As a result, a *and* a^* *are* L*-coercive on* W *whenever* (M1) *holds.*

Proof. Let $w \in W$. Owing to (2.8),

$$\begin{aligned} a(w,w) &= (Tw,w)_L - \tfrac{1}{2}\langle Dw, w\rangle_{W',W} + \tfrac{1}{2}\langle Mw, w\rangle_{W',W} \\ &= \tfrac{1}{2}((T+T^*)w,w)_L + \tfrac{1}{2}\langle Mw, w\rangle_{W',W}. \end{aligned}$$

Hence, (2.27) follows from (A4). The proof of (2.28) is similar. ∎

Consider the following problems: For $f \in L$,

$$\begin{cases} \text{Seek } u \in W \text{ such that} \\ a(u,v) = (f,v)_L, \quad \forall v \in W, \end{cases} \tag{2.29}$$

and

$$\begin{cases} \text{Seek } u^* \in W \text{ such that} \\ a^*(u^*,v) = (f,v)_L, \quad \forall v \in W. \end{cases} \tag{2.30}$$

THEOREM 2.8. *Assume* (A4) *and* (M1)–(M3)*. Then,*

(i) *there is a unique solution to* (2.29) *and this solution solves* (2.9)*;*

(ii) *there is a unique solution to* (2.30) *and this solution solves* (2.22)*.*

Proof. (i) Owing to Theorem 2.5, there is a unique $u \in V$ solving $Tu = f$. Moreover, since u is in V, $(D-M)u = 0$. Hence, $a(u,v) = (f,v)_L$ for all $v \in W$, i.e., u solves (2.29). In addition, since a is L-coercive on W owing to Lemma 2.7, it is clear that the solution to (2.29) is unique.

(ii) The proof of the second statement is similar. ∎

Remark 2.4. Neither the bilinear form a nor the bilinear form a^* induce an isomorphism between W and W'. In particular, there is no guarantee that (2.29) or (2.30) has a solution if the right-hand is replaced by $\langle f, v\rangle_{W',W}$ whenever $f \in W'$.

3. Examples. This section discusses important examples of Friedrichs' symmetric systems: advection–reaction equations, advection–diffusion–reaction equations, and a simplified version of the Maxwell equations in the diffusive regime.

3.1. Advection–reaction. Let β be a vector field in $\mathbb{R}^d$, assume $\beta \in [L^\infty(\Omega)]^d$, $\nabla\cdot\beta \in L^\infty(\Omega)$, and define

$$\partial\Omega^- = \{x \in \partial\Omega;\, \beta(x)\cdot n(x) < 0\}, \tag{3.1}$$

$$\partial\Omega^+ = \{x \in \partial\Omega;\, \beta(x)\cdot n(x) > 0\}, \tag{3.2}$$

$$\partial\Omega^0 = \partial\Omega \backslash (\overline{\partial\Omega^-} \cup \overline{\partial\Omega^+}). \tag{3.3}$$

$\partial\Omega^-$ is the inflow boundary and $\partial\Omega^+$ is the outflow boundary. $\partial\Omega^0$ is the interior of the set $\{x \in \partial\Omega;\, \beta(x)\cdot n(x) = 0\}$.

Let μ be a function in $L^\infty(\Omega)$, and consider the advection–reaction equation

$$\mu u + \beta\cdot\nabla u = f. \tag{3.4}$$

This PDE falls into the category studied above by setting $Kv = \mu v$ for all $v \in L^2(\Omega)$, and $\mathcal{A}^k = \beta^k$ for $k \in \{1, \ldots, d\}$. It is clear that (A1)–(A3) hold with $m = 1$. Define the graph space

$$W = \{w \in L^2(\Omega);\, \beta\cdot\nabla w \in L^2(\Omega)\} \subset L^2(\Omega). \tag{3.5}$$

W is a Hilbert space when equipped with the norm $\|w\|_W = \|w\|_{L^2(\Omega)} + \|\beta\cdot\nabla w\|_{L^2(\Omega)}$. Define the differential operators

$$T : W \ni u \longmapsto \mu u + \beta\cdot\nabla u \in L^2(\Omega), \tag{3.6}$$

$$T^* : W \ni u \longmapsto \mu u - \nabla\cdot(\beta u) \in L^2(\Omega). \tag{3.7}$$

It is clear that T and T^* are continuous.

Without additional hypotheses on Ω, μ, and β, the operator T is unlikely to be an isomorphism (think of $\partial\Omega^- = \partial\Omega^+ = \emptyset$ and $\mu = 0$). Henceforth, we assume

$$\mathfrak{C}^1(\overline{\Omega}) \text{ is dense in } W, \tag{H1}$$

$$\partial\Omega^- \text{ and } \partial\Omega^+ \text{ are well-separated, i.e., } \operatorname{dist}(\partial\Omega^-, \partial\Omega^+) > 0, \tag{H2}$$

$$\mu(x) - \tfrac{1}{2}\nabla\cdot\beta(x) \geq \mu_0 > 0 \quad \text{a.e. in } \Omega. \tag{H3}$$

Hypothesis (H3) implies that (A4) holds. Hypothesis (H1) is a regularity assumption on Ω. It can be shown to hold by using Friedrichs' mollifier whenever Ω is smooth.

Let $L^2(\partial\Omega; |\beta\cdot n|)$ be the space of real-valued functions that are square integrable with respect to the measure $|\beta\cdot n|\,\mathrm{d}x$ where $\mathrm{d}x$ is the Lebesgue measure on $\partial\Omega$.

LEMMA 3.1. *Provided* (H1)–(H2) *hold, the trace operator* $\gamma : \mathfrak{C}^1(\overline{\Omega}) \ni v \longrightarrow v \in L^2(\partial\Omega; |\beta\cdot n|)$ *extends uniquely to a continuous operator on* W.

Proof. Since $\partial\Omega^-$ and $\partial\Omega^+$ are well-separated, there are two non-negative functions ψ^- and ψ^+ in $\mathfrak{C}^1(\overline{\Omega})$ such that

$$\psi^- + \psi^+ = 1 \ \text{ on } \Omega, \qquad \psi^-|_{\partial\Omega^+} = 0, \qquad \psi^+|_{\partial\Omega^-} = 0. \tag{3.8}$$

Let u be a function in $\mathfrak{C}^1(\overline{\Omega})$. Then,

$$\begin{aligned}
\int_{\partial\Omega} u^2|\beta\cdot n| &= \int_{\partial\Omega} u^2(\psi^- + \psi^+)|\beta\cdot n| = \int_{\partial\Omega^-\cup\partial\Omega^0} u^2\psi^-|\beta\cdot n| + \int_{\partial\Omega^+\cup\partial\Omega^0} u^2\psi^+|\beta\cdot n| \\
&= -\int_{\partial\Omega} u^2\psi^-(\beta\cdot n) + \int_{\partial\Omega} u^2\psi^+(\beta\cdot n) \\
&= -\int_{\Omega} \nabla\cdot(u^2\psi^-\beta) + \int_{\Omega} \nabla\cdot(u^2\psi^+\beta).
\end{aligned}$$

Hence,

$$0 \leq \int_{\partial\Omega} u^2|\beta\cdot n| \leq c(\psi^+, \psi^-)\|u\|_W^2.$$

The result follows from the density of $\mathfrak{C}^1(\overline{\Omega})$ in W. ∎

As an immediate consequence of the existence of traces in $L^2(\partial\Omega; |\beta\cdot n|)$, we deduce

COROLLARY 3.2. *Under the hypotheses* (H1)–(H2)*, the operator* D *has the following representation*

$$\forall (u, v) \in W \times W, \quad \langle Du, v\rangle_{W',W} = \int_{\partial\Omega} uv(\beta\cdot n). \tag{3.9}$$

To specify boundary conditions, define for $(u,v) \in W \times W$,

$$\langle Mu, v\rangle_{W',W} = \int_{\partial\Omega} uv|\beta\cdot n|. \tag{3.10}$$

LEMMA 3.3. *Let $M \in \mathcal{L}(W;W')$ be defined in* (3.10)*. Then,* (M1)–(M3) *hold.*

Proof. (1) (M1) directly results from (3.10).
(2) Let ψ^+, ψ^- be the partition of unity introduced in (3.8). Let $w \in W$ and write $w = \psi^+ w + \psi^- w$. It is clear that $\psi^+ w \in \mathrm{Ker}(D-M)$ since for all $v \in W$,

$$\langle (D-M)\psi^+ w, v\rangle_{W',W} = \int_{\partial\Omega} \psi^+ vw(\beta\cdot n - |\beta\cdot n|) = \int_{\partial\Omega^+} \psi^+ vw(\beta\cdot n - |\beta\cdot n|) = 0.$$

Similarly, $\psi^- w \in \mathrm{Ker}(D+M)$. Hence, (M2) holds.
(3) (M3) directly results from (M2) since M is self-adjoint. □

LEMMA 3.4. *Let $M \in \mathcal{L}(W;W')$ be defined in* (3.10)*. Set $V = \mathrm{Ker}(D-M)$ and $V^* = \mathrm{Ker}(D+M^*)$. Then,*

$$V = \{v \in W;\, v|_{\partial\Omega^-} = 0\}, \tag{3.11}$$

$$V^* = \{v \in W;\, v|_{\partial\Omega^+} = 0\}. \tag{3.12}$$

Proof. (1) Let $v \in \mathrm{Ker}(D-M)$. Then, for all $w \in W$, $-2\int_{\partial\Omega^-} |\beta\cdot n| vw = 0$. Take $w = v$ to infer $v|_{\partial\Omega^-} = 0$.
(2) Conversely, if $v|_{\partial\Omega^-} = 0$, it is clear that for all $w \in W$, $\langle (D-M)v, w\rangle_{W',W} = -2\int_{\partial\Omega^-} |\beta\cdot n| vw = 0$, i.e., $v \in \mathrm{Ker}(D-M)$. Hence, (3.11) holds.
(3) Proceed similarly to prove (3.12). □

As a consequence of Theorem 2.5, we deduce

PROPOSITION 3.5. *Assume that* (H1)–(H3) *hold. Let V and V^* be defined in* (3.11) *and* (3.12)*, respectively. Then, $T : V \to L^2(\Omega)$ and $T^* : V^* \to L^2(\Omega)$ are isomorphisms.*

3.2. Advection–diffusion–reaction equations. Let $\beta : \Omega \longrightarrow \mathbb{R}^d$ be a vector field such that $\beta \in [L^\infty(\Omega)]^d$ and $\nabla\cdot\beta \in L^\infty(\Omega)$. Let μ be a function in $L^\infty(\Omega)$, and consider the advection–diffusion–reaction equation

$$-\Delta u + \beta\cdot\nabla u + \mu u = f. \tag{3.13}$$

This equation can be written as a system of first-order PDEs by setting

$$\begin{cases} \sigma + \nabla u = 0, \\ \mu u + \nabla\cdot\sigma + \beta\cdot\nabla u = f. \end{cases} \tag{3.14}$$

The above differential operator can be cast into the form of a symmetric Friedrichs' operator by setting $K(\sigma,u) = (\sigma,\mu u)$ for all $(\sigma,u) \in [L^2(\Omega)]^{d+1}$, and

$$\mathcal{A}^k = \left[\begin{array}{c:c} 0 & e^k \\ \hdashline (e^k)^t & \beta^k \end{array}\right], \tag{3.15}$$

where e^k is the k-th vector in the canonical basis of $\mathbb{R}^d$. It is clear that hypotheses (A1)–(A3) hold with $m = d+1$. Moreover, under the assumption (H3), one readily checks that (A4) holds.

Upon observing the norm equivalence

$$c_1(\|\nabla u\|_{L^2(\Omega)} + \|\nabla\cdot\sigma\|_{L^2(\Omega)}) \le \|\nabla u\|_{L^2(\Omega)} + \|\beta\cdot\nabla u - \nabla\cdot\sigma\|_{L^2(\Omega)} \le c_2(\|\nabla u\|_{L^2(\Omega)} + \|\nabla\cdot\sigma\|_{L^2(\Omega)}),$$

the graph space

$$W = \{(\sigma,u) \in [L^2(\Omega)]^{d+1};\, A(\sigma,u) \in [L^2(\Omega)]^{d+1}\} \tag{3.16}$$

is a Hilbert space when equipped with the norm

$$\|(\sigma,u)\|_W = \|(\sigma,u)\|_{[L^2(\Omega)]^{d+1}} + \|\nabla u\|_{[L^2(\Omega)]^d} + \|\nabla\cdot\sigma\|_{L^2(\Omega)}. \tag{3.17}$$

In other words, $W = H(\mathrm{div};\Omega)\times H^1(\Omega)$.

Now, define the differential operators

$$T : W \ni (\sigma,u) \longmapsto (\sigma + \nabla u, \mu u + \nabla\cdot\sigma + \beta\cdot\nabla u) \in [L^2(\Omega)]^d \times L^2(\Omega), \tag{3.18}$$

$$T^* : W \ni (\tau,v) \longmapsto (\tau - \nabla v, \mu v - \nabla\cdot\tau - \nabla\cdot(\beta v)) \in [L^2(\Omega)]^d \times L^2(\Omega), \tag{3.19}$$

and observe that, for all $(\sigma,u),(\tau,v)\in W$,

$$\langle D(\sigma,u),(\tau,v)\rangle_{W',W} = \langle\sigma\cdot n, v\rangle_{-\frac12,\frac12} + \langle\tau\cdot n, u\rangle_{-\frac12,\frac12} + \int_{\partial\Omega}(\beta\cdot n)uv, \tag{3.20}$$

where $\langle,\rangle_{-\frac12,\frac12}$ denotes the duality pairing between $H^{-\frac12}(\partial\Omega)$ and $H^{\frac12}(\partial\Omega)$. Note that (3.20) makes sense since functions in $H^1(\Omega)$ have traces in $H^{\frac12}(\partial\Omega)$ and vector fields in $H(\mathrm{div};\Omega)$ have normal traces in $H^{-\frac12}(\partial\Omega)$.

3.2.1. Dirichlet boundary conditions. A suitable operator M to weakly enforce Dirichlet boundary conditions is such that, for all $(\sigma,u),(\tau,v)\in W$,

$$\langle M(\sigma,u),(\tau,v)\rangle_{W',W} = \langle\sigma\cdot n, v\rangle_{-\frac12,\frac12} - \langle\tau\cdot n, u\rangle_{-\frac12,\frac12}. \tag{3.21}$$

LEMMA 3.6. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.21). *Then,* (M1)–(M3) *hold.*

Proof. (1) (M1) clearly holds since $M + M^* = 0$.
(2) Let $w = (\sigma,u)\in W$ and write $w = w^+ + w^-$ with $w^+ = (-\frac12\beta u, u)$ and $w^- = (\sigma + \frac12\beta u, 0)$. By assumption on β, the vector-valued field βu is in $H(\mathrm{div};\Omega)$ if $u\in H^1(\Omega)$; hence, $w^\pm$ are in W. Moreover, a straightforward calculation shows that $w^\pm\in\mathrm{Ker}(D\pm M)$. Hence, (M2) holds.
(3) (M3) results from (M2) and the fact that $M + M^* = 0$. □

LEMMA 3.7. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.21). *Set $V = \mathrm{Ker}(D-M)$ and $V^* = \mathrm{Ker}(D+M^*)$. Then,*

$$V = V^* = \{(\sigma,u)\in W;\, u|_{\partial\Omega} = 0\}. \tag{3.22}$$

Proof. (1) The identity $V = V^*$ results from the fact that $M + M^* = 0$.
(2) Let $(\sigma,u)\in\mathrm{Ker}(D-M)$. Then, for all $(\tau,v)\in W$,

$$2\langle\tau\cdot n, u\rangle_{-\frac12,\frac12} + \int_{\partial\Omega}(\beta\cdot n)uv = 0.$$

Let $\gamma\in H^{-\frac12}(\partial\Omega)$. There exists $\tau\in H(\mathrm{div};\Omega)$ such that $\tau\cdot n = \gamma$ in $H^{-\frac12}(\partial\Omega)$. Then, using $(\tau,0)$ in the above equation yields $\langle\gamma,u\rangle_{-\frac12,\frac12} = 0$. Since γ is arbitrary, this

implies $u|_{\partial\Omega}=0$. Hence, $V\subset\{(\sigma,u)\in W;\, u|_{\partial\Omega}=0\}$.
(3) Conversely, let $(\sigma,u)\in W$ be such that $u|_{\partial\Omega}=0$. Then, for all $(\tau,v)\in W$,

$$\langle (D-M)(\sigma,u),(\tau,v)\rangle_{W',W}=2\langle\tau{\cdot}n,u\rangle_{-\frac12,\frac12}+\int_{\partial\Omega}(\beta{\cdot}n)uv=0,$$

i.e., $(\sigma,u)\in\mathrm{Ker}(D-M)=V$. The proof is complete. □

As a consequence of Theorem 2.5, we deduce

PROPOSITION 3.8. *Assume* (H1) *and* (H3). *Let V be defined in* (3.22). *Then, $T:V\to[L^2(\Omega)]^d\times L^2(\Omega)$ and $T^*:V\to[L^2(\Omega)]^d\times L^2(\Omega)$ are isomorphisms.*

Remark 3.1. The choice of the operator M is not unique. For instance, assume there exists $\psi\in[L^\infty(\Omega)]^d$ with $\nabla{\cdot}\psi\in L^\infty(\Omega)$ such that $\psi{\cdot}n|_{\partial\Omega}=1$ (a sufficient condition for ψ to exist is Ω be Lipschitz). Then, denoting by ς a non-negative function in $L^\infty(\partial\Omega)$, the operator M defined by

$$\langle M(\sigma,u),(\tau,v)\rangle_{W',W}=\langle\sigma{\cdot}n,v\rangle_{-\frac12,\frac12}-\langle\tau{\cdot}n,u\rangle_{-\frac12,\frac12}+\int_{\partial\Omega}\varsigma uv,$$

can be used to enforce Dirichlet boundary conditions weakly. Indeed, (M1) clearly holds whereas (M2) and (M3) result from the following identities

$$\begin{aligned}&(\sigma+\tfrac12(\beta+\varsigma\psi)u,0)\in\mathrm{Ker}(D-M),\qquad &&(-\tfrac12(\beta+\varsigma\psi)u,u)\in\mathrm{Ker}(D+M),\\&(-\tfrac12(\beta-\varsigma\psi)u,u)\in\mathrm{Ker}(D-M^*),\qquad &&(\sigma+\tfrac12(\beta-\varsigma\psi)u,0)\in\mathrm{Ker}(D+M^*).\end{aligned}$$

Moreover, one readily verifies that V and V^* are still given by (3.22).

3.2.2. Neumann boundary conditions. To simplify, assume $\beta{\cdot}n|_{\partial\Omega}=0$. A suitable operator M to enforce Neumann boundary conditions weakly is such that, for all $(\sigma,u),(\tau,v)\in W$,

$$\langle M(\sigma,u),(\tau,v)\rangle_{W',W}=\langle\tau{\cdot}n,u\rangle_{-\frac12,\frac12}-\langle\sigma{\cdot}n,v\rangle_{-\frac12,\frac12}.\tag{3.23}$$

LEMMA 3.9. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.23). *Then,* (M1)–(M3) *hold.*

Proof. (1) (M1) clearly holds since $M+M^*=0$.
(2) Let $w=(\sigma,u)\in W$ and write $w=w^+ + w^-$ with $w^+=(\sigma,0)$ and $w^-=(0,u)$. It is readily verified that $w^\pm\in\mathrm{Ker}(D\pm M)$. Hence, (M2) holds.
(3) (M3) results from (M2) and the fact that $M+M^*=0$. □

LEMMA 3.10. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.23). *Set $V=\mathrm{Ker}(D-M)$ and $V^*=\mathrm{Ker}(D+M^*)$. Then,*

$$V=V^*=\{(\sigma,u)\in W;\,(\sigma{\cdot}n)|_{\partial\Omega}=0\}.\tag{3.24}$$

Proof. (1) The identity $V=V^*$ results from the fact that $M+M^*=0$.
(2) Let $(\sigma,u)\in\mathrm{Ker}(D-M)$. Then, for all $v\in H^1(\Omega)$, $\langle\sigma{\cdot}n,v\rangle_{-\frac12,\frac12}=0$. Since v is arbitrary and traces are surjective from $H^1(\Omega)$ to $H^{\frac12}(\partial\Omega)$, it comes that $(\sigma{\cdot}n)|_{\partial\Omega}=0$.
(3) Conversely, if $(\sigma{\cdot}n)|_{\partial\Omega}=0$, it is clear that $(\sigma,u)\in\mathrm{Ker}(D-M)$. The proof is complete. □

As a consequence of Theorem 2.5, we deduce

PROPOSITION 3.11. *Assume* (H1) *and* (H3) *and that $\beta{\cdot}n|_{\partial\Omega}=0$. Let V be defined in* (3.24). *Then, $T:V\to[L^2(\Omega)]^d\times L^2(\Omega)$ and $T^*:V\to[L^2(\Omega)]^d\times L^2(\Omega)$ are isomorphisms.*

3.2.3. Robin boundary conditions. As in §3.2.2, assume $\beta{\cdot}n|_{\partial\Omega}=0$. Let ϱ be a non-negative function in $L^\infty(\partial\Omega)$. A suitable operator M to weakly enforce Robin boundary conditions is such that, for all $(\sigma,u),(\tau,v)\in W$,

$$\langle M(\sigma,u),(\tau,v)\rangle_{W',W}=\langle\tau{\cdot}n,u\rangle_{-\frac12,\frac12}-\langle\sigma{\cdot}n,v\rangle_{-\frac12,\frac12}+2\int_{\partial\Omega}\varrho uv. \tag{3.25}$$

LEMMA 3.12. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.25)*. Then,* (M1)–(M3) *hold.*

Proof. (1) To prove (M1), observe that for all $(\sigma,u)\in W$, $\langle M(\sigma,u),(\sigma,u)\rangle_{W',W}=2\int_{\partial\Omega}\varrho u^2\ge0$.
(2) Let $w=(\sigma,u)\in W$. Since Ω is Lipschitz, there exists $\psi\in[L^\infty(\Omega)]^d$ with $\nabla{\cdot}\psi\in L^\infty(\Omega)$ such that $\psi{\cdot}n|_{\partial\Omega}=1$. Then, write $w=w^+{+}w^-$ with $w^+=(\sigma{-}\varrho u\psi,0)$ and $w^-=(\varrho u\psi,u)$. It is readily verified that $w^\pm\in\mathrm{Ker}(D\pm M)$. Hence, (M2) holds.
(3) To prove (M3), use the decomposition $w=w^++w^-$ with $w^+=(-\varrho u\psi,u)$ and $w^-=(\sigma+\varrho u\psi,0)$; then $w^\pm\in\mathrm{Ker}(D\pm M^*)$. The proof is complete. □

LEMMA 3.13. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.25)*. Set $V=\mathrm{Ker}(D-M)$ and $V^*=\mathrm{Ker}(D+M^*)$. Then,*

$$V=\{(\sigma,u)\in W;\ -(\sigma{\cdot}n)|_{\partial\Omega}+\varrho u|_{\partial\Omega}=0\}, \tag{3.26}$$

$$V^*=\{(\sigma,u)\in W;\ (\sigma{\cdot}n)|_{\partial\Omega}+\varrho u|_{\partial\Omega}=0\}. \tag{3.27}$$

Proof. (1) Let $(\sigma,u)\in\mathrm{Ker}(D-M)$. Then, for all $v\in H^1(\Omega)$, $2\langle\sigma{\cdot}n,v\rangle_{-\frac12,\frac12}-2\int_{\partial\Omega}\varrho uv=0$. Since v is arbitrary and traces are surjective from $H^1(\Omega)$ to $H^{\frac12}(\partial\Omega)$, it comes that $(\sigma{\cdot}n)|_{\partial\Omega}-\varrho u|_{\partial\Omega}=0$.
(2) Conversely, if $(\sigma{\cdot}n)|_{\partial\Omega}-\varrho u|_{\partial\Omega}=0$, it is clear that $(\sigma,u)\in\mathrm{Ker}(D-M)$.
(3) Proceed similarly to prove (3.27). □

As a consequence of Theorem 2.5, we deduce

PROPOSITION 3.14. *Assume* (H1) *and* (H3) *and that $\beta{\cdot}n|_{\partial\Omega}=0$. Let V and V^* be defined in* (3.26) *and* (3.27)*, respectively. Then, $T:V\to[L^2(\Omega)]^d\times L^2(\Omega)$ and $T^*:V^*\to[L^2(\Omega)]^d\times L^2(\Omega)$ are isomorphisms.*

3.3. Maxwell's equations in diffusive regime. We close this series of examples by considering a simplified form of Maxwell's equations in $\mathbb{R}^3$ in the diffusive regime, i.e., when displacement currents are negligible. Let σ and μ be two positive functions in $L^\infty(\Omega)$ uniformly bounded away from zero. Consider the following problem

$$\begin{cases}\mu H+\nabla{\times}E=f,\\ \sigma E-\nabla{\times}H=g.\end{cases} \tag{3.28}$$

This problem can be cast into the form of a symmetric Friedrichs system by setting $K(H,E)=(\mu H,\sigma E)$ for all $(H,E)\in[L^2(\Omega)]^3\times[L^2(\Omega)]^3$ and by introducing the matrices $\mathcal{A}^k\in\mathbb{R}^{6,6}$ given by

$$\mathcal{A}^k=\left[\begin{array}{c:c}0&\mathcal{R}^k\\ \hdashline (\mathcal{R}^k)^t&0\end{array}\right],\quad 1\le k\le3. \tag{3.29}$$

The entries of the matrices $\mathcal{R}^k\in\mathbb{R}^{3,3}$ are those of the Levi-Civita permutation tensor, i.e., $\mathcal{R}^k_{ij}=\epsilon_{ikj}$ for $1\le i,j,k\le3$.

Define $W = H(\mathrm{curl};\Omega)\times H(\mathrm{curl};\Omega)$ and equip W with the corresponding norm. It is clear that the following differential operators are associated with (3.28):

$$T : W \ni (H,E) \longmapsto (\mu H + \nabla\times E, \sigma E - \nabla\times H) \in [L^2(\Omega)]^3\times[L^2(\Omega)]^3. \tag{3.30}$$

$$T^* : W \ni (H,E) \longmapsto (\mu H - \nabla\times E, \sigma E + \nabla\times H) \in [L^2(\Omega)]^3\times[L^2(\Omega)]^3. \tag{3.31}$$

Hypotheses (A1)–(A3) obviously hold with $m = 6$. Hypothesis (A4) is a consequence of the fact that σ and μ are positive functions in $L^\infty(\Omega)$ uniformly bounded away from zero.

Owing to (2.8), (3.30), and (3.31), the boundary operator D is defined for all $(H,E),(h,e)\in W$ as follows

$$\begin{aligned}\langle D(H,E),(h,e)\rangle_{W',W} = {}& (\nabla\times E,h)_{[L^2(\Omega)]^3} - (E,\nabla\times h)_{[L^2(\Omega)]^3}\\ &+ (H,\nabla\times e)_{[L^2(\Omega)]^3} - (\nabla\times H,e)_{[L^2(\Omega)]^3}.\end{aligned} \tag{3.32}$$

When H and E are smooth the above duality product can be interpreted as the boundary integral $\int_{\partial\Omega}(n\times E)\cdot h + (n\times e)\cdot H$.

Let us now define acceptable boundary conditions for (3.28). One possibility (among many others) consists of setting for all $(H,E),(h,e)\in W$,

$$\begin{aligned}\langle M(H,E),(h,e)\rangle_{W',W} = {}& -(\nabla\times E,h)_{[L^2(\Omega)]^3} + (E,\nabla\times h)_{[L^2(\Omega)]^3}\\ &+ (H,\nabla\times e)_{[L^2(\Omega)]^3} - (\nabla\times H,e)_{[L^2(\Omega)]^3}.\end{aligned} \tag{3.33}$$

LEMMA 3.15. *Let M be defined in* (3.33). *Then,* (M1)–(M3) *hold.*

Proof. (1) Observe that $M + M^* = 0$; hence M is positive.
(2) Let $w = (H,E)\in W$. Write $w = w^+ + w^-$ with $w^+ = (0,E)$ and $w^- = (H,0)$. One easily verifies that $w^\pm \in \mathrm{Ker}(D\pm M)$, i.e., (M2) holds.
(3) (M3) is a consequence of (M2) and the fact that $M + M^* = 0$. □

LEMMA 3.16. *Let $M\in\mathcal{L}(W;W')$ be defined in* (3.33). *Set $V = \mathrm{Ker}(D-M)$ and $V^* = \mathrm{Ker}(D+M^*)$. Then,*

$$V = V^* = \{(H,E)\in W;\ (E\times n)|_{\partial\Omega} = 0\}. \tag{3.34}$$

Proof. (1) The identity $V = V^*$ results from the fact that $M + M^* = 0$.
(2) Let $(H,E)\in\mathrm{Ker}(D-M)$. Then, for all $(h,e)\in W$,

$$\langle (D-M)(H,E),(h,e)\rangle_{W',W} = 2(\nabla\times E,h)_{[L^2(\Omega)]^3} - 2(E,\nabla\times h)_{[L^2(\Omega)]^3} = 0.$$

Since vector fields in $H(\mathrm{curl};\Omega)$ have tangential traces in $[H^{-\frac12}(\partial\Omega)]^3$, we infer that for all $h\in[H^1(\Omega)]^3$, $\langle (E\times n),h\rangle_{-\frac12,\frac12} = 0$. Since h is arbitrary and the traces of vectors fields in $[H^1(\Omega)]^3$ span $[H^{\frac12}(\partial\Omega)]^3$, we conclude that $(E\times n)|_{\partial\Omega} = 0$.
(3) Conversely, let $(H,E)\in W$ be such that $(E\times n)|_{\partial\Omega} = 0$. Then, it is clear that $\langle (D-M)(H,E),(h,e)\rangle_{W',W} = 0$ for all $h\in[H^1(\Omega)]^3$ and all $e\in H(\mathrm{curl};\Omega)$. Since $[H^1(\Omega)]^3$ is dense in $H(\mathrm{curl};\Omega)$ and both D and M are in $\mathcal{L}(W;W')$, it comes that $(H,E)\in\mathrm{Ker}(D-M)$. The proof is complete. □

As a consequence of Theorem 2.5, we deduce

PROPOSITION 3.17. *Let V be defined in* (3.34). *Then, $T : V\to[L^2(\Omega)]^3\times[L^2(\Omega)]^3$ and $T^* : V\to[L^2(\Omega)]^3\times[L^2(\Omega)]^3$ are isomorphisms.*

4. Discontinuous Galerkin. The goal of this section is to introduce a generic DG bilinear form, see (4.12), together with its design constraints, see (DG1) to (DG8). This bilinear form is then used to approximate the abstract problem (2.9). The main convergence result is stated in Theorem 4.6.

4.1. The discrete setting. Let $\{\mathcal{T}_h\}_{h>0}$ be a family of meshes of Ω. The meshes are assumed to be affine to avoid unnecessary technicalities, i.e., Ω is assumed to be a polyhedron. However, we do not make any assumption on the matching of element interfaces.

Let p be a non-negative integer. Define

$$W_h = \{v_h \in [L^2(\Omega)]^m; \forall K \in \mathcal{T}_h,\, v_h|_K \in [\mathbb{P}_p]^m\}, \tag{4.1}$$

$$W(h) = [H^1(\Omega)]^m + W_h. \tag{4.2}$$

We denote by $\mathcal{F}_h^{\mathrm{i}}$ the set of interior faces (or interfaces), i.e., $F \in \mathcal{F}_h^{\mathrm{i}}$ if F is a $(d-1)$-manifold and there are $K_1(F)$, $K_2(F) \in \mathcal{T}_h$ such that $F = K_1(F) \cap K_2(F)$. We denote by $\mathcal{F}_h^{\partial}$ the set of the faces that separate the mesh from the exterior of Ω, i.e., $F \in \mathcal{F}_h^{\partial}$ if F is a $(d-1)$-manifold and there is $K(F) \in \mathcal{T}_h$ such that $F = K(F) \cap \partial\Omega$. Finally, we set $\mathcal{F}_h = \mathcal{F}_h^{\mathrm{i}} \cup \mathcal{F}_h^{\partial}$. Since every function v in $W(h)$ has a (possibly two-valued) trace almost everywhere on $F \in \mathcal{F}_h^{\mathrm{i}}$, it is meaningful to set

$$v^1(x) = \lim_{\substack{y \to x \\ y \in K_1(F)}} v(y), \qquad v^2(x) = \lim_{\substack{y \to x \\ y \in K_2(F)}} v(y), \qquad \text{for a.e. } x \in F, \tag{4.3}$$

$$[\![v]\!] = v^1 - v^2, \qquad \{v\} = \tfrac{1}{2}(v^1 + v^2), \qquad \text{a.e. on } F. \tag{4.4}$$

The arbitrariness in the choice of $K_1(F)$ and $K_2(F)$ could be avoided by choosing an intrinsic notation that would, however, unnecessarily complicate the presentation. For instance, we could have chosen to set $[\![v]\!] = v^1 \otimes n^1 + v^2 \otimes n^2$ where n^1, n^2 are the unit outward normals of $K_1(F)$ and $K_2(F)$, respectively. Although having to choose $K_1(F)$ and $K_2(F)$ may seem cumbersome, nothing that is said hereafter depends on the choice that is made.

For any measurable subset of Ω or $\mathcal{F}_h$, say E, $(\cdot,\cdot)_{L,E}$ denotes the scalar product induced by $[L^2(\Omega)]^m$ or $[L^2(\mathcal{F}_h)]^m$ on E, respectively, and $\|\cdot\|_{L,E}$ the associated norm. Similarly, $\|\cdot\|_{L^d,E}$ denotes the norm induced by $[L^2(\Omega)]^{m\times d}$ or $[L^2(\mathcal{F}_h)]^{m\times d}$ on E. For $K \in \mathcal{T}_h$ (resp., $F \in \mathcal{F}_h$), h_K (resp., h_F) denotes the diameter of K (resp., F).

The mesh family $\{\mathcal{T}_h\}_{h>0}$ is assumed to be shape-regular so that there is a constant c, independent of $h = \max_{K\in\mathcal{T}_h} h_K$, such that for all $v_h \in W_h$ and for all $K \in \mathcal{T}_h$,

$$\|\nabla v_h\|_{L^d,K} \le c\, h_K^{-1} \|v_h\|_{L,K}, \tag{4.5}$$

$$\|v_h\|_{L,F} \le c\, h_K^{-\frac{1}{2}} \|v_h\|_{L,K}, \quad \forall F \subset \partial K. \tag{4.6}$$

4.2. Boundary operators. Henceforth we denote $\mathcal{D}_{\partial\Omega} = \sum_{k=1}^d n_k \mathcal{A}^k$ and we assume that the boundary operator M is associated with a matrix-valued field $\mathcal{M} : \partial\Omega \longrightarrow \mathbb{R}^{m,m}$. Hence, for all functions u, v smooth enough (e.g., $u, v \in [H^1(\Omega)]^m$), the following holds:

$$\langle Du, v\rangle_{W',W} = \int_{\partial\Omega} v^t \mathcal{D}_{\partial\Omega} u, \qquad \langle Mu, v\rangle_{W',W} = \int_{\partial\Omega} v^t \mathcal{M} u. \tag{4.7}$$

To enforce boundary conditions weakly, we introduce for all $F \in \mathcal{F}_h^\partial$ a linear operator $M_F \in \mathcal{L}([L^2(F)]^m; [L^2(F)]^m)$. Assume

$$\forall F \in \mathcal{F}_h^\partial, \qquad M_F \text{ is monotone.} \tag{DG1}$$

Hence, it is meaningful to define for all $v \in W(h)$ the following semi-norms:

$$|v|_M^2 = \sum_{F \in \mathcal{F}_h^\partial} |v|_{M,F}^2 \qquad \text{with} \qquad |v|_{M,F}^2 = (M_F(v), v)_{L,F}. \tag{4.8}$$

In addition to (DG1), the analysis below will show that the design of the boundary operators $\{M_F\}_{F \in \mathcal{F}_h^\partial}$ must comply with the following conditions:

$$\forall v \in [L^2(\partial\Omega)]^m, \quad (\mathcal{M} v = \mathcal{D}_{\partial\Omega} v) \implies (\forall F \in \mathcal{F}_h^\partial,\ M_F(v|_F) = \mathcal{D}_{\partial\Omega} v|_F), \tag{DG2}$$

$$\exists c, \quad \forall v, w \in [L^2(F)]^m, \quad |(M_F(v) - \mathcal{D}_\Omega v, w)_{L,F}| \le c |v|_{M,F} \|w\|_{L,F}, \tag{DG3}$$

$$\exists c, \quad \forall v, w \in [L^2(F)]^m, \quad |(M_F(v) + \mathcal{D}_\Omega v, w)_{L,F}| \le c \|v\|_{L,F} |w|_{M,F}, \tag{DG4}$$

where c is a mesh-independent constant.

Remark 4.1.

(i) Examples of boundary operators M_F are presented in §5 for all the model problems introduced in §3.

(ii) Assumption (DG2) is a consistency assumption while assumptions (DG3) and (DG4) are related to the stability and continuity of the discrete bilinear form; see 4.5.

4.3. Interface operators. For $K \in \mathcal{T}_h$, define the matrix-valued field $\mathcal{D}_{\partial K} : \partial K \to \mathbb{R}^{m,m}$ as

$$\mathcal{D}_{\partial K}(x) = \sum_{k=1}^d n_{K,k} \mathcal{A}^k(x) \quad \text{a.e. on } \partial K, \tag{4.9}$$

where $n_K = (n_{K,1}, \ldots, n_{K,d})^t$ is the unit outward normal to K on ∂K. Note that this definition is compatible with that of $\mathcal{D}_{\partial\Omega}$ in (4.7) if $\partial K \cap \partial\Omega \neq \emptyset$. Moreover, observe that for all u, v in $W(h)$ and for all $K \in \mathcal{T}_h$,

$$(\mathcal{D}_{\partial K} u, v)_{L,\partial K} = (Tu, v)_{L,K} - (u, T^* v)_{L,K}. \tag{4.10}$$

We denote by $\mathcal{D}$ the matrix-valued field defined on $\mathcal{F}_h = \mathcal{F}_h^{\mathrm{i}} \cup \mathcal{F}_h^\partial$ as follows. On $\mathcal{F}_h^\partial$, $\mathcal{D}$ is single-valued and coincides with $\mathcal{D}_{\partial\Omega}$. On $\mathcal{F}_h^{\mathrm{i}}$, $\mathcal{D}$ is two-valued and for all $F \in \mathcal{F}_h^{\mathrm{i}}$, its two values are $\mathcal{D}_{\partial K_1(F)}$ and $\mathcal{D}_{\partial K_2(F)}$. Note that $\{\mathcal{D}\} = 0$ a.e. on $\mathcal{F}_h^{\mathrm{i}}$.

To control the jumps of functions in W_h across mesh interfaces, we introduce for all $F \in \mathcal{F}_h^{\mathrm{i}}$ a linear operator $S_F \in \mathcal{L}([L^2(F)]^m; [L^2(F)]^m)$. Assume

$$\forall F \in \mathcal{F}_h^{\mathrm{i}}, \qquad S_F \text{ is monotone.} \tag{DG5}$$

Hence, it is meaningful to define for all $v \in W(h)$ the following semi-norms:

$$|v|_S^2 = \sum_{F \in \mathcal{F}_h^{\mathrm{i}}} |v|_{S,F}^2 \qquad \text{with} \qquad |v|_{S,F}^2 = (S_F(v), v)_{L,F}. \tag{4.11}$$

In addition to (DG5), the analysis below will show that the design of the interface operators $\{S_F\}_{F\in\mathcal{F}_h^{\mathrm{i}}}$ must comply with the following conditions:

$$S_F \text{ is uniformly bounded, i.e., } \|S_F(v)\|_{L,F} \le c\|v\|_{L,F}, \tag{DG6}$$

$$\exists c, \quad \forall v, w \in [L^2(F)]^m, \quad |(S_F(v), w)_{L,F}| \le c|v|_{S,F}|w|_{S,F}, \tag{DG7}$$

$$\exists c, \quad \forall v, w \in [L^2(F)]^m, \quad |(\mathcal{D}_{\partial K(F)}v, w)_{L,F}| \le c|v|_{S,F}\|w\|_{L,F}, \tag{DG8}$$

where c is a mesh-independent constant and where $K(F)$ denotes any of the two elements sharing F and $\partial K(F)$ its boundary.

Remark 4.2.

(i) Examples of interface operators S_F are presented in §5 for all the model problems introduced in §3.

(ii) Since S_F is positive, a sufficient condition for (DG7) to hold with $c = 1$ is S_F be self-adjoint.

4.4. The discrete problem. We now turn our attention to the construction of a discrete counterpart of (2.29). To this end we introduce the bilinear form a_h such that for all u, v in $W(h)$,

$$\begin{aligned} a_h(v,w) = &\sum_{K\in\mathcal{T}_h} (Tv, w)_{L,K} + \sum_{F\in\mathcal{F}_h^{\partial}} \tfrac{1}{2}(M_F(v) - \mathcal{D}v, w)_{L,F} \\ &- \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} 2(\{\mathcal{D}v\}, \{w\})_{L,F} + \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} (S_F([\![v]\!]), [\![w]\!])_{L,F}. \end{aligned} \tag{4.12}$$

Then, we construct an approximate solution to (2.29) as follows: For $f \in L$,

$$\begin{cases} \text{Seek } u_h \in W_h \text{ such that} \\ a_h(u_h, v_h) = (f, v_h)_L, \quad \forall v_h \in W_h. \end{cases} \tag{4.13}$$

Remark 4.3. In the definition of a_h, the second term weakly enforces the boundary conditions. The purpose of the third term is to ensure that a coercivity property holds, see Lemma 4.1. The last term controls the jump of the discrete solution across interfaces. Some user-dependent arbitrariness appear in the second and fourth term through the definition of the operators M_F and S_F. The design constraints on M_F and S_F are (DG1)–(DG4) and (DG5)–(DG8), respectively.

4.5. Convergence analysis. To perform the error analysis we introduce the following discrete norms on $W(h)$,

$$\|v\|_{h,A}^2 = \|v\|_L^2 + |v|_J^2 + |v|_M^2 + \sum_{K\in\mathcal{T}_h} h_K\|Av\|_{L,K}^2, \tag{4.14}$$

$$\|v\|_{h,\frac{1}{2}}^2 = \|v\|_{h,A}^2 + \sum_{K\in\mathcal{T}_h} [h_K^{-1}\|v\|_{L,K}^2 + \|v\|_{L,\partial K}^2], \tag{4.15}$$

where we have introduced the jump semi-norms

$$|v|_J^2 = \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} |v|_{J,F}^2 \qquad \text{with} \qquad |v|_{J,F}^2 = |[\![v]\!]|_{S,F}. \tag{4.16}$$

The norm $\|\cdot\|_{h,A}$ is used to measure the approximation error, and the norm $\|\cdot\|_{h,\frac{1}{2}}$ serves to measure the interpolation properties of the discrete space W_h.

Throughout this section, we assume that:

- Problem (2.29) is well-posed.
- The mesh family $\{\mathcal{T}_h\}_{h>0}$ is shape-regular so that (4.5) and (4.6) hold.
- The design assumptions (DG1)–(DG8) on M_F and S_F hold.

LEMMA 4.1 (L-coercivity). *For all h and for all v in $W(h)$,*

$$a_h(v,v) \geq \mu_0 \|v\|_L^2 + |v|_J^2 + \tfrac{1}{2}|v|_M^2. \tag{4.17}$$

Proof. Let v in $W(h)$. Using (4.10) and summing over the mesh elements yields

$$\sum_{F\in\mathcal{F}_h^\partial} \tfrac{1}{2}(\mathcal{D}v,v)_{L,F} + \sum_{F\in\mathcal{F}_h^{\rm i}} \int_F \{v^t\mathcal{D}v\} = \tfrac{1}{2}\sum_{K\in\mathcal{T}_h}[(Tv,v)_{L,K} - (v,T^*v)_{L,K}].$$

Subtracting this equation to (4.12) and using the fact that $\{v^t\mathcal{D}v\} = 2\{v^t\}\{\mathcal{D}v\}$ leads to

$$a_h(v,v) = \tfrac{1}{2}\sum_{K\in\mathcal{T}_h}[(Tv,v)_{L,K} + (v,T^*v)_{L,K}] + |v|_J^2 + \tfrac{1}{2}|v|_M^2.$$

Then, the desired result follows using (A4). ∎

LEMMA 4.2. *There is $c>0$, independent of h, such that for all F in $\mathcal{F}_h^{\rm i}$ and for all $v,w\in W(h)$,*

$$|(S_F([\![v]\!]),[\![w]\!])_{L,F}| + |(\{\mathcal{D}v\},\{w\})_{L,F}| \leq c|v|_{J,F}(\|\{w\}\|_{L,F} + \|[\![w]\!]\|_{L,F}). \tag{4.18}$$

Proof. (1) Owing to (DG7), $(S_F([\![v]\!]),[\![w]\!])_{L,F} \leq c|v|_{J,F}|w|_{J,F}$, and owing to (DG6), $|w|_{J,F}\leq c\|[\![w]\!]\|_{L,F}$. Hence, $(S_F([\![v]\!]),[\![w]\!])_{L,F}\leq c|v|_{J,F}\|[\![w]\!]\|_{L,F}$.
(2) Let $K_1(F)$ and $K_2(F)$ be the two mesh elements such that $F=K_1(F)\cap K_2(F)$. Then, $2\{\mathcal{D}v\} = \mathcal{D}_{K_1(F)}[\![v]\!]$ since $\{D\}=0$. Using (DG8) yields

$$|(\{\mathcal{D}v\},\{w\})_{L,F}| = |(\mathcal{D}_{K_1(F)}[\![v]\!],\{w\})_{L,F}| \leq c|v|_{J,F}\|\{w\}\|_{L,F}.$$

The proof is complete. ∎

LEMMA 4.3 (Stability). *Assume that for all $k\in\{1,\ldots,d\}$, $\mathcal{A}^k\in[\mathfrak{C}^{0,\frac{1}{2}}(\overline{\Omega})]^{m,m}$. Then, there is $c>0$, independent of h, such that*

$$\inf_{v_h\in W_h\setminus\{0\}}\sup_{w_h\in W_h\setminus\{0\}}\frac{a_h(v_h,w_h)}{\|v_h\|_{h,A}\|w_h\|_{h,A}} \geq c. \tag{4.19}$$

Proof. (1) Let v_h be an arbitrary element in W_h. Let $K\in\mathcal{T}_h$. Denote by $\overline{\mathcal{A}_K^k}$ the mean-value of $\mathcal{A}^k$ on K; then,

$$\|\mathcal{A}^k - \overline{\mathcal{A}_K^k}\|_{[L^\infty(K)]^{m,m}} \leq \|\mathcal{A}^k\|_{[\mathfrak{C}^{0,\frac{1}{2}}(\overline{\Omega})]^{m,m}} h_K^{\frac{1}{2}}. \tag{4.20}$$

Set $\overline{A}_K v_h = \sum_{k=1}^d \overline{\mathcal{A}_K^k}\partial_k v_h$ and $\pi_h = \sum_{K\in\mathcal{T}_h} h_K\overline{A}_K v_h$. Clearly, $\pi_h\in W_h$. Using (4.20), together with the inverse inequalities (4.5) and (4.6), leads to

$$\begin{cases} \|\overline{A}_K v_h\|_{L,F} \leq c\, h_K^{-\frac{1}{2}}\|\overline{A}_K v_h\|_{L,K}, & \text{if } F\in\mathcal{F}_h^\partial,\\ \|\{\overline{A}_K v_h\}\|_{L,F} + \|[\![\overline{A}_K v_h]\!]\|_{L,F} \leq c\, h_K^{-\frac{1}{2}}\|\overline{A}_K v_h\|_{L,K_1\cup K_2}, & \text{if } F\in\mathcal{F}_h^{\rm i}, \end{cases} \tag{4.21}$$

$$\|\overline{A}_K v_h\|_{L,K} \leq c\,\min(\|Av_h\|_{L,K} + h_K^{-\frac{1}{2}}\|v_h\|_{L,K}, h_K^{-1}\|v_h\|_{L,K}). \tag{4.22}$$

Note that (4.22) implies $\|\pi_h\|_L \leq c\|v_h\|_L$. From the definition of a_h it follows that

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} h_K\|Av_h\|^2_{L,K} &= a_h(v_h,\pi_h) - (Kv_h,\pi_h)_L - \sum_{F\in\mathcal{F}_h^\partial} \tfrac{1}{2}(M_F(v_h)-\mathcal{D}v_h,\pi_h)_{L,F}\\ &\quad + \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} [2(\{\mathcal{D}v_h\},\{\pi_h\})_{L,F} - (S_F([\![v_h]\!]),[\![\pi_h]\!])_{L,F}]\\ &\quad + \sum_{K\in\mathcal{T}_h} h_K(Av_h,(A-\overline{A}_K)v_h)_{L,K}\\ &= a_h(v_h,\pi_h) + R_1 + R_2 + R_3 + R_4,\end{aligned}$$

where R_1, R_2, R_3, and R_4 denote the second, third, fourth, and fifth term in the right-hand side of the above equation, respectively. Each of these terms is bounded from above as follows. Using (4.22) yields $\|\pi_h\|_L \leq c\|v_h\|_L$ and hence,

$$|R_1| \leq c\|v_h\|_L\|\pi_h\|_L \leq c\|v_h\|^2_L.$$

Using (DG3) together with (4.21) and (4.22) leads to

$$\begin{aligned}|R_2| &\leq \sum_{F\in\mathcal{F}_h^\partial} [c_\gamma(M_F(v_h),v_h)_{L,F} + \gamma\|\pi_h\|^2_{L,F}]\\ &\leq c(\|v_h\|^2_L + |v_h|^2_M) + \gamma\sum_{K\in\mathcal{T}_h} h_K\|Av_h\|^2_{L,K},\end{aligned}$$

where $\gamma > 0$ can be chosen as small as needed. For the third term, use Lemma 4.2, together with inequalities (4.21) and (4.22), as follows

$$\begin{aligned}|R_3| &\leq \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} c_\gamma|v_h|^2_{J,F} + \gamma\sum_{K\in\mathcal{T}_h} h_K\|\overline{A}_K v_h\|^2_{L,K}\\ &\leq c(\|v_h\|^2_L + |v_h|^2_J) + \gamma\sum_{K\in\mathcal{T}_h} h_K\|Av_h\|^2_{L,K}.\end{aligned}$$

For the last term, (4.5) and (4.20) yield

$$\begin{aligned}|R_4| &\leq \sum_{K\in\mathcal{T}_h} h_K\|Av_h\|_{L,K} c h_K^{\frac{1}{2}}\|\nabla v_h\|_{L^d,K}\\ &\leq c\sum_{K\in\mathcal{T}_h} h_K^{\frac{1}{2}}\|Av_h\|_{L,K}\|v_h\|_{L,K} \leq c\|v_h\|^2_L + \gamma\sum_{K\in\mathcal{T}_h} h_K\|Av_h\|^2_{L,K}.\end{aligned}$$

Using the above four bounds, $\gamma = \frac{1}{6}$, and Lemma 4.1 leads to

$$\tfrac{1}{2}\sum_{K\in\mathcal{T}_h} h_K\|Av_h\|^2_{L,K} \leq a_h(v_h,\pi_h) + c\,a_h(v_h,v_h). \tag{4.23}$$

(2) Let us now prove that $\|\pi_h\|_{h,A} \leq c\,\|v_h\|_{h,A}$. We have already seen that $\|\pi_h\|_L \leq c\|v_h\|_L$. Using (4.5), together with inequalities (4.20) and (4.22), leads to

$$\sum_{K\in\mathcal{T}_h} h_K\|A\pi_h\|^2_{L,K} \leq c\sum_{K\in\mathcal{T}_h} h_K^{-1}\|\pi_h\|^2_{L,K} \leq c\sum_{K\in\mathcal{T}_h}[h_K\|Av_h\|^2_{L,K} + \|v_h\|^2_{L,K}].$$

Moreover, the inverse inequality (4.6), assumption (DG6), and inequalities (4.21) and (4.22) yield

$$|\pi_h|_J^2 = \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} |\pi_h|_{J,F}^2 \le c \sum_{K\in\mathcal{T}_h} h_K^{-1}\|\pi_h\|_{L,K}^2 \le c \sum_{K\in\mathcal{T}_h} [h_K\|Av_h\|_{L,K}^2 + \|v_h\|_{L,K}^2].$$

Proceed similarly to control $|\pi_h|_M$. In conclusion,

$$\|\pi_h\|_{h,A} \le c\,\|v_h\|_{h,A}. \tag{4.24}$$

(3) Owing to (4.17) and (4.23), there is $c_1 > 0$ such that

$$\|v_h\|_{h,A}^2 \le c_1 a_h(v_h, v_h) + a_h(v_h, \pi_h) = a_h(v_h, \pi_h + c_1 v_h).$$

Then, setting $w_h = \pi_h + c_1 v_h$ and using (4.24) yields

$$\|v_h\|_{h,A}\|w_h\|_{h,A} \le c\,\|v_h\|_{h,A}^2 \le c\,a_h(v_h, w_h).$$

The conclusion is straightforward. □

LEMMA 4.4 (Continuity). *Under the hypotheses of Lemma* 4.3*, there is c, independent of h, such that*

$$\forall (v,w)\in W(h)\times W(h), \quad a_h(v,w) \le c\,\|v\|_{h,\frac12}\|w\|_{h,A}. \tag{4.25}$$

Proof. The general principle of the proof is to integrate by parts $a_h(v,w)$ by making use of the formal adjoint T^*. Observing that

$$\sum_{K\in\mathcal{T}_h} [(Tv,w)_{L,K} - (v,T^*w)_{L,K}] = \sum_{F\in\mathcal{F}_h^{\partial}} (\mathcal{D}v,w)_{L,F} + \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} \int_F 2\left\{w^t\mathcal{D}v\right\},$$

and $2\left\{w^t\mathcal{D}v\right\} = 2\left\{w^t\right\}\left\{\mathcal{D}v\right\} + \frac12[\![w^t]\!][\![\mathcal{D}v]\!]$, it is clear that

$$\begin{aligned} a_h(v,w) = &\sum_{K\in\mathcal{T}_h} (v,T^*w)_{L,K} + \sum_{F\in\mathcal{F}_h^{\partial}} \tfrac12(M_F(v)+\mathcal{D}v,w)_{L,F} \\ &+ \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} \tfrac12([\![\mathcal{D}v]\!],[\![w]\!])_{L,F} + \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} (S_F([\![v]\!]),[\![w]\!])_{L,F}. \end{aligned} \tag{4.26}$$

Let R_1 to R_4 be the four terms in the right-hand side. Using the Cauchy–Schwarz inequality yields

$$|R_1| \le c \sum_{K\in\mathcal{T}_h} \|v\|_{L,K}(\|w\|_{L,K} + \|Aw\|_{L,K}) \le c\|v\|_{h,\frac12}\|w\|_{h,A}.$$

Use (DG4) together with the Cauchy–Schwarz inequality to infer

$$|R_2| \le c \sum_{F\in\mathcal{F}_h^{\partial}} \|v\|_{L,F}|w|_{M,F} \le c\|v\|_{h,\frac12}\|w\|_{h,A}.$$

For the third and fourth term, use (DG6) and (DG7), together with the fact that $[\![\mathcal{D}v]\!] = 2\mathcal{D}_{\partial K_1(F)}\{v\}$, to obtain

$$|R_3| + |R_4| \le c \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} (\|\{v\}\|_{L,F} + \|[\![v]\!]\|_{L,F})|w|_{J,F} \le c\|v\|_{h,\frac12}\|w\|_{h,A}.$$

The result follows easily. □

LEMMA 4.5 (Consistency). *Let u solve* (2.29) *and let u_h solve* (4.13). *If $u \in [H^1(\Omega)]^m$, then,*

$$\forall v_h \in W_h, \quad a_h(u - u_h, v_h) = 0. \tag{4.27}$$

Proof. Since $u \in [H^1(\Omega)]^m$ solves (2.29), $\mathcal{M}u = \mathcal{D}u$ a.e. on $\partial\Omega$ and $Tu = f$ in L. Assumption (DG2) yields $M_F(u|_F) = \mathcal{D}u|_F$ for all $F \in \mathcal{F}_h^\partial$. Moreover, $u \in [H^1(\Omega)]^m$ implies that $\{\mathcal{D}u\} = 0$ and $[\![u]\!] = 0$ a.e. on $\mathcal{F}_h^{\mathrm{i}}$. As a result,

$$\forall v_h \in W_h, \quad a_h(u, v_h) = (Tu, v_h)_L = (f, v_h)_L = a_h(u_h, v_h).$$

The conclusion follows readily. □

THEOREM 4.6 (Convergence). *Under the hypotheses of Lemmas* 4.3 *and* 4.5, *there is c, independent of h, such that*

$$\|u - u_h\|_{h,A} \le c \inf_{v_h \in W_h} \|u - v_h\|_{h,\frac{1}{2}}. \tag{4.28}$$

Proof. Simple application of Strang's Second Lemma; see, e.g., [15, p. 94]. Let $v_h \in W_h$. Owing to Lemmas 4.3, 4.4, and 4.5,

$$\begin{aligned}\|v_h - u_h\|_{h,A} &\le c \sup_{w_h \in W_h \setminus \{0\}} \frac{a_h(v_h - u_h, w_h)}{\|w_h\|_{h,A}} \\ &\le c \sup_{w_h \in W_h \setminus \{0\}} \frac{a_h(v_h - u, w_h)}{\|w_h\|_{h,A}} \le c\,\|u - v_h\|_{h,\frac{1}{2}}.\end{aligned}$$

Conclude using the triangle inequality. □

Owing to the definition of W_h, and the regularity of the mesh family $\{\mathcal{T}_h\}_{h>0}$, the following interpolation property holds: There is c, independent of h, such that for all $v \in [H^{p+1}(\Omega)]^m$, there is $v_h \in W_h$ satisfying

$$\|v - v_h\|_{h,\frac{1}{2}} \le c h^{p+\frac{1}{2}} \|v\|_{[H^{p+1}(\Omega)]^m}. \tag{4.29}$$

COROLLARY 4.7. *If u is in $[H^{p+1}(\Omega)]^m$, there is c, independent of h, such that*

$$\|u - u_h\|_{h,A} \le c\, h^{p+\frac{1}{2}} \|u\|_{[H^{p+1}(\Omega)]^m}. \tag{4.30}$$

In particular,

$$\|u - u_h\|_L \le c\, h^{p+\frac{1}{2}} \|u\|_{[H^{p+1}(\Omega)]^m}, \tag{4.31}$$

and if the mesh family $\{\mathcal{T}_h\}_{h>0}$ is quasi-uniform,

$$\|A(u - u_h)\|_L \le c\, h^p \|u\|_{[H^{p+1}(\Omega)]^m}. \tag{4.32}$$

The above estimates show that, provided the exact solution is smooth enough, the method yields p-order convergence in the graph norm and $(p + \frac{1}{2})$-order convergence in the L-norm.

Remark 4.4.

(i) To apply Strang's Second Lemma, it is actually sufficient that the continuity property established in Lemma 4.4 holds for $(v, w_h) \in W(h) \times W_h$.

(ii) The estimates (4.30) to (4.32) are identical to those that can be obtained by other stabilization methods like GaLS [5, 18, 19] or subgrid viscosity [17] and many other methods.

4.6. Localization, fluxes, and adjoint-fluxes. The purpose of this section is to discuss briefly some equivalent formulations of the discrete problem (4.13) in order to emphasize the link with other formalisms derived previously for DG methods, namely that of Lesaint and Raviart [22, 21] and Johnson et al. [19, 20] for Friedrichs' symmetric systems and that of Arnold et al. [2] for the Laplacian. To this end, we rewrite the bilinear form (4.12) in various equivalent ways and introduce the concept of element fluxes and that of element adjoint-fluxes.

Let $K \in \mathcal{T}_h$. Define the operator $M^{\mathrm{L}}_{\partial K} \in \mathcal{L}([L^2(\partial K)]^m; [L^2(\partial K)]^m)$ as follows: For $v \in [L^2(\partial K)]^m$ and a face $F \subset \partial K$, set

$$M^{\mathrm{L}}_{\partial K}(v)|_F = \begin{cases} M_F(v|_F), & \text{if } F \in \mathcal{F}_h^{\partial}, \\ 2S_F(v|_F), & \text{if } F \in \mathcal{F}_h^{\mathrm{i}}. \end{cases} \tag{4.33}$$

Furthermore, for $v \in W(h)$ and $x \in \partial K$, set

$$v^i(x) = \lim_{\substack{y \to x \\ y \in K}} v(y), \qquad v^e(x) = \lim_{\substack{y \to x \\ y \notin K}} v(y), \tag{4.34}$$

$$[\![v]\!]_{\partial K}(x) = v^i(x) - v^e(x), \qquad \{v\}_{\partial K}(x) = \tfrac{1}{2}(v^i(x) + v^e(x)), \tag{4.35}$$

with $v^e(x) = 0$ if $x \in \partial\Omega$. Then, a straightforward calculation shows that the bilinear form a_h defined in (4.12) can be rewritten as follows:

$$a_h(u,v) = \sum_{K \in \mathcal{T}_h} (Tu, v)_{L,K} + \sum_{K \in \mathcal{T}_h} \tfrac{1}{2}(M^{\mathrm{L}}_{\partial K}([\![u]\!]_{\partial K}) - \mathcal{D}_{\partial K}[\![u]\!]_{\partial K}, v^i)_{L,\partial K}. \tag{4.36}$$

This is the bilinear form analyzed by Lesaint and Raviart [21, 22] and further investigated by Johnson et al. [19] in the particular case where the operator $M^{\mathrm{L}}_{\partial K}$ is defined pointwise using a matrix-valued field on ∂K; see §5.1 for further discussion.

Using (4.36) in the discrete problem (4.13) and localizing the test functions to the mesh elements yields the following local formulation

$$\begin{cases} \text{Seek } u_h \in W_h \text{ such that } \forall K \in \mathcal{T}_h \text{ and } \forall v_h \in \mathbb{P}_p(K), \\ (Tu_h, v_h)_{L,K} + \tfrac{1}{2}(M^{\mathrm{L}}_{\partial K}([\![u_h]\!]_{\partial K}) - \mathcal{D}_{\partial K}[\![u_h]\!]_{\partial K}, v_h)_{L,\partial K} = (f, v_h)_{L,K}. \end{cases} \tag{4.37}$$

Likewise, using (4.10) reveals that the bilinear form a_h can also be recast into the following form

$$a_h(u,v) = \sum_{K \in \mathcal{T}_h} (u, T^*v)_{L,K} + \sum_{K \in \mathcal{T}_h} (\tfrac{1}{2}M^{\mathrm{L}}_{\partial K}([\![u]\!]_{\partial K}) + \mathcal{D}_{\partial K}\{u\}_{\partial K}, v^i)_{L,\partial K}. \tag{4.38}$$

By localizing the test functions to the mesh elements we obtain the following equivalent local formulation of (4.13)

$$\begin{cases} \text{Seek } u_h \in W_h \text{ such that } \forall K \in \mathcal{T}_h \text{ and } \forall v_h \in \mathbb{P}_p(K), \\ (u_h, T^*v_h)_{L,K} + (\tfrac{1}{2}M^{\mathrm{L}}_{\partial K}([\![u_h]\!]_{\partial K}) + \mathcal{D}_{\partial K}\{u_h\}_{\partial K}, v_h)_{L,\partial K} = (f, v_h)_{L,K}. \end{cases} \tag{4.39}$$

In view of (4.37) and (4.39), we are led to introduce a concept of flux and adjoint-flux.

DEFINITION 4.8. *Let $K \in \mathcal{T}_h$ and let $v \in W(h)$. The* element flux *of v on ∂K, say $\phi_{\partial K}(v) \in [L^2(\partial K)]^m$, is defined on a face $F \subset \partial K$ by*

$$\begin{aligned} \phi_{\partial K}(v)|_F &= \tfrac{1}{2}M^{\mathrm{L}}_{\partial K}([\![v]\!]_{\partial K}) + \mathcal{D}_{\partial K}\{v\}_{\partial K} \\ &= \begin{cases} \tfrac{1}{2}M_F(v|_F) + \tfrac{1}{2}\mathcal{D}_{\partial\Omega}v, & \text{if } F \subset \partial K^{\partial}, \\ S_F([\![v]\!]_{\partial K}|_F) + \mathcal{D}_{\partial K}\{v\}_{\partial K}, & \text{if } F \subset \partial K^{\mathrm{i}}, \end{cases} \end{aligned} \tag{4.40}$$

where ∂K^{i} denotes that part of ∂K that lies in Ω and ∂K^{∂} that part of ∂K that lies on $\partial\Omega$. Likewise, the element adjoint-flux *of v on ∂K, say $\phi^*_{\partial K}(v) \in [L^2(\partial K)]^m$, is defined on a face $F \subset \partial K$ by*

$$\begin{aligned}\phi^*_{\partial K}(v)|_F &= \tfrac{1}{2} M^{\mathrm{L}}_{\partial K}(\llbracket v \rrbracket_{\partial K}) - \tfrac{1}{2}\mathcal{D}_{\partial K}\llbracket v \rrbracket_{\partial K} \\ &= \begin{cases} \tfrac{1}{2} M_F(v|_F) - \tfrac{1}{2}\mathcal{D}_{\partial\Omega} v, & \text{if } F \subset \partial K^{\partial}, \\ S_F(\llbracket v \rrbracket_{\partial K}|_F) - \tfrac{1}{2}\mathcal{D}_{\partial K}\llbracket v \rrbracket_{\partial K}, & \text{if } F \subset \partial K^{\mathrm{i}}. \end{cases}\end{aligned} \tag{4.41}$$

The relevance of the notion of flux and adjoint-flux is clarified by the following

PROPOSITION 4.9. *The discrete problem* (4.37) *is equivalent to each of the following two formulations:*

$$\begin{cases} \text{Seek } u_h \in W_h \text{ such that } \forall K \in \mathcal{T}_h \text{ and } \forall v_h \in [\mathbb{P}_p(K)]^m, \\ (u_h, T^* v_h)_{L,K} + (\phi_{\partial K}(u_h), v_h)_{L,\partial K} = (f, v_h)_{L,K}. \end{cases} \tag{4.42}$$

$$\begin{cases} \text{Seek } u_h \in W_h \text{ such that } \forall K \in \mathcal{T}_h \text{ and } \forall v_h \in [\mathbb{P}_p(K)]^m, \\ (T u_h, v_h)_{L,K} + (\phi^*_{\partial K}(u_h), v_h)_{L,\partial K} = (f, v_h)_{L,K}. \end{cases} \tag{4.43}$$

Proof. Straightforward consequence of (4.37) and (4.39) together with Definition 4.8. □

Let v be a function in $W(h)$. We define the *interface fluxes* (resp., *interface adjoint-fluxes*) of v, say $\phi^{\mathrm{i}}(v)$ (resp., say $\phi^{*,\mathrm{i}}(v)$), to be the two-valued function defined on $\mathcal{F}_h^{\mathrm{i}}$ that collects all the element fluxes (resp. adjoint-fluxes) of v on the interior faces. Likewise we define the *boundary fluxes* (resp., *boundary adjoint-fluxes*) of v, say $\phi^{\partial}(v)$ (resp., say $\phi^{*,\partial}(v)$), to be the single-valued function defined on $\mathcal{F}_h^{\partial}$ that collects all the element fluxes (resp., adjoint-fluxes) of v on the boundary faces.

Remark 4.5.

(i) The link between DG methods and the concept of element fluxes has been explored recently in [2] for the Laplacian (in [2], the terminology "numerical fluxes" is employed instead).

(ii) In engineering practice, approximation schemes such as (4.42) are often designed by a priori specifying the element fluxes. The above analysis then provides a practical means to assess the stability and convergence properties of the scheme. Indeed, once the element fluxes are given, the boundary operators M_F and the interface operators S_F can be directly retrieved from (4.40). Then, properties (DG1)–(DG8) provide sufficient conditions to analyze the scheme.

(iii) The interface fluxes are such that $\{\phi^{\mathrm{i}}(v)\} = 0$ a.e. on $\mathcal{F}_h^{\mathrm{i}}$. Approximation schemes where the interface fluxes satisfy this property are often termed *conservative*. Note that the concept of conservativity as such does not play any role in the present analysis of the method, although it can play a role when deriving improved L^2-error estimates by using the Aubin–Nitsche lemma; see, e.g., Arnold et al. [2] and the second part of this work [14].

(iv) The following relation links the element fluxes and the element adjoint-fluxes

$$\phi_{\partial K}(v) - \phi^*_{\partial K}(v) = \mathcal{D}_{\partial K} v^i. \tag{4.44}$$

In particular, the element adjoint-fluxes are not conservative.

(v) Both the element fluxes and the element adjoint-fluxes are associated with

the operator T, i.e., they are derived from a DG discretization of (2.29). It is also possible to design a DG discretization of the adjoint problem (2.30) involving the operator T^* and the bilinear form a^*. This would lead to two new families of fluxes, the element fluxes for T^* and the element adjoint-fluxes for T^*. It should be noted that the element adjoint-fluxes for T are not the element fluxes for T^*. In particular, the former are not conservative whereas the latter are conservative.

5. Applications. This section shows how the conditions (DG1)–(DG8) can be used to design DG approximations of the model problems introduced in §3.

5.1. Pointwise boundary and interface operators. For ease of presentation, the boundary and interface operators discussed in this section are constructed from matrix-valued fields defined on all the mesh faces. This simpler construction is sufficient to recover several DG methods considered in the literature. Examples where the more general form for the boundary and interface operators is needed will be presented in a forthcoming work [14].

Let $\widehat{\mathcal{M}} \in L^\infty(\partial\Omega;\mathbb{R}^{m,m})$ be a matrix-valued field such that, a.e. in $\partial\Omega$,

$$\widehat{\mathcal{M}} \text{ is positive,} \tag{DG1a}$$

$$\mathrm{Ker}(\mathcal{M}-\mathcal{D}_{\partial\Omega}) \subset \mathrm{Ker}(\widehat{\mathcal{M}}-\mathcal{D}_{\partial\Omega}), \tag{DG2a}$$

$$\exists c,\quad \forall \xi,\zeta\in\mathbb{R}^m,\quad |\zeta^t(\widehat{\mathcal{M}}-\mathcal{D}_{\partial\Omega})\xi| \le c(\xi^t\widehat{\mathcal{M}}\xi)^{\frac12}\|\zeta\|_{\mathbb{R}^m}, \tag{DG3a}$$

$$\exists c,\quad \forall \xi,\zeta\in\mathbb{R}^m,\quad |\zeta^t(\widehat{\mathcal{M}}+\mathcal{D}_{\partial\Omega})\xi| \le c(\zeta^t\widehat{\mathcal{M}}\zeta)^{\frac12}\|\xi\|_{\mathbb{R}^m}, \tag{DG4a}$$

where $\|\cdot\|_{\mathbb{R}^m}$ denotes the Euclidean norm in $\mathbb{R}^m$. A straightforward verification yields the following

PROPOSITION 5.1. *For all $F\in\mathcal{F}_h^\partial$, define*

$$M_F : [L^2(F)]^m \ni v \mapsto \widehat{\mathcal{M}}|_F v \in [L^2(F)]^m. \tag{5.1}$$

Then, the operator M_F satisfies (DG1)–(DG4).

Let $\Omega_{\mathcal{F}}$ be the set of points located on an interior face of the mesh. Let $\mathcal{S}\in L^\infty(\Omega_{\mathcal{F}};\mathbb{R}^{m,m})$ be a matrix-valued field such that, a.e. in $\Omega_{\mathcal{F}}$,

$$\mathcal{S} \text{ is positive,} \tag{DG5a}$$

$$\mathcal{S} \text{ is uniformly bounded,} \tag{DG6a}$$

$$\exists c,\quad \forall \xi,\zeta\in\mathbb{R}^m,\quad |\zeta^t\mathcal{S}\xi| \le c(\xi^t\mathcal{S}\xi)^{\frac12}(\zeta^t\mathcal{S}\zeta)^{\frac12}, \tag{DG7a}$$

$$\exists c,\quad \forall \xi,\zeta\in\mathbb{R}^m,\quad |\zeta^t\mathcal{D}\xi| \le c(\xi^t\mathcal{S}\xi)^{\frac12}\|\zeta\|_{\mathbb{R}^m}. \tag{DG8a}$$

A straightforward verification yields the following

PROPOSITION 5.2. *For all $F\in\mathcal{F}_h^{\mathrm{i}}$, define*

$$S_F : [L^2(F)]^m \ni v \mapsto \mathcal{S}|_F v \in [L^2(F)]^m. \tag{5.2}$$

Then, the operator S_F satisfies (DG5)–(DG8).

Remark 5.1.

(i) Whenever the matrix-valued field $\mathcal{M}$ defined in (4.7) satisfies (DG3a)–(DG4a), one simply sets $\widehat{\mathcal{M}}=\mathcal{M}$; otherwise, it is necessary to strengthen $\mathcal{M}$. This last situation occurs, for instance, when approximating advection–diffusion–reaction problems and the Maxwell equations in the diffusive regime; see §5.3 and §5.4.

(ii) One possible way of constructing $\mathcal{S}$ is as follows. Since $\mathcal{D}$ is symmetric, $\mathcal{D}$ is diagonalizable; hence, the absolute value of $\mathcal{D}$, say $|\mathcal{D}|$, can be defined. Moreover, observing that $|\mathcal{D}|$ is single-valued on $\mathcal{F}_h^{\mathrm{i}}$, one can set $\mathcal{S}=|\mathcal{D}|$.

5.2. Advection–reaction. Consider the advection–reaction problem introduced in §3.1. Assume that all the faces in $\mathcal{F}_h^\partial$ are in $\partial\Omega^-$, in $\partial\Omega^+$, or in $\partial\Omega\backslash(\partial\Omega^-\cup\partial\Omega^+)$. Let $K\in\mathcal{T}_h$. The scalar-valued field $\mathcal{D}_{\partial K}$ is $\mathcal{D}_{\partial K}=\beta{\cdot}n_K$, and the scalar-valued field $\mathcal{M}$ associated with the boundary operator M is $\mathcal{M}=|\beta{\cdot}n|$. It is straightforward to verify the following results.

LEMMA 5.3. *Properties* (DG1a)–(DG4a) *hold for* $\widehat{\mathcal{M}}=\mathcal{M}=|\beta{\cdot}n|$.

LEMMA 5.4. *Let* $\alpha>0$. *For all* $F\in\mathcal{F}_h^{\mathrm{i}}$ *and for a.e.* $x\in F$, *define* $\mathcal{S}=\alpha|\beta{\cdot}n|$ *where* n *is a unit normal vector to* F *(its orientation is clearly irrelevant to define* $\mathcal{S}$*). Then, properties* (DG5a)–(DG8a) *hold.*

Owing to Definition 4.8 and Proposition 4.9, the local formulation (4.42) takes the following form: Seek $u_h\in W_h$ such that $\forall K\in\mathcal{T}_h$ and $\forall v_h\in\mathbb{P}_p(K)$,

$$((\mu-\nabla{\cdot}\beta)u_h,v_h)_{L,K}-(u_h,\beta{\cdot}\nabla v_h)_{L,K}+(\phi_{\partial K}(u_h),v_h)_{L,\partial K}=(f,v_h)_{L,K},\tag{5.3}$$

with the interface and boundary fluxes

$$\phi^{\mathrm{i}}(u_h)|_{\partial K}=(\beta{\cdot}n_K)\,\{u_h\}+\alpha|\beta{\cdot}n_K|[\![u_h]\!]_{\partial K},\tag{5.4}$$

$$\phi^{\partial}(u_h)=|\beta{\cdot}n|u_h1_{\partial\Omega^+},\tag{5.5}$$

where $1_{\partial\Omega^+}$ denotes the characteristic function of $\partial\Omega^+$.

Likewise, the local formulation (4.43) takes the following form: Seek $u_h\in W_h$ such that $\forall K\in\mathcal{T}_h$ and $\forall v_h\in\mathbb{P}_p(K)$,

$$(\mu u_h+\beta{\cdot}\nabla u_h,v_h)_{L,K}+(\phi^*_{\partial K}(u_h),v_h)_{L,\partial K^{\mathrm{i}}}=(f,v_h)_{L,K},\tag{5.6}$$

with the interface and boundary adjoint-fluxes

$$\phi^{*,\mathrm{i}}(u_h)|_{\partial K}=(\alpha|\beta{\cdot}n_K|-\tfrac{1}{2}\beta{\cdot}n_K)[\![u_h]\!]_{\partial K},\tag{5.7}$$

$$\phi^{*,\partial}(u_h)=-|\beta{\cdot}n|u_h1_{\partial\Omega^-},\tag{5.8}$$

where $1_{\partial\Omega^-}$ denotes the characteristic function of $\partial\Omega^-$.

Remark 5.2.

(i) The design parameter α can vary from face to face.

(ii) The specific value $\alpha=\frac{1}{2}$ has received considerable attention in the literature. When working with formulation (5.6) with this value of α, one obtains the DG method analyzed by Lesaint and Raviart [21, 22]; in this case the interface adjoint-flux $\phi^{*,\mathrm{i}}$ is nonzero only on that part of the boundary ∂K where $\beta{\cdot}n_K<0$. When working with the formulation (5.3), the particular choice $\alpha=\frac{1}{2}$ leads to

$$\phi^{\mathrm{i}}(u_h)=(\beta{\cdot}n_K)u_h^{\uparrow}\quad\text{where}\quad u_h^{\uparrow}=\begin{cases}u_h^i, & \text{if } \beta{\cdot}n_K>0,\\ u_h^e, & \text{otherwise,}\end{cases}\tag{5.9}$$

i.e., the well-known upwind flux is recovered as a particular case of the above analysis which is valid for any $\alpha>0$. This coincidence has lead many authors to believe that DG methods are methods of choice to solve hyperbolic problems. Actually DG methods, as presented herein, are tailored to solve symmetric systems of first-order PDEs, and as pointed out by Friedrichs, the notion of symmetric systems goes beyond the hyperbolic/elliptic traditional classification of PDEs. Moreover, the presence of the user-dependent interface operator S_F (see (DG5)–(DG8)) points out to the fact that DG methods are merely stabilization techniques. This fact is even clearer when one realizes that the error estimates (4.30)–(4.32) are identical to those that can be obtained by using other stabilization techniques like GaLS (also sometimes called streamline diffusion) [5, 18, 19] or subgrid viscosity [17] methods.

5.3. Advection–diffusion–reaction. Consider the advection–diffusion–reaction problem introduced in §3.2. Let $K \in \mathcal{T}_h$. Then, the $\mathbb{R}^{d+1,d+1}$-valued field $\mathcal{D}_{\partial K}$ is

$$\mathcal{D}_{\partial K} = \left[\begin{array}{c:c} 0 & n_K \\ \hdashline n_K^t & \beta{\cdot}n_K \end{array}\right]. \tag{5.10}$$

To simplify, we assume that the parameters β and μ are of order one, i.e., we hide the dependency on these coefficients in the constants. Special cases such as advection-dominated problems go beyond the scope of the present work.

We begin with the interface operator S_F since its design is independent of the boundary conditions imposed. For a vector $\xi \in \mathbb{R}^{d+1}$, denote by $\xi = (\xi_\sigma, \xi_u)$ its canonical decomposition in $\mathbb{R}^d \times \mathbb{R}$ and use a similar notation for $\zeta = (\zeta_\sigma, \zeta_u) \in \mathbb{R}^{d+1}$.

LEMMA 5.5. *Let $\alpha > 0$, $\eta > 0$, and $\delta \in \mathbb{R}^d$. For all $F \in \mathcal{F}_h^{\mathrm{i}}$ and for a.e. $x \in F$, define*

$$\mathcal{S} = \left[\begin{array}{c:c} \alpha n{\otimes}n & (\delta{\cdot}n)n \\ \hdashline -(\delta{\cdot}n)n^t & \eta \end{array}\right] \tag{5.11}$$

where n is a unit normal vector to F (its orientation is clearly irrelevant to define $\mathcal{S}$). Then, properties (DG5a)–(DG8a) *hold.*

Proof. The field $\mathcal{S}$ is clearly positive and bounded, i.e., (DG5a) and (DG6a) hold. Moreover, for $\xi, \zeta \in \mathbb{R}^{d+1}$,

$$\begin{aligned} \zeta^t \mathcal{S} \xi &= \alpha(\xi_\sigma{\cdot}n)(\zeta_\sigma{\cdot}n) + (\delta{\cdot}n)(\zeta_\sigma{\cdot}n)\xi_u - (\delta{\cdot}n)(\xi_\sigma{\cdot}n)\zeta_u + \eta\xi_u\zeta_u, \\ (\xi^t \mathcal{S} \xi)^{\frac12} &= (\alpha(\xi_\sigma{\cdot}n)^2 + \eta\xi_u^2)^{\frac12}, \end{aligned}$$

whence (DG7a) is readily deduced. Finally, since

$$\zeta^t \mathcal{D}_{\partial K} \xi = (\xi_\sigma{\cdot}n_K)\zeta_u + (\zeta_\sigma{\cdot}n_K)\xi_u + (\beta{\cdot}n_K)\xi_u\zeta_u,$$

it is clear that (DG8a) holds. ∎

Owing to Definition 4.8 and Proposition 4.9, the local formulation (4.42) takes the following form: Seek $(\sigma_h, u_h) \in W_h$ such that $\forall K \in \mathcal{T}_h$, $\forall \tau_h \in [\mathbb{P}_p(K)]^d$, and $\forall v_h \in \mathbb{P}_p(K)$,

$$\left\{\begin{array}{l} (\sigma_h, \tau_h)_{L,K} - (u_h, \nabla{\cdot}\tau_h)_{L,K} + (\phi^\sigma_{\partial K}(\sigma_h, u_h), \tau_h)_{L,\partial K} = 0, \\ -(\sigma_h, \nabla v_h)_{L,K} + ((\mu - \nabla{\cdot}\beta)u_h, v_h)_{L,K} - (u_h, \beta{\cdot}\nabla v_h)_{L,K} \\ + (\phi^u_{\partial K}(\sigma_h, u_h), v_h)_{L,\partial K} = (f, v_h)_{L,K}, \end{array}\right. \tag{5.12}$$

with the interface fluxes

$$\phi^{\sigma,\mathrm{i}}(\sigma_h, u_h)|_{\partial K} = (\{u_h\} + \alpha n_K{\cdot}[\![\sigma_h]\!]_{\partial K} + (\delta{\cdot}n_K)[\![u_h]\!]_{\partial K})n_K, \tag{5.13}$$

$$\phi^{u,\mathrm{i}}(\sigma_h, u_h)|_{\partial K} = n_K{\cdot}\{\sigma_h\} - (\delta{\cdot}n_K)n_K{\cdot}[\![\sigma_h]\!]_{\partial K} + \eta[\![u_h]\!]_{\partial K} + \beta{\cdot}n_K\{u_h\}. \tag{5.14}$$

The boundary fluxes are specified below for the various boundary conditions.

Likewise, the local formulation (4.43) takes the following form: Seek $(\sigma_h, u_h) \in W_h$ such that $\forall K \in \mathcal{T}_h$, $\forall \tau_h \in [\mathbb{P}_p(K)]^d$, and $\forall v_h \in \mathbb{P}_p(K)$,

$$\left\{\begin{array}{l} (\sigma_h + \nabla u_h, \tau_h)_{L,K} + (\phi^{*,\sigma}_{\partial K}(\sigma_h, u_h), \tau_h)_{L,\partial K} = 0, \\ (\nabla{\cdot}\sigma_h + \beta{\cdot}\nabla u_h + \mu u_h, v_h)_{L,K} + (\phi^{*,u}_{\partial K}(\sigma_h, u_h), v_h)_{L,\partial K} = (f, v_h)_{L,K}, \end{array}\right. \tag{5.15}$$

with the interface adjoint-fluxes

$$\phi^{*,\sigma,\mathrm{i}}(\sigma_h,u_h)|_{\partial K} = (((\delta\cdot n_K)-\tfrac{1}{2})[\![u_h]\!]_{\partial K} + \alpha n_K\cdot[\![\sigma_h]\!]_{\partial K})n_K, \tag{5.16}$$

$$\phi^{*,u,\mathrm{i}}(\sigma_h,u_h)|_{\partial K} = -((\delta\cdot n_K)+\tfrac{1}{2})n_K\cdot[\![\sigma_h]\!]_{\partial K} + (\eta-\tfrac{1}{2}\beta\cdot n_K)[\![u_h]\!]_{\partial K}. \tag{5.17}$$

The boundary adjoint-fluxes are specified below for the various boundary conditions.

Remark 5.3.

(i) We stress the fact that (5.12) and (5.15) yield $(p+\frac{1}{2})$-order estimates in the L-norm for both u_h and σ_h.

(ii) Owing to the fact that $\alpha\neq 0$ in (5.11), the first equation in (5.12) or (5.15) cannot be used to derive a local reconstruction formula where $\sigma_h|_K$ is expressed solely in terms of u_h. To this end, the coefficient α has to be set to zero, and this requires a nontrivial modification of the analysis that will be reported in [14]. With this modification, the DG approximation does not yield a $(p+\frac{1}{2})$-order estimate for σ_h in the L-norm.

(iii) The design parameters α, δ, and η can vary from face to face. In particular, one can take δ to be any bounded vector-valued field on $\mathcal{F}_h^{\mathrm{i}}$; $\delta=0$ is a suitable choice. Other particular choices lead to DG methods already reported in the literature for advection–diffusion–reaction problems; for instance, the comparison with the method of Bassi and Rebay [7] and with the LDG method of Cockburn and Shu [10] will be discussed in [14]. However, the method of Baumann and Oden and its variants [8, 23, 26] cannot be directly recovered from (5.12) and (5.15) since these methods eliminate the unknown $\sigma_h|_K$ locally, and, therefore, require the design parameter α to be set to zero. A more detailed discussion is postponed to [14].

5.3.1. Dirichlet boundary conditions. The $\mathbb{R}^{d+1,d+1}$-valued field $\mathcal{M}$ associated with the operator M defined in (3.21) is

$$\mathcal{M} = \left[\begin{array}{c:c} 0 & -n \\ \hdashline n^t & 0 \end{array}\right]. \tag{5.18}$$

Lemma 5.6. *Let $\varsigma>0$. For all $x\in\partial\Omega$, set*

$$\widehat{\mathcal{M}} = \left[\begin{array}{c:c} 0 & -n \\ \hdashline n^t & \varsigma \end{array}\right]. \tag{5.19}$$

Then, properties (DG1a)–(DG4a) *hold.*

Proof. Properties (DG1a) and (DG2a) obviously hold. A straightforward calculation shows that for all $\xi,\,\zeta\in\mathbb{R}^{d+1}$,

$$|\zeta^t(\widehat{\mathcal{M}}-\mathcal{D}_{\partial\Omega})\xi| = |-2(\zeta_\sigma\cdot n)\xi_u + (\varsigma-\beta\cdot n)\zeta_u\xi_u| \le c|\xi_u|\,\|\zeta\|_{\mathbb{R}^{d+1}}, \tag{5.20}$$

and hence, (DG3a) holds since $|\xi_u|\le c(\xi^t\widehat{\mathcal{M}}\xi)^{\frac{1}{2}}$. The proof of (DG4a) is similar. □

The field $\widehat{\mathcal{M}}$ defined in (5.19) yields the boundary fluxes

$$\phi^{\sigma,\partial}(\sigma_h,u_h) = 0, \tag{5.21}$$

$$\phi^{u,\partial}(\sigma_h,u_h) = \tfrac{1}{2}(\varsigma+\beta\cdot n)u_h + \sigma_h\cdot n, \tag{5.22}$$

and the boundary adjoint-fluxes

$$\phi^{*,\sigma,\partial}(\sigma_h,u_h) = -u_h n, \tag{5.23}$$

$$\phi^{*,u,\partial}(\sigma_h,u_h) = \tfrac{1}{2}(\varsigma-\beta\cdot n)u_h. \tag{5.24}$$

Remark 5.4. Observe that setting $\widehat{\mathcal{M}}=\mathcal{M}$ is not adequate here since with this choice (DG3a) does not hold.

5.3.2. Neumann boundary conditions. To simplify, assume $(\beta\cdot n)|_{\partial\Omega} = 0$. The $\mathbb{R}^{d+1,d+1}$-valued field $\mathcal{M}$ associated with the operator M defined in (3.23) is

$$\mathcal{M} = \left[\begin{array}{c:c} 0 & n \\ \hdashline -n^t & 0 \end{array}\right]. \tag{5.25}$$

LEMMA 5.7. *Let $\lambda > 0$. For all $x \in \partial\Omega$, set*

$$\widehat{\mathcal{M}} = \left[\begin{array}{c:c} \lambda n\otimes n & n \\ \hdashline -n^t & 0 \end{array}\right]. \tag{5.26}$$

Then, properties (DG1a)–(DG4a) *hold.*

Proof. Clearly, $\widehat{\mathcal{M}}$ is positive since for all $\xi \in \mathbb{R}^{d+1}$, $\xi^t\widehat{\mathcal{M}}\xi = \lambda(\xi_\sigma\cdot n)^2$, i.e., (DG1a) holds. Moreover, if $\xi \in \mathrm{Ker}(\mathcal{M}-\mathcal{D}_{\partial\Omega})$, then $\xi_\sigma\cdot n = 0$ and hence, $\widehat{\mathcal{M}}\xi = \mathcal{M}\xi = \mathcal{D}_{\partial\Omega}\xi$, i.e., (DG2a) holds. To verify (DG3a), observe that for all ξ, $\zeta \in \mathbb{R}^{d+1}$,

$$|\zeta^t(\widehat{\mathcal{M}} - \mathcal{D}_{\partial\Omega})\xi| = |\lambda(\zeta_\sigma\cdot n)(\xi_\sigma\cdot n) - 2\zeta_u(\xi_\sigma\cdot n)| \le c|\xi_\sigma\cdot n|\|\zeta\|_{\mathbb{R}^{d+1}}, \tag{5.27}$$

showing that (DG3a) holds. Proceed similarly to verify (DG4a). ∎

The field $\widehat{\mathcal{M}}$ defined in (5.26) yields the boundary fluxes

$$\phi^{\sigma,\partial}(\sigma_h, u_h) = (\tfrac{1}{2}\lambda(\sigma_h\cdot n) + u_h)n, \tag{5.28}$$

$$\phi^{u,\partial}(\sigma_h, u_h) = 0. \tag{5.29}$$

and the boundary adjoint-fluxes

$$\phi^{*,\sigma,\partial}(\sigma_h, u_h) = \tfrac{1}{2}\lambda(\sigma_h\cdot n)n, \tag{5.30}$$

$$\phi^{*,u,\partial}(\sigma_h, u_h) = -(\sigma_h\cdot n). \tag{5.31}$$

Remark 5.5. The bilinear form $(u, v) \longmapsto \int_{\partial\Omega} v^t\widehat{\mathcal{M}}u$ cannot be extended to $W\times W$ due to the presence of the upper-left block in $\widehat{\mathcal{M}}$. The difficulty stems from the fact that vectors fields in $H(\mathrm{div};\Omega)$ may not have normal traces in $L^2(\partial\Omega)$. As a consequence, the approximate method is meaningful only if the exact solution is smooth enough; see the definition of $W(h)$ in (4.2).

5.3.3. Robin boundary conditions. As in §5.3.2, assume $(\beta\cdot n)|_{\partial\Omega} = 0$. Assume that the function $\rho \in L^\infty(\partial\Omega)$ is uniformly bounded away from zero. The $\mathbb{R}^{d+1,d+1}$-valued field $\mathcal{M}$ associated with the operator M defined in (3.25) is

$$\mathcal{M} = \left[\begin{array}{c:c} 0 & n \\ \hdashline -n^t & 2\varrho \end{array}\right]. \tag{5.32}$$

LEMMA 5.8. *Let $0 < \varsigma < 1$, set $\theta = 2\varsigma - 1$ and $\lambda = 2\frac{1-\varsigma}{\varrho}$. For all $x \in \partial\Omega$, set*

$$\widehat{\mathcal{M}} = \left[\begin{array}{c:c} \lambda n\otimes n & \theta n \\ \hdashline -\theta n^t & 2\varrho\varsigma \end{array}\right]. \tag{5.33}$$

Then, properties (DG1a)–(DG4a) *hold.*

Proof. Clearly, $\widehat{\mathcal{M}}$ is positive since for all $\xi \in \mathbb{R}^{d+1}$, $\xi^t\mathcal{M}\xi = \lambda(\xi_\sigma\cdot n)^2 + 2\varrho\varsigma\xi_u^2$ and both λ and ς are positive by assumption. Hence, (DG1a) holds. Moreover, if

$\xi \in \mathrm{Ker}(\mathcal{M} - \mathcal{D}_{\partial\Omega})$, then $\xi_\sigma{\cdot}n = \varrho\xi_u$ and a direct calculation yields $(\widehat{\mathcal{M}} - \mathcal{D}_{\partial\Omega})\xi = 0$. Hence, (DG2a) holds. To verify (DG3a), observe that for all ξ, $\zeta \in \mathbb{R}^{d+1}$,

$$|\zeta^t(\widehat{\mathcal{M}} - \mathcal{D}_{\partial\Omega})\xi| \le c(\xi_u^2 + (\xi_\sigma{\cdot}n)^2)^{\frac{1}{2}}\|\zeta\|_{\mathbb{R}^{d+1}}, \tag{5.34}$$

and proceed similarly to verify (DG4a). □

The field $\widehat{\mathcal{M}}$ defined in (5.33) yields the boundary fluxes

$$\phi^{\sigma,\partial}(\sigma_h, u_h) = \tfrac{1}{2}(\lambda(\sigma_h{\cdot}n) + (\theta+1)u_h)n, \tag{5.35}$$

$$\phi^{u,\partial}(\sigma_h, u_h) = \tfrac{1}{2}(1-\theta)(\sigma_h{\cdot}n) + \varrho\varsigma u_h, \tag{5.36}$$

and the boundary adjoint-fluxes

$$\phi^{*,\sigma,\partial}(\sigma_h, u_h) = \tfrac{1}{2}(\lambda(\sigma_h{\cdot}n) + (\theta-1)u_h)n, \tag{5.37}$$

$$\phi^{*,u,\partial}(\sigma_h, u_h) = -\tfrac{1}{2}(1+\theta)(\sigma_h{\cdot}n) + \varrho\varsigma u_h. \tag{5.38}$$

Remark 5.6.

(i) A simple choice for the field $\widehat{\mathcal{M}}$ is $\varsigma = \frac{1}{2}$, $\theta = 0$, and $\lambda = \frac{1}{\varrho}$, yielding

$$\widehat{\mathcal{M}} = \left[\begin{array}{c:c} \frac{1}{\varrho}n\otimes n & 0 \\ \hdashline 0 & \varrho \end{array}\right]. \tag{5.39}$$

(ii) As for Neumann boundary conditions, the bilinear form $(u,v) \longmapsto \int_{\partial\Omega} v^t\widehat{\mathcal{M}}u$ cannot be extended to $W \times W$ due to the presence of the upper-left block in $\widehat{\mathcal{M}}$.

5.4. Maxwell's equations in diffusive regime. We close this series of applications with Maxwell's equations in the diffusive regime; see §3.3. Let $K \in \mathcal{T}_h$. Let $n_K = (n_{K,1}, n_{K,2}, n_{K,3})^t$ be the unit outward normal to K on ∂K and introduce the $\mathbb{R}^{6,6}$-valued field $\mathcal{N}_K$ such that $\mathcal{N}_K = \sum_{k=1}^3 n_{K,k}\mathcal{R}^k$. Observe that for all $\xi \in \mathbb{R}^3$, $\mathcal{N}_K\xi = n_K \times \xi$. Then, the $\mathbb{R}^{6,6}$-valued field $\mathcal{D}_{\partial K}$ defined in (4.9) is given by

$$\mathcal{D}_{\partial K} = \left[\begin{array}{c:c} 0 & \mathcal{N}_K \\ \hdashline \mathcal{N}_K^t & 0 \end{array}\right]. \tag{5.40}$$

Furthermore, the $\mathbb{R}^{6,6}$-valued field $\mathcal{M}$ associated with the boundary operator M defined in (3.33) is given by

$$\mathcal{M} = \left[\begin{array}{c:c} 0 & -\mathcal{N} \\ \hdashline \mathcal{N}^t & 0 \end{array}\right], \tag{5.41}$$

where $\mathcal{N} = \sum_{k=1}^3 n_k\mathcal{R}^k$ and $n = (n_1, n_2, n_3)^t$ is the unit outward normal to Ω on $\partial\Omega$. It can be verified that the field $\mathcal{M}$ satisfies neither (DG3a) nor (DG4a). To remedy this weakness, we introduce a positive constant ς and we set

$$\widehat{\mathcal{M}} = \left[\begin{array}{c:c} 0 & -\mathcal{N} \\ \hdashline \mathcal{N}^t & \varsigma\mathcal{N}^t\mathcal{N} \end{array}\right]. \tag{5.42}$$

LEMMA 5.9. *Provided* $\varsigma > 0$, *properties* (DG1a)–(DG4a) *hold.*

Proof. For all $\xi = (\xi_h, \xi_e) \in \mathbb{R}^3 \times \mathbb{R}^3$, it is clear that $(\widehat{\mathcal{M}}\xi, \xi)_{\mathbb{R}^6} = \varsigma\|\mathcal{N}\xi_e\|_{\mathbb{R}^3}^2$, showing that (DG1a) holds. Moreover, if $\xi = (\xi_h, \xi_e) \in \mathrm{Ker}(\mathcal{M} - \mathcal{D})$, then $\mathcal{N}\xi_e = 0$

yielding $\xi \in \mathrm{Ker}(\widehat{\mathcal{M}} - \mathcal{D})$, i.e., (DG2a) holds. To prove that (DG3a) holds, let $\xi = (\xi_h, \xi_e) \in \mathbb{R}^3 \times \mathbb{R}^3$ and let $\zeta = (\zeta_h, \zeta_e) \in \mathbb{R}^3 \times \mathbb{R}^3$. A straightforward calculation yields

$$|\zeta^t(\widehat{\mathcal{M}} - \mathcal{D})\xi| = |-2(\mathcal{N}\xi_e, \zeta_h)_{\mathbb{R}^3} + \varsigma(\mathcal{N}\xi_e, \mathcal{N}\zeta_e)_{\mathbb{R}^3}| \le c\|\mathcal{N}\xi_e\|_{\mathbb{R}^3}\|\zeta\|_{\mathbb{R}^6}.$$

Since $\|\mathcal{N}\xi_e\|^2_{\mathbb{R}^3} = \varsigma^{-1}(\widehat{\mathcal{M}}\xi, \xi)_{\mathbb{R}^6}$, (DG3a) holds. Proceed similarly to prove (DG4a). ∎

Let $\alpha_1 > 0$ and $\alpha_2 > 0$. For all $F \in \mathcal{F}_h^{\mathrm{i}}$ and for a.e. $x \in F$, define

$$\mathcal{S} = \left[\begin{array}{c:c} \alpha_1 \mathcal{N}^t\mathcal{N} & 0 \\ \hdashline 0 & \alpha_2 \mathcal{N}^t\mathcal{N} \end{array}\right], \tag{5.43}$$

where $\mathcal{N} = \sum_{k=1}^3 n_k \mathcal{R}^k$ and $n = (n_1, n_2, n_3)^t$ is a unit outward normal to F (its orientation is clearly irrelevant to define $\mathcal{S}$).

LEMMA 5.10. *Provided $\alpha_1 > 0$ and $\alpha_2 > 0$, properties* (DG5a)–(DG8a) *hold.*

Proof. Observe that for all $\xi = (\xi_h, \xi_e) \in \mathbb{R}^3 \times \mathbb{R}^3$ and $\zeta = (\zeta_h, \zeta_e) \in \mathbb{R}^3 \times \mathbb{R}^3$,

$$\zeta^t\mathcal{S}\xi = \alpha_1(n\times\xi_h)\cdot(n\times\zeta_h) + \alpha_2(n\times\xi_e)\cdot(n\times\zeta_e).$$

Hence, $\mathcal{S}$ is positive, i.e., (DG5a) holds. Moreover, $\mathcal{S}$ is uniformly bounded, i.e., (DG6a) holds. In addition, since $\mathcal{S}$ is symmetric, (DG7a) results from (DG5a). Finally, for all $\xi = (\xi_h, \xi_e) \in \mathbb{R}^3 \times \mathbb{R}^3$ and $\zeta = (\zeta_h, \zeta_e) \in \mathbb{R}^3 \times \mathbb{R}^3$,

$$|\zeta^t\mathcal{D}_{\partial K}\xi| = |(\mathcal{N}_K\xi_e, \zeta_h)_{\mathbb{R}^3} + (\mathcal{N}_K^t\xi_h, \zeta_e)_{\mathbb{R}^3}| \le c(\|\mathcal{N}_K\xi_e\|_{\mathbb{R}^3} + |\|\mathcal{N}_K\xi_h\|_{\mathbb{R}^3})\|\zeta\|_{\mathbb{R}^6},$$

showing that (DG8a) holds. ∎

Owing to Definition 4.8 and Proposition 4.9, the local formulation (4.42) takes the following form: Seek $(H_h, E_h) \in W_h$ such that $\forall K \in \mathcal{T}_h$, $\forall \eta_h \in [\mathbb{P}_p(K)]^3$, and $\forall \psi_h \in [\mathbb{P}_p(K)]^3$,

$$\begin{cases} (\mu H_h, \eta_h)_{L,K} - (H_h, \nabla\times\psi_h)_{L,K} + (\phi^H_{\partial K}(H_h, E_h), \eta_h)_{L,\partial K} = (f, \eta_h)_{L,K}, \\ (\sigma E_h, \psi_h)_{L,K} + (E_h, \nabla\times\eta_h)_{L,K} + (\phi^E_{\partial K}(H_h, E_h), \psi_h)_{L,\partial K} = (g, \psi_h)_{L,K}, \end{cases} \tag{5.44}$$

with the interface fluxes

$$\phi^{H,\mathrm{i}}(H_h, E_h)|_{\partial K} = n_K\times(-\alpha_1 n_K\times[\![H_h]\!]_{\partial K} + \{E_h\}), \tag{5.45}$$

$$\phi^{E,\mathrm{i}}(H_h, E_h)|_{\partial K} = -n_K\times(\alpha_2 n_K\times[\![E_h]\!]_{\partial K} + \{H_h\}), \tag{5.46}$$

and the boundary fluxes

$$\phi^{H,\partial}(H_h, E_h) = 0, \tag{5.47}$$

$$\phi^{E,\partial}(H_h, E_h) = -n\times(H_h + \tfrac{1}{2}\varsigma(n\times E_h)). \tag{5.48}$$

Likewise, the local formulation (4.43) takes the following form: Seek $(H_h, E_h) \in W_h$ such that $\forall K \in \mathcal{T}_h$, $\forall \eta_h \in [\mathbb{P}_p(K)]^3$, and $\forall \psi_h \in [\mathbb{P}_p(K)]^3$,

$$\begin{cases} (\mu H_h + \nabla\times E_h, \eta_h)_{L,K} + (\phi^{*,H}_{\partial K}(H_h, E_h), \eta_h)_{L,\partial K} = (f, \eta_h)_{L,K}, \\ (\sigma E_h - \nabla\times H_h, \psi_h)_{L,K} + (\phi^{*,E}_{\partial K}(H_h, E_h), \psi_h)_{L,\partial K} = (g, \psi_h)_{L,K}, \end{cases} \tag{5.49}$$

with the interface adjoint-fluxes

$$\phi^{*,H,\mathrm{i}}(H_h, E_h)|_{\partial K} = -n_K\times(\alpha_1 n_K\times[\![H_h]\!]_{\partial K} + \tfrac{1}{2}[\![E_h]\!]_{\partial K}), \tag{5.50}$$

$$\phi^{*,E,\mathrm{i}}(H_h, E_h)|_{\partial K} = -n_K\times(\alpha_2 n_K\times[\![E_h]\!]_{\partial K} - \tfrac{1}{2}[\![H_h]\!]_{\partial K}), \tag{5.51}$$

and the boundary adjoint-fluxes

$$\phi^{*,H,\partial}(H_h, E_h) = -n\times E_h, \tag{5.52}$$

$$\phi^{*,E,\partial}(H_h, E_h) = -\tfrac{1}{2}\varsigma n\times(n\times E_h). \tag{5.53}$$

Remark 5.7. The design parameters α_1 and α_2 can vary from face to face.

REFERENCES

[1] D. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[2] D. Arnold, F. Brezzi, B. Cockburn, and L. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001/02), pp. 1749–1779.

[3] I. Babuška, *The finite element method with penalty*, Math. Comp., 27 (1973), pp. 221–228.

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

[5] C. Baiocchi, F. Brezzi, and L. Franca, *Virtual bubbles and Galerkin-Least-Squares type methods (GaLS)*, Comput. Methods Appl. Mech. Engrg., 105 (1993), pp. 125–141.

[6] G. Baker, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp., 31 (1977), pp. 45–59.

[7] 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 (1997), pp. 267–279.

[8] C. E. Baumann and J. T. Oden, *A discontinuous hp finite element method for convection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., 175 (1999), pp. 311–341.

[9] B. Cockburn, G. Karniadakis, and C. Shu, *Discontinuous Galerkin Methods - Theory, Computation and Applications*, vol. 11 of Lecture Notes in Computer Science and Engineering, Springer, 2000.

[10] B. Cockburn and C. Shu, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463.

[11] C. Dawson, *Godunov-mixed methods for advection-diffusion equations in multidimensions*, SIAM J. Numer. Anal., 30 (1993), pp. 1315–1332.

[12] ———, *Analysis of an upwind-mixed finite element methodfor nonlinear contaminant transport equations*, SIAM J. Numer. Anal., 35 (1998), pp. 1709–1724.

[13] J. Douglas Jr. and T. Dupont, *Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods*, vol. 58 of Lecture Notes in Physics, Springer-Verlag, Berlin, 1976.

[14] A. Ern and J.-L. Guermond, *Discontinuous Galerkin methods for Friedrichs' symmetric systems. II. Second-order PDEs*, SIAM J. Numer. Anal., (2004). to be submitted.

[15] ———, *Theory and Practice of Finite Elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, NY, 2004.

[16] K. Friedrichs, *Symmetric positive linear differential equations*, Comm. Pure Appl. Math., 11 (1958), pp. 333–418.

[17] J.-L. Guermond, *Subgrid stabilization of Galerkin approximations of linear monotone operators*, IMA J. Numer. Anal., 21 (2001), pp. 165–197.

[18] T. Hughes, L. Franca, and G. Hulbert, *A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/Least-Squares method for advection-diffusive equations*, Comput. Methods Appl. Mech. Engrg., 73 (1989), pp. 173–189.

[19] C. Johnson, U. Nävert, and J. Pitkäranta, *Finite element methods for linear hyperbolic equations*, Comput. Methods Appl. Mech. Engrg., 45 (1984), pp. 285–312.

[20] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp., 46 (1986), pp. 1–26.

[21] P. Lesaint, *Sur la résolution des sytèmes hyperboliques du premier ordre par des méthodes d'éléments finis*, PhD thesis, University of Paris VI, 1975.

[22] 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, Math. Res. Center, Univ. of Wisconsin-Madison, Academic Press, New York, 1974, pp. 89–123. Publication No. 33.

[23] J. Oden, I. Babuška, and C. Baumann, *A discontinuous hp finite element method for diffusion problems*, J. Comput. Phys., 146 (1998), pp. 491–519.

[24] J. Rauch, *Symmetric positive systems with boundary characteristic of constant multiplicity*, Trans. Amer. Math. Soc., 291 (1985), pp. 167–187.

[25] W. Reed and T. Hill, *Triangular mesh methods for the neutron transport equation*, Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.

[26] B. Rivière, M. Wheeler, and V. Girault, *Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. I*, Comput. Geosci., 8 (1999), pp. 337–360.

[27] E. Süli, C. Schwab, and P. Houston, *hp-DGFEM for partial differential equations with nonnegative characteristic form*, vol. 11 of Lecture Notes in Comput. Sci. Engrg., Springer-Verlag, New York, 2000, pp. 221–230. B. Cockburn, G.E. Karniadakis, and C.-W. Shu, eds.

[28] M. Wheeler, *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161.