Journal of Elliptic and Parabolic Equations (2024) 10:711–730
https://doi.org/10.1007/s41808-024-00274-8

On first order elliptic systems

Michaël Ndjinga[1] · Sédrick Kameni Ngwamou[2]

Received: 29 August 2022 / Accepted: 4 March 2024 / Published online: 1 April 2024

Abstract

The aim of this paper is the study of first-order stationary systems of PDEs of the form $\sum_k A_k \partial_k U + KU = 0$ with $K \not> 0$ on $\Omega = \mathbb{R}^d$ and $\Omega \subset \mathbb{R}^d$ bounded. We prove that the classical assumption $K > 0$ is not necessary for the well-posedness of the system and is violated in the particular case of the first-order Poisson problem. In the case $\Omega = \mathbb{R}^d$, we use Fourier analysis for the existence and uniqueness of solutions. For $\Omega \subset \mathbb{R}^d$ bounded, we use a complex analog of the *Banach–Nečas–Babuška* theorem to obtain the existence and uniqueness of a solution in a setting that encompasses both Friedrichs' systems and the first order reduction of the Poisson problem. The techniques used to prove the classical inf-sup conditions are inspired by harmonic analysis arguments that are consistent with the case $\Omega = \mathbb{R}^d$. In order to illustrate our approach, we study in detail the reduction of the Poisson equation to a first-order system.

Keywords First order systems · Elliptic systems · Friedrichs' systems · Inf-sup conditions

Mathematics Subject Classification 35F45 · 35J46 · 35J05

Michaël Ndjinga and Sédrick Kameni Ngwamou contributed equally to this work.

✉ Sédrick Kameni Ngwamou
sedrick.ngwamou@aims-cameroon.org

Michaël Ndjinga
michael.ndjinga@cea.fr

1 Université de Paris-Saclay, CEA Saclay, ISAS, DM2S, STMF, 91191 Gif-sur-Yvette, France

2 Department of Mathematics, Faculty of Science, University of Yaoundé I, BP: 812, Yaoundé, Cameroon

1 Introduction

Let $d, m \in \mathbb{N}^*$. We consider systems of first-order PDEs

$$\sum_{k=1}^{d} A_k \partial_k U(x) + KU(x) = F(x), \tag{1}$$

with unknown $U \in \left(L^2(\mathbb{R}^d)\right)^m$, where A_k, $k = 1, \ldots d$, and K are $m \times m$ real matrices and $F \in \left(L^2(\mathbb{R}^d)\right)^m$. The system (1) takes the conservative form $\nabla \cdot \mathbb{F}(U) + KU = F$ with linear flux $\mathbb{F}(U)\vec{\xi} = A(\vec{\xi})U$ and jacobian matrix $A(\vec{\xi}) = \sum_{k=1}^{d} \xi_k A_k$ where $\vec{\xi} = (\xi_1, \ldots, \xi_d)$.

The classical theory of Friedrichs systems ([1], [2, section 5.2]) covers the case of symmetric systems with $K > 0$, using a variational formulation and Lax-Milgram theorem. The coercivity required for the Lax-Milgram theorem is a consequence of the assumption $K > 0$.

However, in several important cases, one has to consider systems (1) with $K \not> 0$. The first class of examples is conservation laws with source terms in the stationary regime. In the particular case of gas dynamics, taking into account a friction force [3], a Coriolis force [4] or a chemical reaction [5, chapter 2 section 5] yields $K \not> 0$.

The second class consists of the mixed formulation of stationary diffusion problems $\nabla \cdot (D\vec{\nabla} u) = f$. The prototypical example of a diffusion equation is the Poisson problem whose first-order reduction is given in Sect. 4. Finite volume schemes for stationary diffusion on unstructured meshes can have a very complex design (see for instance [6]). The authors however believe that the discretization of the mixed formulation of stationary diffusion on unstructured meshes will yield simpler schemes with smaller stencils and linear systems that are larger but with a better condition number ($\mathcal{O}(\frac{1}{h})$ instead of $\mathcal{O}(\frac{1}{h^2})$).

The main objective of this paper is to lay the ground for the numerical analysis of finite volume methods for stationary first-order systems with $K \not> 0$. In this first account of our research, we emphasize the relations between the symbol A, the friction matrix K and the boundary condition operator $P_{\partial\Omega}$ that yield a well-posed problem.

We quickly review in Sect. 2 the well-posedness of the problem (1) on $\mathbb{R}^d$. We prove that $K > 0$ is not a necessary condition but only a particular case. There is even no need for K to be invertible as shown in Theorem 1 and Corollary 3. Then in Sect. 3 we investigate the case $\Omega \subset \mathbb{R}^d$ bounded. We use a complex analog of the BNB theorem [2], instead of the Lax-Milgram theorem to obtain a general existence result (see Theorem 5 and Corollary 2). In Sect. 4 we investigate the first order reduction of the Poisson problem with Dirichlet boundary condition. An existence result is given in Theorem 7 as an application of the approach laid in Sect. 3.

2 The case $\Omega = \mathbb{R}^d$

In this section, the solutions are sought for in $\left(L^2(\mathbb{R}^d)\right)^m$ and we give three theorems (Theorems 1, 2 and 6) that are straightforward applications of the Fourier transform on first order systems (1) as done in [7, section 3.1] for hyperbolic systems.
Existence of solutions to (1) is guaranteed provided $iA(\vec{\xi}) + K$ is an invertible matrix for almost every $\xi \in \mathbb{R}^d$ and $(iA(\vec{\xi}) + K)^{-1}\hat{F} \in \left(L^2(\mathbb{R}^d)\right)^m$.

Theorem 1 (Existence for first order systems) *Let $m, d \in \mathbb{N}^*$, $A_1, \ldots, A_d, K$ be $m \times m$ real matrices and $F \in \left(L^2(\mathbb{R}^d)\right)^m$ such that $iA(\vec{\xi}) + K$ is an invertible matrix for almost every $\xi \in \mathbb{R}^d$ and*

$$\left(iA(\vec{\xi}) + K\right)^{-1} \hat{F}(\vec{\xi}) \in \left(L^2(\mathbb{R}^d)\right)^m. \tag{2}$$

Then (1) admits a unique solution $U \in \left(L^2(\mathbb{R}^d)\right)^m$.

Proof Taking the Fourier transform, (1) is equivalent to $\left(iA(\vec{\xi}) + K\right)\hat{U}(\vec{\xi}) = \hat{F}(\vec{\xi})$, hence $\hat{U}(\vec{\xi}) = \left(iA(\vec{\xi}) + K\right)^{-1} \hat{F}(\vec{\xi})$ should be in $\left(L^2(\mathbb{R}^d)\right)^m$ for a unique solution to exist. □

We propose to define elliptic systems as systems such that $\det\left(iA(\vec{\xi}) + K\right) \neq 0$ almost everywhere. This definition is different from both the Petrovski [8–11] and ADN-ellipticity definitions [8, 9, 12], since they do not involve zero order terms, i.e. the matrix K in our case. In the case of the first order Poisson system (Sect. 4), the quantity $\det\left(iA(\vec{\xi}) + K\right) \neq 0$ matches exactly the symbol of the scalar Poisson problem (see Eq. 46).

Let $\sigma_{min}(M)$ denote the smallest singular value of a matrix M. The following theorem states that solutions to (1) exist provided the ratio between $\hat{F}$ and $\sigma_{min}(iA(\vec{\xi})+K)^{-1}$ is bounded in L^2. This condition is similar to the condition (40) found when studying the Poisson problem on $\mathbb{R}^d$.

Theorem 2 *Let $m, d \in \mathbb{N}^*$.*
Let $A_1, \ldots, A_d, K$ be $m \times m$ real matrices and $F \in \left(L^2(\mathbb{R}^d)\right)^m$ such that

$$\frac{\hat{F}}{\sigma_{\min}\left(iA(\vec{\xi}) + K\right)} \in \left(L^2(\mathbb{R}^d)\right)^m. \tag{3}$$

Then the system (1) admits a unique solution in $\left(L^2(\mathbb{R}^d)\right)^m$.

Proof Since $\left\|\left(iA(\vec{\xi}) + K\right)^{-1} \hat{F}\right\| \leq \frac{\|\hat{F}\|}{\sigma_{\min}\left(iA(\vec{\xi})+K\right)}$, (3) implies $\left(iA(\vec{\xi}) + K\right)^{-1} \hat{F} \in \left(L^2(\mathbb{R}^d)\right)^m$. Hence Theorem 1 yields the existence of a unique solution to (1) in $\left(L^2(\mathbb{R}^d)\right)^m$. □

The following corollary shows that Theorem 2 generalizes the well-known results about Friedrichs' systems (see [2, Section 5.2]) to the cases where $K \geq 0$.

Corollary 1 (Case of Friedrichs' systems) *Assume $A_k, k = 1, \dots, d$ and K are symmetric matrices, and that $K > 0$. Then for any $F \in \left(L^2(\mathbb{R}^d)\right)^m$, the system (1) admits a unique solution U in $\left(L^2(\mathbb{R}^d)\right)^m$.*

Proof We prove that $\sigma_{min}(iA(\vec{\xi}) + K)^{-1} \in L^\infty(\mathbb{R}^d)$ and then apply Theorem 2.

Since $A_k, k = 1, \dots, d$ and K are symmetric matrices, they are diagonalizable with real eigenvalues in an orthonormal basis of $\mathbb{R}^d$ and therefore

$$\begin{aligned} &\forall X \in \mathbb{C}^m, \quad {}^t\bar{X} A_k X \in \mathbb{R}, k = 1, \dots, d, \\ &\forall X \in \mathbb{C}^m, \quad {}^t\bar{X} K X \in \mathbb{R}, \end{aligned}$$

from which we deduce

$$\forall X \in \mathbb{C}^m, \vec{\xi} \in \mathbb{R}^d \quad |{}^t\bar{X}(iA(\vec{\xi}) + K)X| = |i{}^t\bar{X}A(\vec{\xi})X + {}^t\bar{X}KX)| \geq \lambda_{min}(K)||X||^2.$$

Since

$$\forall X \in \mathbb{C}^m, \vec{\xi} \in \mathbb{R}^d \quad |{}^t\bar{X}(iA(\vec{\xi}) + K)X| \leq \|X\| \, \|(iA(\vec{\xi}) + K)X\|,$$

we deduce

$$\forall X \in \mathbb{C}^m, \vec{\xi} \in \mathbb{R}^d \quad \lambda_{min}(K)\|X\|^2 \leq \|X\| \, \|(iA(\vec{\xi}) + K)X\|$$

and finally since $K > 0$ implies $\lambda_{min}(K) > 0$ we deduce

$$\sigma_{min}(iA(\vec{\xi}) + K)^{-1} \leq \frac{1}{\lambda_{min}(K)}.$$

Hence $\sigma_{min}(iA(\vec{\xi}) + K)^{-1} \in L^\infty(\mathbb{R}^d)$. As a consequence for any $F \in \left(L^2(\mathbb{R}^d)\right)^m$, $\|(iA(\vec{\xi}) + K)^{-1}\hat{F}\|_2 \leq \sigma_{min}(iA(\vec{\xi}) + K)^{-1}\|\hat{F}\|_2 \in L^2(\mathbb{R}^d)$ and Theorem 1 yields the existence of a unique solution U in $\left(L^2(\mathbb{R}^d)\right)^m$ to the system (1). □

Corollary 1 cannot be used for first-order systems with $K \not> 0$. An example is given with the first order reduction of the Poisson equation in Sect. 4.1, where the more general Theorem 1 is used instead of Corollary 1.

3 The case of a bounded Ω

In this section, Ω is assumed bounded with a Lipschitz boundary. As $\Omega \neq \mathbb{R}^d$, we can no longer make straightforward use of the Fourier transform to obtain a linear algebraic system. Also in this case of bounded Ω, boundary conditions are fundamental for solutions' existence and uniqueness.

A particular feature of bounded domains Ω is the existence of a Poincaré inequality, which allows for the coercivity of bilinear forms arising from weak formulations. In some simple cases such as the Poisson equation, Poincaré inequality enables the use of the Lax-Milgram theorem to derive the existence of a unique weak solution. However, in the case of first-order elliptic systems, existence and uniqueness are more tricky to prove. The classical theory [1, 13–15] of Friedrichs' systems makes some assumptions such as $K > 0$ that enable the use of the Lax–Milgram theorem to derive existence results (see for instance theorem 5.4 in [2]).

However, the assumption $K > 0$ is a serious obstacle to the analysis of linear symmetric hyperbolic systems in stationary regimes since these do not necessarily include a non-zero friction operator K (see for instance conservation laws in [7]). Another interesting example is the reduction of the Poisson equation to a first-order system with $K \not> 0$ (see Sect. 4). Giving up the assumption $K > 0$ yields a loss of coercivity and we can no longer use the Lax-Milgram theorem. Therefore, Theorem 5 and Corollary 2 use a complex analog of the more general *Banach–Nečas–Babuška* theorem (Sect. 3.5) to obtain the existence and uniqueness of a solution in a setting that encompasses both Friedrichs' systems and the first order reduction of the Poisson problem. The techniques used to prove the classical inf-sup conditions are inspired by harmonic analysis arguments that are consistent with the case $\Omega = \mathbb{R}^d$.

We consider a first order system of $m \in \mathbb{N}^*$ partial differential equations on $\Omega \subset \mathbb{R}^d$ taking the form

$$\sum_{k=1}^{d} A_k \partial_k U + KU = F \quad \text{on } \Omega \quad \text{(First order system)}, \tag{4}$$

$$P_{\partial\Omega}(\vec{\xi}_x)U(x) = 0 \quad \text{on } \partial\Omega \quad \text{(System Boundary conditions)}, \tag{5}$$

where $U \in \left(L^2(\Omega)\right)^m$ is the unknown and $F \in \left(L^2(\Omega)\right)^m$ the source term. The function $P_{\partial\Omega}$ is a symmetric projection matrix depending on $\vec{\xi}_x$, the outward normal vector to $\partial\Omega$ at $x \in \partial\Omega$.

3.1 Trace operator

The assumption that Ω is an open bounded domain with Lipschitz boundary $\partial\Omega$ allows for the existence of a trace operator $u \to u_{|\partial\Omega}$ for functions $u \in H^1(\Omega)$. The trace operator enables us to retrieve the boundary values of any $u \in H^1(\Omega)$ (see comment 7 of chapter 9 in [16]) and to set boundary conditions.

Thanks to this operator, it is also possible to state the Green–Ostrogradski formula for $f \in H^1(\Omega, \mathbb{C})$, $\vec{g} \in \left(H^1(\Omega, \mathbb{C})\right)^d$:

$$\int_{\Omega} \left(\vec{g} \cdot \vec{\nabla} f + f \nabla \cdot \vec{g}\right) dv = \int_{\partial\Omega} f \vec{g} \cdot \mathrm{d}\vec{s}. \tag{6}$$

3.2 Boundary conditions

Unlike scalar elliptic equations, we cannot impose a global boundary condition $U = U_b$ on $\partial\Omega$. Instead we can impose a condition on some components of U on some parts of $\partial\Omega$, as for instance: $u_1 = 0$ on $\partial\Omega_1 \subset \partial\Omega$.

In order to be able to impose such conditions on the boundary we assume the existence of a symmetric matrix-valued function $P_{\partial\Omega} : \mathbb{R}^d \to \mathcal{M}_m(\mathbb{R})$ such that $P_{\partial\Omega}$ is an orthogonal projector:

$$^tP_{\partial\Omega} = P_{\partial\Omega}, \quad P_{\partial\Omega}^2 = P_{\partial\Omega}.$$

The boundary condition (5) can be equivalently formulated as $U \in \ker P_{\partial\Omega}$ or $U \in \mathrm{Im}(\mathbb{I} - P_{\partial\Omega})$. We are thus led to seek a solution to (5) in the space

$$H^1_{P_{\partial\Omega}}(\Omega) = \left\{U \in \left(H^1(\Omega)\right)^m, U_{|\partial\Omega} \in \mathrm{Im}(\mathbb{I} - P_{\partial\Omega})\right\}.$$

3.3 Weak formulation

We define the $\mathbb{C}$-valued bilinear form a on $H^1_{P_{\partial\Omega}}(\Omega, \mathbb{C}) \times \left(L^2(\Omega, \mathbb{C})\right)^m$ as

$$a : H^1_{P_{\partial\Omega}}(\Omega, \mathbb{C}) \times \left(L^2(\Omega, \mathbb{C})\right)^m \to \mathbb{C} \tag{7}$$

$$(U, V) \mapsto \int_\Omega \left(\sum_{k=1}^d A_k \partial_k U + KU\right) \cdot \bar{V} dx, \tag{8}$$

U and V are taken $\mathbb{C}$-valued because harmonic analysis tools will be required later in our approach.

The weak formulation of problem (4), is to find $U \in H^1_{P_{\partial\Omega}}(\Omega, \mathbb{C})$ such that

$$\forall V \in \left(H^1(\Omega, \mathbb{C})\right)^m, \quad a(U, V) = \int_\Omega F \cdot \bar{V} dx. \tag{9}$$

We will then prove that the solution U is indeed in $H^1_{P_{\partial\Omega}}(\Omega, \mathbb{R})$ provided F is real valued (see Remark 2 later).

3.4 The first order operator $\mathcal{W}_{A,K}$

In the sequel, we study the first order operator $U \mapsto \sum_{k=1}^d A_k \partial_k U + KU$ by means of Fourier analysis. Since Ω is bounded, $L^2(\Omega) \subset L^1(\Omega)$ and we can use the expression of the L^1-Fourier transform on $\mathbb{R}^d$ to define $\hat{U}$ by extending U by zero outside Ω. We therefore define the Fourier transform $\mathcal{F}$ of a function $U \in \left(L^2(\Omega)\right)^m$ as

$$\mathcal{F}(U)(\vec{\xi}) = \hat{U}(\vec{\xi}) \equiv \frac{1}{(2\pi)^{\frac{d}{2}}} \int_\Omega U(x) e^{-i\vec{x}\cdot\vec{\xi}} \mathrm{d}x. \tag{10}$$

Similarly, the inverse Fourier transform is defined for $V \in \left(L^2(\mathbb{R}^d, \mathbb{C})\right)^m$ as the limit of the Fourier transform on $L^2 \cap L^1$:

$$\mathcal{F}^{-1}(V)(\vec{x}) = \hat{V}(-\vec{x}) \equiv \lim_{n\to\infty} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{B(0,n)} V(\vec{\xi})e^{i\vec{x}\cdot\vec{\xi}}\mathrm{d}\vec{\xi}. \tag{11}$$

We start by defining the following operator

$$\begin{aligned} \mathcal{W}_{A,K} : \left(H^1(\Omega)\right)^m &\to \left(L^2(\mathbb{R}^d), \mathbb{C}\right)^m \\ U &\mapsto \mathcal{F}\left(\sum_{k=1}^{d} A_k \partial_k U + KU\right). \end{aligned} \tag{12}$$

In order to find another expression for $\mathcal{W}_{A,K}$, we can use the Green–Ostrogradski formula (6) expressed as follows: for $U \in \left(H^1(\Omega, \mathbb{R})\right)^m$ and $v \in H^1(\Omega, \mathbb{R})$

$$\int_{\Omega} \left(\tilde{\mathrm{F}} \cdot \vec{\nabla} v + v\nabla \cdot \tilde{\mathrm{F}}\right) \mathrm{d}x = \int_{\partial\Omega} v\tilde{\mathrm{F}} \cdot \mathrm{d}\vec{s},$$

with the matrix $\tilde{\mathrm{F}} = (A_1U, \ldots, A_dU)$. For a given $\vec{\xi} \in \mathbb{R}^d$, we consider $v_\xi(\vec{x}) = e^{-i\vec{x}\cdot\vec{\xi}}$ and obtain

$$\begin{aligned} \mathcal{W}_{A,K}(U)(\vec{\xi}) &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\Omega} \left(\sum_{k=1}^{d} A_k\partial_k U + KU\right) e^{-i\vec{x}\cdot\vec{\xi}}\mathrm{d}x \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\Omega} \left(-\sum_{k=1}^{d} A_k\partial_k v_\xi(\vec{x}) + v_\xi(\vec{x})K\right) U\mathrm{d}x \\ &\quad + \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} v_\xi(\vec{x})A(\vec{\xi}_{\vec{x}})U(\vec{x})\mathrm{d}s \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\Omega} \left(\sum_{k=1}^{d} iA_k\xi_k + K\right) Ue^{-i\vec{x}\cdot\vec{\xi}}\mathrm{d}x \\ &\quad + \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} e^{-i\vec{x}\cdot\vec{\xi}} A(\vec{\xi}_{\vec{x}})U(\vec{x})\mathrm{d}s \\ &= (iA(\vec{\xi}) + K)\hat{U}(\vec{\xi}) + \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} e^{-i\vec{x}\cdot\vec{\xi}} A(\vec{\xi}_{\vec{x}})U(\vec{x})\mathrm{d}s. \end{aligned} \tag{13}$$

As was the case with $\Omega = \mathbb{R}^d$, we see with Eq. (13) for a bounded set $\Omega \subset \mathbb{R}^d$ that the Fourier transform of $\sum_{k=1}^{d} A_k\partial_k U + KU$ displays the symbol $(iA(\vec{\xi}) + K)\hat{U}(\vec{\xi})$ of the operator $\sum_{k=1}^{d} A_k\partial_k U + KU$. However, unlike the case with $\Omega = \mathbb{R}^d$, we see with Eq. (13) an extra contribution arising from the boundary of Ω.

We first give in Lemma 1 an extension of the Poincaré inequality to first-order systems, that will allow us to derive the first complex BNB condition (Lemma 4).

Lemma 1 (Generalised Poincaré inequality) *Assume that $\mathcal{W}_{A,K}$ is injective. Then*

$$\exists\alpha > 0, \forall U \in \left(H^1(\Omega)\right)^m, \quad \|\mathcal{W}_{A,K}(U)\|_{(L^2(\mathbb{R}^d))^m} \geq \alpha\|U\|_{(H^1(\Omega))^m}.$$

Proof The result follows from the open mapping theorem applied to the operator $\mathcal{W}_{A,K}$. Indeed

- $\mathcal{W}_{A,K}$ is a linear operator (see Eq. 12)
- $\mathcal{W}_{A,K}$ is a bounded operator since Parseval's identity yields

$$\begin{aligned}\|\mathcal{W}_{A,K}(U)\|_{(L^2(\mathbb{R}^d))^m} &= \|\mathcal{F}^{-1}(\mathcal{W}_{A,K}(U))\|_{(L^2(\mathbb{R}^d))^m}\\ &= \left\|\sum_{k=1}^{d} A_k\partial_k U + KU\right\|_{(L^2(\Omega))^m}\\ &\leq \sup\{|||K|||, |||A_1|||, \ldots, |||A_d|||\}\, \|U\|_{(H^1(\Omega))^m}.\end{aligned} \tag{14}$$

Therefore assuming that $\mathcal{W}_{A,K}$ is injective, $\mathcal{W}_{A,K}^{-1}$ is surjective from $\mathcal{W}_{A,K}((H^1(\Omega))^m) \subset (L^2(\mathbb{R}^d, \mathbb{C}))^m$ to $(H^1(\Omega))^m$. The open mapping theorem implies that $\mathcal{W}_{A,K}^{-1}$ is a continuous linear map, hence the result. □

Remark 1 (Connection with Poincaré inequality) We called Lemma 1 "Generalised Poincaré inequality" because using Parseval equality it yields

$$\forall U \in \left(H^1(\Omega)\right)^m, \quad \int_\Omega \left\|\sum_{k=1}^{d} A_k\partial_k U + KU\right\|_2^2 \mathrm{d}x \geq c\int_\Omega \|U\|_2^2 \mathrm{d}x.$$

We recover Poincaré inequality if $m = d$, $K = 0$, $\quad A_k = \vec{e}_k \otimes \vec{e}_k$ where $\vec{e}_k, k = 1, \ldots, d$ are the vectors forming the canonical basis of $\mathbb{R}^d$.

Note that the boundary conditions required for the classical Poincaré inequality are hidden behind the assumption that $\mathcal{W}_{A,K}$ be injective

3.5 The complex BNB theorem

Considering two Banach spaces V, W and a bilinear form $a \in \mathcal{L}(V \times W', \mathbb{R})$, the classical BNB conditions [2, Theorem 2.6]

$$\inf_{v\in V}\sup_{w\in W'} \frac{a(v,w)}{\|v\|_V\|w\|_{W'}} \geq \alpha, \tag{15}$$

$$\forall w \in W', \quad (\forall v \in V,\ a(v,w) = 0) \Rightarrow w = 0, \tag{16}$$

are deduced from the characterisation of bijective Banach operators (see [2, Theorem A.43]):

Theorem 3 (Characterisation of bijective operators) *Let V, W be two Banach spaces and $A \in \mathcal{L}(V; W)$ a continuous linear operator.*

A is bijective if and only if

1. *$A^T : W' \to V'$ is injective and*
2. *there exists $\alpha > 0$ such that*

$$\forall v \in V, \quad \|Av\|_W \geq \alpha \|v\|_V. \tag{17}$$

Then the classical BNB conditions are obtained under the assumption that W is reflexive and that $A \in \mathcal{L}(V; W)$ is associated with a real-valued bilinear form $a \in \mathcal{L}(V \times W', \mathbb{R})$ such that $< Av, w >_{W,W'} = a(v, w)$ (see [2, Corollary A.46]).

The first BNB condition (15) is therefore a direct consequence of the condition (17) in the real case where $a \in \mathcal{L}(V \times W', \mathbb{R})$.

In the complex case, we must assume instead that $A \in \mathcal{L}(V; W)$ is associated with a complex-valued bilinear form $a \in \mathcal{L}(V \times W^\natural, \mathbb{C})$ such that $< Av, w >_{W,W^\natural} = a(v, w)$, where $W^\natural$ is the space of complex-valued linear applications on W.

The analog of [2, Corollary A.46] in the complex case now follows.

Theorem 4 (The complex BNB conditions) *We assume that W is reflexive and that $A \in \mathcal{L}(V; W)$ is associated with a bilinear form $a \in \mathcal{L}(V \times W^\natural, \mathbb{C})$ such that $\langle Av, w \rangle_{W,W^\natural} = a(v, w)$.*

The following statements are equivalent:

1. *For all $f \in W$, there is a unique $u \in V$ such that $a(u, w) = \langle f, w \rangle_{W,W^\natural}$ for all $w \in W^\natural$.*
2. *There is $\alpha > 0$ such that*

$$\inf_{v \in V} \sup_{w \in W^\natural} \frac{|\ a(v, w)\ |}{\|v\|_V \|w\|_{W^\natural}} \geq \alpha, \tag{18}$$

$$\forall w \in W^\natural, \quad (\forall v \in V,\ a(v, w) = 0) \Rightarrow w = 0. \tag{19}$$

The proof of Theorem 4 will require the following lemma.

Lemma 2 *Let V be a normed vector space.*

$$\forall v \in V, \quad \|v\|_V = \sup_{l \in V^\natural, \|l\|_{V^\natural} = 1} |\langle v, l \rangle|_{V,V^\natural} = \sup_{l \in V^\natural} \frac{|\langle v, l \rangle|_{V,V^\natural}}{\|l\|_{V^\natural}}. \tag{20}$$

Lemma 2 is a complex analog of [2, Corollary A.17], which is a consequence of the Hahn-Banach theorem. Lemma 2 therefore relies on a complex version of the Hahn-Banach theorem that can be found in [17, Theorem 3.3]. A proof of Lemma 2 can indeed be found in the corollary following [17, Theorem 3.3].

Proof of Theorem 4 The first statement means that A is bijective, which thanks to Theorem 3 is equivalent to the two conditions

1.

$$\forall w \in W^\natural, \quad (\forall v \in V,\ a(v,w) = 0) \Rightarrow w = 0, \tag{21}$$

which is identical to (19).

2.

$$\exists \alpha > 0, \forall v \in V, \quad \|Av\|_W \geq \alpha \|v\|_V, \tag{22}$$

which is equivalent thanks to Lemma 2 to

$$\exists \alpha > 0, \forall v \in V, \quad \sup_{w \in W^\natural} \frac{|\langle Av, w\rangle|_{W,W^\natural}}{\|w\|_{W^\natural}} \geq \alpha \|v\|_V. \tag{23}$$

Taking the infimum over v in (23) yields the equivalence with (18).

□

3.6 The main result

In the classical study of first-order Friedrichs systems, the assumption $K > 0$ yields the coercivity and the Lax–Milgram theorem [2, lemma 2.2] yields well-posedness (see assumption F2 in [2, theorem 5.7]). Unfortunately, when $K \not> 0$ we do not have coercivity. We can prove the well-posedness thanks to a complex analog of the BNB theorem (Theorem 4 in Sect. 3.5), which has weaker assumptions than the Lax–Milgram theorem.

In order to use the complex BNB theorem (Theorem 4 in Sect. 3.5) and obtain the existence of a solution to the weak formulation (9), we need to prove the following two conditions:

$$\exists \alpha > 0, \forall U \in H^1_{P_{\partial\Omega}}(\Omega), \quad \sup_{V \in (L^2(\Omega,\mathbb{C}))^m} \frac{|\ a(U,V)\ |}{\|V\|_{(L^2(\Omega,\mathbb{C}))^m}} \geq \alpha \|U\|_{(H^1(\Omega,\mathbb{C}))^m}, \tag{24}$$

$$\forall V \in \left(L^2(\Omega,\mathbb{C})\right)^m, \quad (\forall U \in H^1_{P_{\partial\Omega}}(\Omega),\ a(U,V) = 0) \Rightarrow V = 0. \tag{25}$$

Since we do not have the coercivity property $a(U,U) \geq \gamma \|U\|^2_{H_1}$, we are going to find an injective function $\mathcal{W}(U)$ such that $a(U,\mathcal{W}(U)) \geq \gamma \|U\|^2_{H_1}$. The operator $\mathcal{W}(U) = \mathcal{F}^{-1}(\mathcal{W}_{A,K}(U)) = \sum_{k=1}^d A_k \partial_k U + KU$ with $\mathcal{W}_{A,K}(U)$ defined in (13) is our candidate. The following lemma connects the bilinear form a to $\mathcal{W}$, and will play an important role in proving the first BNB condition (24).

Lemma 3 (Connecting a and $\mathcal{W}$). *For any $U \in H^1_{P_{\partial\Omega}}(\Omega)$, we define*

$$\mathcal{W}(U) = \mathcal{F}^{-1}(\mathcal{W}_{A,K}(U)) = \sum_{k=1}^d A_k \partial_k U + KU.$$

Then $\mathcal{W}(U) \in \left(L^2(\Omega,\mathbb{C})\right)^m$ *and we have*

$$a(U,\mathcal{W}(U)) = \|\mathcal{W}_{A,K}(U)\|^2_{L^2(\Omega,\mathbb{C})}. \tag{26}$$

Proof $\mathcal{W}(U) \in \left(L^2(\Omega,\mathbb{C})\right)^m$ is a consequence of $U \in \left(H^1(\Omega,\mathbb{C})\right)^m$.
From the definition of a (Eq. 8) we have

$$\begin{aligned} a(U,\mathcal{W}(U)) &= \int_\Omega \left(\sum_{k=1}^d A_k\partial_k U + KU\right)\cdot\left(\sum_{k=1}^d A_k\partial_k U + KU\right)\mathrm{d}x \\ &= \int_\Omega \left\|\sum_{k=1}^d A_k\partial_k U + KU\right\|^2 \mathrm{d}x \end{aligned}$$

and the result follows from Parseval's identity:

$$a(U,\mathcal{W}(U)) = \int_{\mathbb{R}^d} \|\mathcal{F}^{-1}(\mathcal{W}_{A,K}(U))(\vec{\xi})\|^2\mathrm{d}\vec{\xi}. \tag{27}$$

□

Thanks to Lemmas 1 and 3 we have the following lemma on the first complex BNB condition.

Lemma 4 (First complex BNB condition) *Assume that* $\mathcal{W}_{A,K}$ *is injective on* $H^1_{P_{\partial\Omega}}(\Omega)$. *Then the first complex BNB condition* (24) *is true*:

$$\exists\alpha>0,\quad \forall U\in H^1_{P_{\partial\Omega}}(\Omega),\qquad \sup_{V\in\left(L^2(\Omega,\mathbb{C})\right)^m}\frac{|\,a(U,V)\,|}{\|V\|_{\left(L^2(\Omega,\mathbb{C})\right)^m}} \geq \alpha\|U\|_{\left(H^1(\Omega,\mathbb{C})\right)^m}.$$

Proof Starting from

$$\forall U\in H^1_{P_{\partial\Omega}}(\Omega),\qquad \sup_{V\in\left(L^2(\Omega,\mathbb{C})\right)^m}\frac{|\,a(U,V)\,|}{\|V\|_{\left(L^2(\Omega,\mathbb{C})\right)^m}} \geq \frac{a(U,\mathcal{W}(U))}{\|\mathcal{W}(U)\|_{\left(L^2(\Omega,\mathbb{C})\right)^m}},$$

Lemma 3 and the Parseval's identity yield

$$\begin{aligned} \forall U\in H^1_{P_{\partial\Omega}}(\Omega),\qquad \sup_{V\in\left(L^2(\Omega,\mathbb{C})\right)^m}\frac{|\,a(U,V)\,|}{\|V\|_{\left(L^2(\Omega,\mathbb{C})\right)^m}} &\geq \frac{\|\mathcal{W}_{A,K}\|^2_{L^2(\Omega,\mathbb{C})}}{\|\mathcal{W}_{A,K}\|_{L^2(\Omega,\mathbb{C})}} \\ &\geq \|\mathcal{W}_{A,K}\|_{L^2(\Omega,\mathbb{C})}. \end{aligned}$$

Now the Generalised Poincaré identity (Lemma 1) yields the result. □

The following theorem proves that the invertibility of $iA(\vec{\xi})+K$ yields the second complex BNB condition (25).

Lemma 5 (Second complex BNB condition) *Assume that* $iA(\vec{\xi}) + K$ *is invertible for almost every* $\vec{\xi}$*. Then the second complex BNB condition (25) is true*:

$$\forall V \in \left(L^2(\Omega, \mathbb{C})\right)^m, \quad (\forall U \in H^1_{P_{\partial\Omega}}(\Omega),\ a(U, V) = 0) \Rightarrow V = 0.$$

Proof Let $V \in \left(L^2(\Omega, \mathbb{C})\right)^m$ such that

$$\forall U \in H^1_{P_{\partial\Omega}}(\Omega), \quad a(U, V) = 0.$$

We define

$$W_V(\vec{\xi}) = {}^t\left(-iA(\vec{\xi}) + K\right)\hat{V}(\vec{\xi}).$$

Since $e^{i\vec{x}\cdot\vec{\xi}}W_V(\vec{\xi}) \notin H^1_{P_{\partial\Omega}}$, we cannot write "$a(e^{i\vec{x}\cdot\vec{\xi}}W_V(\vec{\xi}), V) = 0 = \|W_V\|^2$" to deduce $V = 0$. Hence we are first going to approximate $e^{i\vec{x}\cdot\vec{\xi}}W_V(\vec{\xi})$ with functions in $\left(\mathcal{C}^\infty_K(\Omega)\right)^m \subset H^1_{P_{\partial\Omega}}$, (where $\mathcal{C}^\infty_K$ is the space of smooth functions with compact support in Ω), and then pass to the limit.

$\mathcal{C}^\infty_K(\Omega)$ is dense in $H^1(\Omega)$ (see [16, Corollary 9.8]). For any $\vec{\xi} \in \mathbb{R}^m$, let a sequence $\Phi_n \in \left(\mathcal{C}^\infty_K(\Omega)\right)^m$, $n \in \mathbb{N}$ such that $\Phi_n(\cdot, \vec{\xi}) \to e^{i\vec{x}\cdot\vec{\xi}}$ in $H^1_{P_{\partial\Omega}}$. The function $\Phi_n(\cdot, \vec{\xi})W_V(\vec{\xi})$ is in $H^1_{P_{\partial\Omega}}$ so we can compute

$$\begin{aligned}
&a\left(\frac{1}{(2\pi)^{\frac{d}{2}}}\Phi_n(\cdot, \vec{\xi})W_V(\vec{\xi}), V\right) \\
&= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_\Omega \left(\sum_{k=1}^d A_k\partial_k\Phi_n(\cdot, \vec{\xi})W_V(\vec{\xi}) + K\Phi_n(\cdot, \vec{\xi})W_V(\vec{\xi})\right)\cdot\bar{V}\mathrm{d}x \\
&= {}^tW_V(\vec{\xi})\frac{1}{(2\pi)^{\frac{d}{2}}}\int_\Omega \left(\sum_{k=1}^d {}^tA_k\partial_k\Phi_n(\cdot, \vec{\xi}) + {}^tK\Phi_n(\cdot, \vec{\xi})\right)\bar{V}\mathrm{d}x.
\end{aligned}$$

Passing to the limit, since $\Phi_n(\cdot, \vec{\xi}) \to e^{i\vec{x}\cdot\vec{\xi}}$ in $H^1_{P_{\partial\Omega}}$ we obtain

$$\begin{aligned}
&\lim_{n\to\infty} a\left(\frac{1}{(2\pi)^{\frac{d}{2}}}\Phi_n(\cdot, \vec{\xi})W_V(\vec{\xi}), V\right) \\
&= {}^tW_V(\vec{\xi})\frac{1}{(2\pi)^{\frac{d}{2}}}\int_\Omega \left(\sum_{k=1}^d {}^tA_k\partial_k e^{i\vec{x}\cdot\vec{\xi}} + {}^tKe^{i\vec{x}\cdot\vec{\xi}}\right)\bar{V}\mathrm{d}x \\
&= {}^tW_V(\vec{\xi})\left(i{}^tA(\vec{\xi}) + {}^tK\right)\frac{1}{(2\pi)^{\frac{d}{2}}}\int_\Omega e^{i\vec{x}\cdot\vec{\xi}}\bar{V}\mathrm{d}x \\
&= {}^tW_V(\vec{\xi})\left(i{}^tA(\vec{\xi}) + {}^tK\right)\hat{\bar{V}}
\end{aligned}$$

$$= {}^t W_V(\vec{\xi})^t \left(\overline{-iA(\vec{\xi}) + K}\right) \bar{\hat{V}}$$
$$= {}^t W_V(\vec{\xi}) \bar{W}_V(\vec{\xi}) = ||W_V(\vec{\xi})\|_2^2.$$

Since $a(U, V) = 0$, we deduce

$${}^t \left(-iA(\vec{\xi}) + K\right) \hat{V} = 0.$$

Hence if $iA(\vec{\xi}) + K$ is invertible for almost every $\vec{\xi}$, so is ${}^t(-iA(\vec{\xi}) + K) = {}^t\overline{(iA(\vec{\xi}) + K)}$, and we can deduce that $V = 0$ almost everywhere. □

We can now prove our main result Theorem 5.

Theorem 5 (Well-posedness of first-order systems) *Let $d, m \in \mathbb{N}^*$.*

Let $A_k, k = 1, \dots d$, $P_{\partial\Omega}$, and K be $m \times m$ real matrices. Let Ω be a bounded open Lipschitz subset of $\mathbb{R}^d$. Assume that

- *$\mathcal{W}_{A,K}$ (Eqs. (12) and (13)) is injective on $H^1_{P_{\partial\Omega}}(\Omega)$.*
- *$iA(\vec{\xi}) + K$ is invertible for almost every $\vec{\xi}$.*

Then there exists a unique solution $U \in H^1_{P_{\partial\Omega}}(\Omega, \mathbb{C})$ to the variational problem (9) and

$$\exists\gamma, \forall F \in \left(H^1(\Omega, \mathbb{C})'\right)^m, \quad \|U\|_{H^1_{P_{\partial\Omega}}(\Omega,\mathbb{C})} \leq \gamma \|F\|_{(H^1(\Omega,\mathbb{C})')^m}. \tag{28}$$

Proof The theorem is a direct consequence of the complex BNB theorem (Theorem 4 in Sect. 3.5) with $W = H^1_{P_{\partial\Omega}}(\Omega, \mathbb{C})$, $V = \left(H^1(\Omega, \mathbb{C})\right)^m$. The constant γ is the inverse of the inf-sup constant from the first complex BNB condition. In our case, it follows from the open mapping theorem (see Lemma 1). □

Remark 2 (On the real character of the solutions) Theorems 5 and 7 give the existence of a complex-valued function $U \in \left(H^1(\Omega, \mathbb{C})\right)^m$. Their proofs do not require real matrices $A_k, k = 1, \dots, d$, $P_{\partial\Omega}$, K nor a real vector F and could be extended to complex coefficient matrices. However the assumption of real matrices $A_k, k = 1, \dots, d$, $P_{\partial\Omega}$, K and a real vector F implies $U \in \left(H^1(\Omega, \mathbb{R})\right)^m$ since otherwise $\bar{U}$ would be another distinct solution.

In the sequel, we investigate the connections between the matrices A, K and $P_{\partial\Omega}$ in order to satisfy the first complex BNB condition (24), and deduce the well-posedness of the system. For simplicity, we consider only symmetric systems. Since we can always multiply Eq. (4) by $sign(K)$, we assume in the following corollaries that $K \geq 0$.

Corollary 2 (Boundary compatibility condition, symmetric case) *Let Ω be a bounded set with a Lipschitz boundary, and $A_k, k = 1, \dots, d$, K, $P_{\partial\Omega}$ be symmetric matrices with $K \geq 0$.*

Assume that

- *$iA(\vec{\xi}) + K$ is invertible for almost every $\vec{\xi} \in \mathbb{R}^d$*

- $$\forall \vec{\xi} \in \mathbb{R}^d,\ \left(\mathbb{I}_m - P_{\partial\Omega}(\vec{\xi})\right) A(\vec{\xi}) \left(\mathbb{I}_m - P_{\partial\Omega}(\vec{\xi})\right) = 0 \tag{29}$$

- $$\forall \vec{\xi} \in \mathbb{R}^d,\ \ker K \cap \ker P_{\partial\Omega}(\vec{\xi}) = \{0\}. \tag{30}$$

Then there exists a unique solution $U \in H^1_{P_{\partial\Omega}}(\Omega, \mathbb{R})$ to the variational problem (9) *and*

$$\exists \gamma, \forall F \in \left(H^1(\Omega, \mathbb{R})'\right)^m, \quad \|U\|_{H^1_{P_{\partial\Omega}}(\Omega,\mathbb{R})} \leq \gamma \|F\|_{(H^1(\Omega,\mathbb{R})')^m}. \tag{31}$$

Proof Let $U \in \ker \mathcal{W}_{A,K}$. We are going to prove that $U = 0$ and deduce that $\mathcal{W}_{A,K}$ is injective. From Eq. (13) we have

$$\text{For a.e. } \vec{\xi} \in \mathbb{R}^d, \quad (iA(\vec{\xi}) + K)\hat{U}(\vec{\xi}) = -\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} e^{-i\vec{x}\cdot\vec{\xi}} A(\vec{\xi}_{\vec{x}}) U(\vec{x}) \mathrm{d}s. \tag{32}$$

Hence taking the inner product with $\hat{U}(\vec{\xi})$ yields for a.e. $\vec{\xi} \in \mathbb{R}^d$,

$$^t\bar{\hat{U}}(\vec{\xi})(iA(\vec{\xi}) + K)\hat{U}(\vec{\xi}) = -\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} e^{-i\vec{x}\cdot\vec{\xi}}\, {}^t\bar{\hat{U}}(\vec{\xi}) A(\vec{\xi}_{\vec{x}}) U(\vec{x}) \mathrm{d}s. \tag{33}$$

after integration with respect to $\vec{\xi}$ on the ball $B(0, n),\ n \in \mathbb{N}$, we obtain

$$\int_{B(0,n)} {}^t\bar{\hat{U}}(\vec{\xi})(iA(\vec{\xi})+K)\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi} = -\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{B(0,n)} \int_{\partial\Omega} e^{-i\vec{x}\cdot\vec{\xi}}\, {}^t\bar{\hat{U}}(\vec{\xi}) A(\vec{\xi}_{\vec{x}}) U(\vec{x}) \mathrm{d}s\mathrm{d}\vec{\xi},$$

and Fubini theorem yields

$$\begin{aligned}&\int_{B(0,n)} {}^t\bar{\hat{U}}(\vec{\xi})(iA(\vec{\xi}) + K)\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi}\\ &\quad = -\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} {}^t\left(\int_{B(0,n)} e^{-i\vec{x}\cdot\vec{\xi}} \bar{\hat{U}}(\vec{\xi}) dx\right) A(\vec{\xi}_{\vec{x}}) U(\vec{x}) \mathrm{d}s.\end{aligned}$$

Taking the limit $n \to \infty$, since $U \in H^1(\Omega)$ we find

$$\int_{\mathbb{R}^d} {}^t\bar{\hat{U}}(\vec{\xi})(iA(\vec{\xi}) + K)\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi} = -\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\partial\Omega} {}^t\bar{U}(\vec{x}) A(\vec{\xi}_{\vec{x}}) U(\vec{x}) \mathrm{d}s.$$

We recall that we seek U in the space $H^1_{P_{\partial\Omega}}(\Omega, \mathbb{R})$ and as such, the boundary condition (5) is equivalent to $U_{|\partial\Omega} \in \mathrm{Im}(\mathbb{I} - P_{\partial\Omega})$ ($P_{\partial\Omega}$ is a projector). As a consequence, the assumption (29) yields ${}^t\bar{U}(\vec{x}) A(\vec{\xi}_{\vec{x}}) U(\vec{x}) = 0$ on $\partial\Omega$.

Hence

$$\int_{\mathbb{R}^d} {}^t\bar{\hat{U}}(\vec{\xi})(iA(\vec{\xi}) + K)\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi} = 0. \tag{34}$$

Now since we assumed that $A(\vec{\xi})$ and K are symmetric matrices, ${}^t\bar{\hat{U}}A\hat{U} \in \mathbb{R}$ and ${}^t\bar{\hat{U}}K\hat{U} \in \mathbb{R}$. (34) thus implies

$$\int_{\mathbb{R}^d} {}^t\bar{\hat{U}}(\vec{\xi})A(\vec{\xi})\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi} = 0, \text{ and } \int_{\mathbb{R}^d} {}^t\bar{\hat{U}}(\vec{\xi})K\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi} = 0.$$

Hence since $K \geq 0$ we have

$$\text{for a.e.} \quad \vec{\xi} \in \mathbb{R}^d, \ \hat{U}(\vec{\xi}) \in \ker K,$$

which in turn implies that the L^2-Fourier transform of $\hat{U}$ is in the kernel of K:

$$U(x) = \lim_{n\to\infty} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{B(0,n)} e^{i\vec{x}\cdot\vec{\xi}}\hat{U}(\vec{\xi})\mathrm{d}\vec{\xi} \in \ker K.$$

We recall that $\hat{U}$ does not necessarily belong to L^1, hence we cannot use the L^1 formula for the Fourier transform. Instead, the L^2 Fourier transform is obtained as a limit using the density of $L^1 \cap L^2$ in L^2.

Now the boundary condition is $U(x) \in \ker P_{\partial\Omega}(\vec{\xi}_x)$ for x on the boundary so

$$\forall x \in \partial\Omega, \ U(x) \in \ker P_{\partial\Omega}(\vec{\xi}_x) \cap \ker K.$$

Using the boundary assumption (30), $U(x)$ is therefore 0 on the boundary. Hence the right-hand side of Eq. (32) vanishes and

$$\text{for a.e. } \vec{\xi} \in \mathbb{R}^d, \quad iA(\vec{\xi})\hat{U}(\xi) + K\hat{U}(\vec{\xi}) = 0.$$

Now from the first assumption, $iA(\vec{\xi}) + K$ is invertible and we conclude that $U = 0$.

Hence $\mathcal{W}_{A,K}$ is injective and from Theorem 5 we deduce the existence of a unique solution $U \in H^1_{P_{\partial\Omega}}(\Omega, \mathbb{C})$. Now since the matrices A_k, $k = 1, \ldots, d$, K, $P_{\partial\Omega}$ and the vector F are real, $\bar{U}$ is also solution of (9). Uniqueness yields $\bar{U} = U$ and therefore U is real valued. □

4 The first order reduction of the Poisson equation

In this section, we illustrate the reduction of the second-order Poisson equation to a first-order system. This first-order reduction yields a symmetric first-order system with $K \geq 0$ but $K \not> 0$, and therefore classical results on Friedrichs systems do not apply. We start with $\Omega = \mathbb{R}^d$ in Sect. 4.1 and then on a bounded Ω endowed with a boundary condition in Sect. 4.2.

Let Ω be an open subset of $\mathbb{R}^d$. The first-order reduction of the Poisson problem

$$-\Delta u = f \quad \text{on } \Omega, \tag{35}$$

amounts to defining $\vec{v} = \vec{\nabla} u$ and solving the first-order symmetric system that takes the form (1)

$$\begin{pmatrix} 0 & -\nabla\cdot \\ -\vec{\nabla} & 0 \end{pmatrix}\begin{pmatrix} u \\ \vec{v} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I}_d \end{pmatrix}\begin{pmatrix} u \\ \vec{v} \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \quad \text{on } \Omega. \tag{36}$$

We introduce the following symmetric matrices, unknown and right-hand side vectors

$$A_{Poisson}(\vec{\xi}) = \begin{pmatrix} 0 & {}^t\vec{\xi} \\ \vec{\xi} & 0 \end{pmatrix}, \quad K_{Poisson} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I}_d \end{pmatrix}, \tag{37}$$

$$U = \begin{pmatrix} u \\ \vec{\nabla} u \end{pmatrix}, \quad F_{Poisson}(x) = \begin{pmatrix} f(x) \\ 0 \end{pmatrix}. \tag{38}$$

4.1 The case $\Omega = \mathbb{R}^d$

In this case the solution is sought for in $\left(L^2(\mathbb{R}^d)\right)^{d+1}$ using Fourier transform. As $\partial\mathbb{R}^d = \emptyset$, there is no need to impose boundary conditions. We cannot use Corollary 1 for the well-posedness of the system (36) since $K \not> 0$. Theorem 1 should be used instead.

Corollary 3 (Existence for first order Poisson system—$\Omega = \mathbb{R}^d$) *Let* $d \in \mathbb{N}^*$, $f \in L^2(\mathbb{R}^d)$ *such that*

$$\frac{\hat{f}}{\|\vec{\xi}\|^2} \in L^2(\mathbb{R}^d), \quad \frac{\hat{f}}{\|\vec{\xi}\|^2}\vec{\xi} \in \left(L^2(\mathbb{R}^d)\right)^d. \tag{39}$$

Let the matrices $A_{poisson}(\vec{\xi})$, $K_{poisson}$ *and* $F_{poisson}(\vec{\xi})$ *defined in* (37). *Then the system* (36) *admits a unique solution* $(u, \vec{v}) \in \left(L^2(\mathbb{R}^d)\right)^{d+1}$.

Proof Since $\left(iA_{Poisson}(\vec{\xi}) + K_{Poisson}\right)^{-1} \hat{F}_{Poisson}(\vec{\xi}) = \frac{\hat{f}(\vec{\xi})}{\|\vec{\xi}\|^2}\begin{pmatrix} 1 \\ -i\vec{\xi} \end{pmatrix}$, according to Theorem 1, the assumptions (39) yield the existence of a unique solution in $L^2(\mathbb{R}^d) \times \left(L^2(\mathbb{R}^d)\right)^d$ (actually $H^1(\mathbb{R}^d) \times \left(L^2(\mathbb{R}^d)\right)^d$). □

The assumption (39) on f in Corollary 3 is to be compared with assumption (40) when the original second order problem is solved in L^2.

Theorem 6 (Existence for the scalar Poisson equation) *Let* $d \in \mathbb{N}^*$, $f \in L^2(\mathbb{R}^d)$ *such that*

$$\frac{\hat{f}}{\|\vec{\xi}\|^2} \in L^2(\mathbb{R}^d). \tag{40}$$

Then there exists a unique $u \in L^2(\mathbb{R}^d)$ *such that* $-\Delta u = f$ *in the weak (distributional) sense.*

Theorem 3 has more stringent assumptions than Theorem 6. Indeed the first order reduction require that $v = \vec{\nabla} u \in \left(L^2(\mathbb{R}^d)\right)^d$, hence the solution u given in Corollary 3 is actually in $H^1(\mathbb{R}^d)$. The first order reduction is not able to represent very weak solutions $u \notin H^1(\mathbb{R}^d)$. However, this should not be a serious issue since $H^1(\mathbb{R}^d)$ is the usual solution space in classical variational formulations.

4.2 The case of bounded domain $\Omega \subset \mathbb{R}^d$

We consider the Poisson equation (35) with the Dirichlet boundary condition

$$u = 0 \quad \text{on } \partial\Omega. \tag{41}$$

The Dirichlet boundary condition $u = 0$ can be enforced on the first order system (36) as $P_{\partial\Omega} U = 0$ using a constant Dirichlet boundary operator defined as follows

$$P_{\partial\Omega}^{Dirichlet}(\vec{\xi}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{42}$$

The following theorem gives an existence result in $H^1_{P_{\partial\Omega}}$ for the first order reduction of the Poisson equation (36–37) with the Dirichlet boundary condition (42). It is based on the variational formulation (9) and the use of the complex BNB theorem (Theorem 4 in Sect. 3.5) through Corollary 2.

Theorem 7 (Existence for the first order Poisson—Ω bounded) *Let* $d, m \in \mathbb{N}^*$, *let* $A_{Poisson}, K_{Poisson}$ *be* $m \times m$ *real matrices defined in* (37) *and boundary operator* $P_{\partial\Omega}^{Dirichlet}$ *defined in* (42). *Let* Ω *be a bounded open Lipschitz subset of* $\mathbb{R}^d$.

Then there exists a unique weak solution $U \in H^1_{P_{\partial\Omega}^{Dirichlet}}(\Omega, \mathbb{R})$ *to the variational problem* (9) *and*

$$\exists \gamma, \forall F \in \left(H^1(\Omega, \mathbb{R})'\right)^m, \quad \|U\|_{H^1_{P_{\partial\Omega}^{Dirichlet}}(\Omega,\mathbb{R})} \leq \gamma \|F\|_{(H^1(\Omega,\mathbb{R})')^m}. \tag{43}$$

Proof We check that the assumptions of Corollary 2 are satisfied to deduce the theorem. For the first condition, we need to prove that $iA_{Poisson}(\vec{\xi}) + K_{Poisson}$ is invertible. We compute the determinant using the cofactor expansion along the second column:

$$\det(iA_{Poisson}(\vec{\xi}) + K_{Poisson}) = -i\xi_1 \begin{vmatrix} i\xi_1 & 0 & 0 & 0 \\ i\xi_2 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ i\xi_d & 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & i\xi_2 & \dots & i\xi_d \\ i\xi_2 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ i\xi_d & 0 & 0 & 1 \end{vmatrix} \tag{44}$$

$$= -(i\xi_1)^2 + \begin{vmatrix} 0 & i\xi_2 & \dots & i\xi_d \\ i\xi_2 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ i\xi_d & 0 & 0 & 1 \end{vmatrix}. \tag{45}$$

By an induction argument we obtain

$$\det(iA(\vec{\xi}) + K) = -\sum_{i=1}^{d}(i\xi_i)^2 = \|\vec{\xi}\|_2^2. \tag{46}$$

Hence we obtain for a.e. $\vec{\xi} \in \mathbb{R}^d$, $\det(iA_{Poisson}(\vec{\xi}) + K_{Poisson}) \neq 0$, thus the first condition of Corollary 2 is met.

We now check the second condition (Eq. 29) using the definition of the Dirichlet boundary operator (42)

$$\begin{aligned}\forall\vec{\xi} \in \mathbb{R}^d,\ (\mathbb{I}_m - P_{\partial\Omega}^{Dirichlet})A_{Poisson}(\vec{\xi})(\mathbb{I}_m - P_{\partial\Omega}^{Dirichlet}) &= \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I}_d \end{pmatrix}\begin{pmatrix} 0 & {}^t\vec{\xi} \\ \vec{\xi} & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I}_d \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ \vec{\xi} & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & \mathbb{I}_d \end{pmatrix} \\ &= 0. \end{aligned} \tag{47}$$

Hence the second compatibility condition in Corollary 2 is true.

We now check the third condition (Eq. 30): $\forall\vec{\xi} \in \mathbb{R}^d$,

$$\ker K = span\langle(1, 0)\rangle \tag{48}$$

$$\ker P_{\partial\Omega}^{Dirichlet}(\vec{\xi}) = (\ker K)^{\perp} = span\langle(1, 0)\rangle^{\perp}, \tag{49}$$

hence $\ker K \cap \ker P_{\partial\Omega}^{Dirichlet}(\vec{\xi}) = \{0\}$ and the third condition in Corollary 2 is satisfied. □

5 Conclusion and perspectives

First-order systems with $K \not> 0$ are of particular interest in many applications. We studied the stationary first-order systems for $\Omega = \mathbb{R}^d$ and $\Omega \subset \mathbb{R}^d$ bounded using tools derived from harmonic and functional analysis. In the case $\Omega = \mathbb{R}^d$, Fourier analysis (Sect. 2) showed that the assumption $K > 0$ is not necessary but only sufficient for the existence and uniqueness of solutions. On bounded domains (Sect. 3), we therefore, looked for an alternative approach to the classical Friedrichs' theory that assumes $K > 0$. We used a complex analog of the more general *Banach–Nečas–Babuška* theorem (Sect. 3.5) to obtain the existence and uniqueness of a solution in a setting that encompasses both Friedrichs' systems and the first order reduction of the Poisson problem on $\Omega \subset \mathbb{R}^d$ (Theorem 5). Indeed as shown in Sect. 4, the assumption $K > 0$ is violated by first-order reduction of the Poisson problem, where the kernel of K

is not trivial. We however obtained an existence result (Theorem 7) thanks to our new approach. This paper lays the ground for the numerical analysis of stationary hyperbolic problems where we expect to use discrete analogs of our new approach to design convergent numerical approximations.

Acknowledgements All authors whose names appear on the submission: made substantial contributions to the conception of the work; drafted the work and revised it critically for important intellectual content; approved the version to be published; and agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. We also declare that this paper is not published and is not under consideration for publication elsewhere.

Data availability Data availability is not applicable to this article as no new data were created or analysed in this study.

References

1. Friedrichs, K.O.: Symmetric positive linear differential equations. Commun. Pure Appl. Math. **11**, 333–418 (1958)
2. Ern, A., Guermond, J.-L.: Theory and Practice of Finite Elements, vol. 159. Springer, New York (2004)
3. Bouchut, F., Ounaissa, H., Perthame, B.: Upwinding of the source term at interfaces for Euler equations with high friction. Comput. Math. Appl. **53**(3–4), 361–375 (2007)
4. Audusse, E., Minh, H.D., Omnes, P., Penel, Y.: Analysis of a modified Godunov scheme for the linear wave equation with Coriolis source term on cartesian meshes. J. Comput. Phys. **373**, 91–129 (2018)
5. Godlewski, E., Raviart, P.A.: Numerical Approximation of Hyperbolic Systems of Conservation Laws. Applied Mathematical Sciences, vol. 118. Springer, New York (1996)
6. Eymard, R., Gallouët, T., Herbin, R.: Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes sushi: a scheme using stabilization and hybrid interfaces. IMA J. Numer. Anal. **30**(4), 1009–1043 (2010)
7. Serre, D.: Systems of Conservation Laws 1: Hyperbolicity, Entropies, Shock Waves. Cambridge University Press, Cambridge (1999)
8. Bochev, P.B., Gunzburger, M.D.: Least-Squares Finite Element Methods. Springer, Berlin (2006)
9. Cosner, C.: On the definition of ellipticity for systems of partial differential equations. J. Math. Anal. Appl. **158**(1), 80–93 (1991)
10. Rus, I.A.: Maximum principles for elliptic systems and the problem of the minimum matrix norm of a characteristic matrix, revisited. Stud. Univ. Babes-Bolyai Math. **58**(2), 199–211 (2013)
11. Krupchyk, K., Seiler, W.M., Tuomela, J.: Overdetermined elliptic systems. Found. Comput. Math. **6**(3), 309–351 (2006)
12. Agmon, S., Douglis, A., Nirenberg, L.: Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II. Commun. Pure Appl. Math. **17**(1), 35–92 (1964)
13. Ern, A., Guermond, J.-L., Caplain, G.: An intrinsic criterion for the bijectivity of Hilbert operators related to Friedrich' systems. Commun. Partial Differ. Equ. **32**(2), 317–341 (2007)
14. Ern, A., Guermond, J.-L.: Discontinuous Galerkin methods for Friedrichs' systems. I. General theory. SIAM J. Numer. Anal. **44**(2), 753–778 (2006)
15. Antonić, N., Burazin, K., Crnjac, I., Erceg, M.: Complex Friedrichs systems and applications. J. Math. Phys. **58**(10), 101508 (2017)
16. Brezis, H.: Functional Analysis, Sobolev Spaces and PDEs. Universitext. Springer, Berlin (2010)
17. Rudin, W.: Functional Analysis. International Series in Pure and Applied Mathematics, vol. 8. McGraw-Hill Science/Engineering/Math., New York (1991)