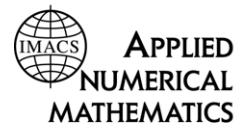




Available online at [www.sciencedirect.com](http://www.sciencedirect.com)



Applied Numerical Mathematics 53 (2005) 437–456



[www.elsevier.com/locate/apnum](http://www.elsevier.com/locate/apnum)

## Perturbation index of linear partial differential-algebraic equations <sup>☆</sup>

J. Rang, L. Angermann \*

*Institute of Mathematics, TU Clausthal, Erzstr. 1, D-38768 Clausthal-Zellerfeld, Germany*

Available online 13 October 2004

---

### Abstract

For numerous problems in science and engineering a refined modeling approach leads to initial-boundary value problems for partial differential-algebraic equations (PDAEs). The paper investigates linear PDAEs from the point of view of weakly differentiable in space solutions. The appropriate treatment of boundary conditions is obtained by the requirement that the spatial differential operator has to satisfy a Gårding-type inequality in suitable function spaces. Based on this, an index concept extending the classical perturbation index is introduced.

© 2004 IMACS. Published by Elsevier B.V. All rights reserved.

**Keywords:** Partial differential-algebraic equations; Coupled systems; Perturbation index

---

### 1. Introduction

Differential-algebraic equations (DAEs) consist of ordinary differential equations (ODEs) coupled with finite-dimensional algebraic equations. According to the survey [6], DAEs are singular implicit ODEs  $F(t, \mathbf{u}, \dot{\mathbf{u}}) = 0$ ,  $t \in J$ , where  $J = (0, \bar{t})$  is a time interval with  $\bar{t} \in (0, \infty]$  and the matrix  $\frac{\partial F(t, \mathbf{u}, \dot{\mathbf{u}})}{\partial \mathbf{v}}$  is singular everywhere in  $J$ . Otherwise the above system leads to an implicit ODE.

DAEs arise in various fields of applications, for example, in the simulation of electrical circuits, constrained dynamical systems (Euler–Lagrange equation and chemical reactions), reduced equations in singularly perturbed systems, and semi-discretization of partial differential equations (PDEs) with the vertical method of lines (MOL) (e.g., the Navier–Stokes equations [17], the semiconductor device

---

<sup>☆</sup> Extended version of a talk presented at NUMDIFF-10 (Halle, September 2003).

\* Corresponding author.

E-mail addresses: [major@math.tu-clausthal.de](mailto:major@math.tu-clausthal.de) (J. Rang), [mala@math.tu-clausthal.de](mailto:mala@math.tu-clausthal.de) (L. Angermann).

equations and density-dependent flows in porous media). Further examples are given in [1]. PDEs the semi-discretization in space of which leads to a DAE are often called partial differential-algebraic equations (PDAEs) and the semi-discretized system a MOL-DAE. For a historical overview we refer to [6] and for a detailed introduction into the theory of DAEs to [7,1].

The notion of the so-called index of a DAE plays a fundamental role in both theoretical and numerical investigations of such problems. It has turned out to give insight into the solution properties, as well as into the numerical difficulties to be expected when solving these problems, i.e., how to obtain consistent initial data if there are hidden constraints. To a certain extent, the DAE index is a measure of the singularity of the DAE. There are various types of indices known, for example, the differentiation index and the perturbation index to mention the best known indices. Griepentrog and März (see [7]) describe the theoretical background and define some indices, whereas Brenan, Campbell and Petzold in [1] present a theoretical overview and numerical aspects. The differentiation index is one of the main topics of that book. Campbell and Gear specify the definition of the differentiation index in [2]. A definition of the perturbation index and an overview about the numerical methods can be found in [9] and [10], respectively.

In contrast to the theory of DAEs, the theory of PDAEs is a comparatively young topic. As it is expected that PDAEs can also be characterized by an index (or some collection of indices), many efforts have been made to define indices for PDAEs. The first concepts deal with linear PDAEs with constant coefficients and in one spatial dimension, i.e., with problems of the form  $A\dot{\mathbf{u}} + B\mathbf{u}'' + C\mathbf{u} = \mathbf{f}$ , where “ $\cdot$ ” and “ $'$ ” denote differentiation with respect to the time  $t$  and the spatial variable  $x$ , respectively. An index for this problem can be determined by means of semi-discretization in space or time. The resulting problem is called a MOL-DAE (see, for example, [3]). An alternative concept defines indices with the help of Laplace- and Fourier-transform (see [13]). A definition of the perturbation index for the above PDAE is given in [12]. In [11], Lucht and Strehmel consider semi-linear problems of the form  $A\dot{\mathbf{u}} + B\mathbf{u}'' + F(\mathbf{u}) = \mathbf{f}$ . Many indices for the PDAE  $A\dot{\mathbf{u}} + B\mathbf{u}'' + C\mathbf{u}' + D\mathbf{u} = \mathbf{f}$  can be found in paper of Campbell and Marszalek [4]. The paper of Martinson and Barton [15] deals with a definition of the differentiation index for general PDEs and PDAEs. Hyperbolic-type problems are considered in the paper of Günther and Wagner [8]. The group of März consider the so-called tractability index for abstract DAEs (ADAEs), see [14].

Most of the above-mentioned papers on PDAEs deal with classes of solutions which are smooth in space. The present paper investigates linear PDAEs within the framework of weak solutions, i.e., we consider the PDAEs as abstract DAEs within suitable function spaces of Sobolev-type. The appropriate treatment of boundary conditions within this framework is motivated by the requirement that the spatial component of the differential operator has to satisfy a Gårding-type inequality. The latter is an important property the usefulness of which is known, for example, from the theory of abstract degenerate parabolic equations, see [5] or [16]. Based on the obtained weak formulation, an index concept extending the classical perturbation index is introduced.

## 2. The problem and its weak formulation

Let  $\Omega \subset \mathbb{R}^d$ ,  $d \in \{2, 3\}$  be a domain with a Lipschitzian boundary  $\partial\Omega$ . Let  $J = (0, \bar{t})$ ,  $\bar{t} \in (0, \infty]$ , be some time interval. We consider the following linear system of  $n \in \mathbb{N}$  partial differential, ordinary differential, and algebraic equations with respect to the unknown  $\mathbf{u} = (u_1, \dots, u_n)^\top : J \times \Omega \rightarrow \mathbb{R}^n$ :

$$A\dot{\mathbf{u}} + \mathcal{L}\mathbf{u} = \mathbf{f} \quad \text{in } J \times \Omega, \tag{1}$$

where  $n \in \mathbb{N}$ ,  $A : \Omega \rightarrow \mathbb{R}^{n,n}$ ,  $\mathbf{f} : \Omega \rightarrow \mathbb{R}^n$ ,

$$\begin{aligned} (\mathcal{L}\mathbf{u})_i &:= \sum_{j=1}^n \mathcal{L}_{ij} u_j, \quad i = 1, \dots, n, \\ \mathcal{L}_{ij} w &:= -\nabla \cdot (K_{ij} \nabla w) + b_{ij} \cdot \nabla w + c_{ij} w, \quad i, j = 1, \dots, n. \end{aligned}$$

Here the coefficients  $K_{ij} : \Omega \rightarrow \mathbb{R}^{d,d}$ ,  $b_{ij} : \Omega \rightarrow \mathbb{R}^d$ ,  $c_{ij} : \Omega \rightarrow \mathbb{R}$ ,  $i, j = 1, \dots, n$ , have the properties

$$A \in L_\infty(\Omega)^{n,n}, \quad \mathbf{f} \in H^s(\Omega)^n, \quad \text{for some } s \geq 0, \quad (2)$$

$$K_{ij} = K_{ij}^\top \in W_\infty^1(\Omega)^{d,d}, \quad b_{ij} \in L_\infty(\Omega)^d, \quad c_{ij} \in L_\infty(\Omega), \quad i, j = 1, \dots, n. \quad (3)$$

We assume that the matrix function  $A$  in (1) is singular in some subset of  $\Omega$  with a positive (Lebesgue) measure. In such a case, the system (1) is called a partial differential-algebraic equation (PDAE). An important ingredient of a systematic theory of PDAEs would be a classification scheme such that its output tells us, among other things, which initial *and* boundary conditions may be imposed. Unfortunately, to our knowledge the actual situation is that there is no such classification scheme.

On the other hand, it is easy to see that if we want to develop a theory of weak solutions, then a naive testing of the PDAE (1) by too arbitrary test functions can cause serious problems since the space of test functions typically depends on the type of boundary conditions.

We propose to formulate the boundary conditions formally as follows. Given  $3n$  piecewise continuous functions  $\mu_i, \omega_i : \partial\Omega \rightarrow \mathbb{R}$  and  $u_{\Gamma_i} : J \times \partial\Omega \rightarrow \mathbb{R}$ , the boundary conditions read as

$$\omega_i u_i + \mu_i \sum_{j=1}^n v \cdot (K_{ij} \nabla u_j) = u_{\Gamma_i} \quad \text{on } J \times \partial\Omega, \quad (4)$$

where  $v$  denotes the outer normal. First we set

$$\Gamma_{Ni} = \text{int}(\text{supp } \mu_i), \quad \Gamma_{Di} = \text{supp } \omega_i \setminus \Gamma_{Ni}, \quad \Gamma_{ij} = \Gamma_{Ni} \cup \text{supp } \omega_j, \quad (5)$$

$$\kappa_{ij} = \text{esssup}_{x \in \Omega} \|K_{ij}(x)\|_2, \quad \beta_{ij} = \text{esssup}_{x \in \Omega} \|b_{ij}(x)\|_2, \quad (6)$$

where  $\|\cdot\|_2$  denotes the Euclidean norm in  $\mathbb{R}^d$  or the corresponding matrix norm depending on the context and  $\text{int}(\cdot)$  is the set of points which are interior as elements of a subset of the boundary.

Moreover the data of (1) and (4) should fulfil the following minimal requirements:

- (C1) If  $\kappa_{ij} > 0$  for some  $i, j$ , i.e., there is some diffusion term in the  $i$ th differential equation, then boundary conditions for  $u_j$  should be prescribed on the whole boundary and so  $\Gamma_{ij} = \Gamma_{Ni} \cup \Gamma_{Dj} = \partial\Omega$  holds.
- (C2) If there is no diffusion term in the  $j$ th column, i.e.,  $\sum_{i=1}^n \kappa_{ij} = 0$ , but some convection terms in the  $j$ th column, i.e.,  $\sum_{i=1}^n \beta_{ij} > 0$ , then  $\emptyset \neq \text{supp } \omega_j \subset \bigcup_{i=1}^n \{x \in \partial\Omega : v(x) \cdot b_{ij}(x) < 0\}$ .
- (C3) If some Neumann part of the boundary is not empty, i.e.,  $\Gamma_{Ni} \neq \emptyset$  for some  $i$ , then one of the diffusion coefficients in the  $i$ th row is non-zero and we have  $\sum_{j=1}^n \kappa_{ij} > 0$ .
- (C4) If there are Dirichlet boundary conditions given, i.e.,  $\text{supp } \omega_j \neq \emptyset$  for some  $j$ , then one of the diffusion or convection coefficients of  $u_j$  is non-zero and so  $\sum_{i=1}^n (\kappa_{ij} + \beta_{ij}) > 0$ .
- (C5) If  $\text{supp } \omega_i \cup \text{supp } \mu_i = \emptyset$  for some  $i$ , then  $u_{\Gamma_i} = 0$ .

Furthermore, given some  $\mathbf{u}_0 : \Omega \rightarrow \mathbb{R}^n$ , we have the implicit initial condition

$$A(\mathbf{u} - \mathbf{u}_0) = 0 \quad \text{for } x \in \Omega. \quad (7)$$

In the following we will transform the PDAE (1) into a so-called functional DAE. We set

$$\begin{aligned} l_i &:= \begin{cases} 1, & \sum_{j=1}^n (\kappa_{ij} + \kappa_{ji} + \beta_{ji}) > 0, \\ 0, & \text{else} \end{cases} \quad \Upsilon := \{i \in \{1, \dots, n\} : l_i = 1\}, \\ V_i &:= \begin{cases} H_{\Gamma_{Di}}^1(\Omega) := \{v_i \in H^1(\Omega) : v_i|_{\Gamma_{Di}} = 0\}, & i \in \Upsilon, \\ H_{\Gamma_{Di}}^0(\Omega) := L_2(\Omega), & i \notin \Upsilon, \end{cases} \\ V &:= \bigotimes_{i=1}^n H_{\Gamma_{Di}}^{l_i}(\Omega) \subset Y := \bigotimes_{i=1}^n H^{l_i}(\Omega), \quad X := L_2(\Omega)^n. \end{aligned} \quad (8)$$

The norms on  $Y, X$  are defined in the usual way:

$$\|\mathbf{v}\|_X^2 := \sum_{i=1}^n \|v_i\|_{0,2,\Omega}^2, \quad \|\mathbf{v}\|_Y^2 := \sum_{i=1}^n \|v_i\|_{l_i,2,\Omega}^2,$$

$\mathbf{v} = (v_1, \dots, v_n)^\top \in Y$ , and the norm  $\|\cdot\|_Y$  is the restriction of  $\|\cdot\|_Y$  to  $V$ . For the convenience of the reader we summarize our notations and assumptions in the following corollary.

### Corollary 2.1.

- (i) If  $\sum_{i=1}^n (\kappa_{ij} + \beta_{ji}) = 0$  for some  $j$ , i.e.,  $u_j$  has neither diffusion nor convection coefficients in all columns, then Dirichlet boundary conditions w.r.t.  $u_j$  must not be prescribed and so  $\omega_j = 0$ .
- (ii) If the  $i$ th row has no diffusion terms, i.e.,  $\sum_{j=1}^n \kappa_{ij} = 0$ , then we need no Neumann boundary conditions and so  $\mu_i = 0$ .
- (iii) If  $l_i = 0$  for some  $i$ , then the  $i$ th row is at most a first-order PDE and  $u_i$  has no diffusion and convection coefficients, i.e.,  $\sum_{j=1}^n (\kappa_{ij} + \kappa_{ji} + \beta_{ji}) = 0$ ,  $\omega_i = 0$ ,  $\mu_i = 0$ , and  $u_{\Gamma i} = 0$ .
- (iv) If  $l_i = 1$  for some  $i$ , then  $\sum_{j=1}^n (\kappa_{ij} + \kappa_{ji} + \beta_{ji}) > 0$  and  $\bigcup_{j=1}^n (\Gamma_{ij} \cup \Gamma_{ji}) \neq \emptyset$ .
- (v) If  $\kappa_{ij} > 0$  for some  $i, j$ , then  $l_i = l_j = 1$  and  $\Gamma_{ij} = \partial\Omega$ .
- (vi) If  $\kappa_{ij} = 0$  and  $\beta_{ij} > 0$  for some  $i, j$ , i.e., there is only a convection coefficient of  $u_j$  in the  $i$ th row, then  $l_j = 1$  and so  $u_j$  belongs to  $H^1$  and  $\text{supp } \omega_j \neq \emptyset$ .
- (vii) Let  $\sum_{j=1}^n (\kappa_{ij} + \kappa_{ji} + \beta_{ji}) = 0$ . Then  $l_i = 0$ ,  $\mu_i = 0$ ,  $\omega_i = 0$ , and  $u_{\Gamma i} = 0$ .

**Proof.** (i) follows from the condition (C4).

(ii) follows from the condition (C3).

(iii) From (8) we get  $\sum_{j=1}^n \kappa_{ij} = 0$  and  $\sum_{j=1}^n (\kappa_{ji} + \beta_{ji}) = 0$ . With (i) we get  $\omega_i = 0$  and with (ii)  $\mu_i = 0$ . The condition on  $u_{\Gamma i}$  is then a consequence of (C5).

(iv) (8) implies  $\sum_{j=1}^n \kappa_{ij} > 0$ ,  $\sum_{j=1}^n \kappa_{ji} > 0$ , or  $\sum_{j=1}^n \beta_{ji} > 0$ . In the first case we get with condition (C1)  $\Gamma_{ij} = \partial\Omega$  and so  $\bigcup_{i=1}^n \Gamma_{ij} \neq \emptyset$ . The second case is equivalent to the first, if we change the role of  $i$  and  $j$ . Finally we get with condition (C2)  $\text{supp } \omega_i \neq \emptyset$ . Hence  $\Gamma_{ji} \neq \emptyset$ .

(v) follows with (8) and (C1).

(vi) follows with (8) and (C2).

(vii) follows with (8), (ii), (i), and (C5).  $\square$

The following examples illustrate these settings.

**Example 2.1.** Let  $\Gamma := \partial\Omega$  and consider the following PDAE

$$\begin{cases} \dot{u}_2 - \Delta u_1 + u_2 = f_1 & \text{in } J \times \Omega, \\ u_1 = f_2 & \text{in } J \times \Omega, \\ u_1(t, x) = f_2(t, x) & \text{on } J \times \Gamma, \\ u_2(0, x) = u_{20}(x) & x \in \Omega. \end{cases} \quad (9)$$

Writing this problem in the form (1), (4), we see that

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (K_{ij}) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad (b_{ij}) = 0, \quad (c_{ij}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\omega_1 = 1, \quad \omega_2 = 0, \quad \mu_1 = \mu_2 = 0 \quad \text{on } \Gamma, \quad u_{\Gamma 2} = 0.$$

Then this PDAE together with the boundary conditions fulfils the assumptions (C1)–(C5). In particular, we have  $l_1 = 1$ ,  $l_2 = 0$ ,  $\Upsilon = \{1\}$ ,  $V_1 = H_0^1(\Omega)$ , and  $V_2 = L_2(\Omega)$ . We also observe that there is no need in requiring an initial condition to  $u_1$ , and this is consistent with (7).

**Example 2.2.** Let  $\Gamma := \partial\Omega$  and consider the following PDAE

$$\begin{cases} -\Delta u_1 + u_2 = f_1 & \text{in } J \times \Omega, \\ \dot{u}_2 + a \cdot \nabla u_1 - \nabla \cdot (D \nabla u_2) + b \cdot \nabla u_2 + u_1 - u_2 = f_2 & \text{in } J \times \Omega, \\ \mathbf{u} = \mathbf{u}_D & \text{on } J \times \Gamma, \\ u_1(0, x) = u_{10}(x) & x \in \Omega, \end{cases} \quad (10)$$

where  $a, b \in \mathbb{R}^d$  and  $D \in \mathbb{R}^{d,d}$  are constant with  $\|a\|_2 = \|b\|_2 = \|D\|_2 = 1$ . Writing this problem in the form (1), (4), we see that

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (K_{ij}) = \begin{pmatrix} I & 0 \\ 0 & 1 \end{pmatrix}, \quad (b_{ij}) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad (c_{ij}) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix},$$

$$\omega_1 = \omega_2 = 1, \quad \mu_1 = \mu_2 = 0 \quad \text{on } \Gamma.$$

Then this PDAE together with the boundary conditions fulfils the assumptions (C1)–(C5). In particular, we have  $l_1 = l_2 = 1$ ,  $\Upsilon = \{1, 2\}$ , and  $V_1 = V_2 = H_0^1(\Omega)$ . We also observe that there is no need in requiring an initial condition to  $u_1$ , and this is consistent with (7).

To motivate the following definitions, we recall the following formula.

**Corollary 2.2.** Suppose  $\kappa_{ij} > 0$  for some  $i, j$  and

$$\mu_i^{-1}, \omega_i \in L_\infty(\Gamma_{Ni}), \quad u_{\Gamma i} \in L_2(\Gamma_{Ni}) \quad \text{for } i \in \{1, \dots, n\} \text{ such that } \Gamma_{Ni} \neq \emptyset. \quad (11)$$

Then, for  $u_j \in H^{l_j}(\Omega)$  and  $v_i \in V_i$ , it holds

$$\int_{\partial\Omega} \sum_{j=1}^n v \cdot (K_{ij} \nabla u_j) v_i \, ds = \int_{\Gamma_{Ni}} \sum_{j=1}^n v \cdot (K_{ij} \nabla u_j) v_i \, ds = \int_{\Gamma_{Ni}} \frac{1}{\mu_i} (u_{\Gamma i} - \omega_i u_i) v_i \, ds.$$

**Proof.** The first equality is a result of the definition of the space  $V_i$ , and the second equality sign follows from the definition of  $\Gamma_{Ni}$  and from the assumption (11).  $\square$

In order to give an abstract formulation of the problem (1), (4) and (7) under the assumptions (2), (3), and (11), first we introduce the following bilinear and linear forms, where  $(\cdot, \cdot)$  denotes the  $L_2(\Omega)$ - or  $L_2(\Omega)^d$ -inner product:

$$a^{(i,j)}(u_j, v_i) := (a_{ij} u_j, v_i), \quad u_j \in H^{l_j}(\Omega), \quad v_i \in V_i,$$

$$b^{(i,j)}(u_j, v_i) := (K_{ij} \nabla u_j, \nabla v_i) + (b_{ij} \cdot \nabla u_j + c_{ij} u_j, v_i) + \delta_{ij} \int_{\Gamma_{Ni}} \frac{\omega_i}{\mu_i} u_i v_i \, ds, \quad i, j \in \{1, \dots, n\},$$

$$f_\Omega^{(i)}(v_i) := (f_i, v_i), \quad i \in \{1, \dots, n\},$$

$$f_N^{(i)}(v_i) := \int_{\Gamma_{Ni}} \frac{1}{\mu_i} u_{\Gamma i} v_i \, ds, \quad i \in \{1, \dots, n\}, \quad \Gamma_{Ni} \neq \emptyset.$$

With the settings (2), (3), and  $\mathbf{f} \in H^s(\Omega)^n$  for some  $s \geq 0$ , we can define the operators  $\mathcal{A}, \mathcal{B}: Y \rightarrow V^*$  and  $\mathbf{f}_\Omega, \mathbf{f}_N \in V^*$  by means of the identities

$$\begin{aligned} \langle \mathcal{A}\mathbf{u}, \mathbf{v} \rangle &:= \sum_{i,j=1}^n a^{(i,j)}(u_j, v_i), \quad \langle \mathcal{B}\mathbf{u}, \mathbf{v} \rangle := \sum_{i,j=1}^n b^{(i,j)}(u_j, v_i), \quad \forall \mathbf{u} \in Y, \mathbf{v} \in V, \\ \langle \mathbf{f}_\Omega, \mathbf{v} \rangle &:= \sum_{i=1}^n f_\Omega^{(i)}(v_i), \quad \langle \mathbf{f}_N, \mathbf{v} \rangle := \sum_{i=1}^n f_N^{(i)}(v_i). \end{aligned}$$

With this, we get the following operator equation in  $V^*$  for the unknown element  $\mathbf{u} \in Y$ :

$$\mathcal{A}\dot{\mathbf{u}} + \mathcal{B}\mathbf{u} = \mathbf{f}_\Omega + \mathbf{f}_N. \quad (12)$$

Given some  $\mathbf{u}_0 \in Y$ , the initial condition reads as

$$\mathcal{A}(\mathbf{u} - \mathbf{u}_0) = 0. \quad (13)$$

Eq. (12) is called an abstract DAE (ADAE). In order to be able to include inhomogeneous Dirichlet boundary conditions, we assume that there exists some abstract function  $\mathbf{u}_D: J \rightarrow Y$  with  $u_{Dj} = u_{\Gamma j}$  on  $\Gamma_{Dj}$  for  $j = 1, \dots, n$ . Using the representation  $\mathbf{u} = \mathbf{u}_{\text{hom}} + \mathbf{u}_D$ , where  $\mathbf{u}_{\text{hom}} \in V$ , and introducing the linear form  $\mathbf{f}_D$  by

$$\langle \mathbf{f}_D, \mathbf{v} \rangle := - \sum_{i,j=1}^n a^{(i,j)}(\dot{u}_{Dj}, v_i) - \sum_{i,j=1}^n b^{(i,j)}(u_{Dj}, v_i), \quad \forall \mathbf{v} \in V,$$

and the functional  $\mathbf{f} \in V^*$  by  $\mathbf{f} := \mathbf{f}_\Omega + \mathbf{f}_D + \mathbf{f}_N$ , we get the following operator equation in  $V^*$  for the unknown element  $\mathbf{u}_{\text{hom}} \in V$ :

$$\mathcal{A}\dot{\mathbf{u}}_{\text{hom}} + \mathcal{B}\mathbf{u}_{\text{hom}} = \mathbf{f}. \quad (14)$$

If there are no Dirichlet boundary conditions at all, then we formally set  $\mathbf{u}_D = 0$ .

In what follows, we will present some sufficient conditions under which the operator  $\mathcal{B}$  satisfies a Gårding-type inequality, i.e., there exist two constants  $\lambda \geq 0, c > 0$  such that it holds:

$$\forall \mathbf{v} \in V: \langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle + \lambda \|\mathbf{v}\|_X^2 \geq c \|\mathbf{v}\|_V^2. \quad (15)$$

To formulate the conditions, we introduce the following notation:

$$\begin{aligned}\gamma_{ij} &:= \operatorname{esssup}_{x \in \Omega} |c_{ij}(x)|, \quad i, j \in \{1, \dots, n\}, \\ \tilde{\beta}_i &:= \sum_{j=1}^n \beta_{ji}, \quad i \in \Upsilon, \quad \hat{\beta}_i := \sum_{j \in \Upsilon} \beta_{ij}, \quad i \in \{1, \dots, n\}, \\ \check{\beta}_i &:= \sum_{\substack{j=1 \\ j \neq i}}^n \beta_{ji}, \quad i \in \Upsilon, \quad \check{\beta}_i := \sum_{\substack{j \in \Upsilon \\ j \neq i}} \beta_{ij}, \quad i \in \{1, \dots, n\}.\end{aligned}$$

**Lemma 2.1.** Let there exist constants  $\underline{\kappa}_{ii} > 0$ ,  $i \in \Upsilon$ , such that

$$\xi^\top K_{ii}(x)\xi \geq \underline{\kappa}_{ii} \|\xi\|_2^2, \quad \forall \xi \in \mathbb{R}^d, \quad \forall x \in \Omega,$$

holds,  $m := |\Upsilon|$ , and let the inequalities

$$\mu_i^{-1} \omega_i \geq 0 \quad \text{on } \Gamma_{Ni}, \quad i \in \{1, \dots, n\} \text{ such that } \Gamma_{Ni} \neq \emptyset,$$

be valid. If the matrix  $\kappa_{\varepsilon, \text{sym}}^1 := \frac{\kappa_\varepsilon^1 + (\kappa_\varepsilon^1)^\top}{2}$ , where

$$\kappa_\varepsilon^1 := \begin{pmatrix} \underline{\kappa}_{11} - \frac{\varepsilon}{2} \tilde{\beta}_1 & -\kappa_{12} & \dots & -\kappa_{1m} \\ -\kappa_{21} & \underline{\kappa}_{22} - \frac{\varepsilon}{2} \tilde{\beta}_2 & & \vdots \\ \vdots & \ddots & & \vdots \\ -\kappa_{m1} & \dots & \dots & \underline{\kappa}_{mm} - \frac{\varepsilon}{2} \tilde{\beta}_m \end{pmatrix} \in \mathbb{R}^{m,m},$$

is positive definite for at least one  $\varepsilon > 0$ , then the Gårding inequality (15) is valid for some positive constants  $c$  and  $\lambda$ , which may depend on  $\varepsilon$ .

**Proof.** For simplicity of the notation, we denote by  $\|\cdot\|$  the  $L_2(\Omega)$ - or  $L_2(\Omega)^d$ -norm. First, with (6), the Cauchy–Schwarz inequality, and the  $\varepsilon$ -inequality we have that

$$\begin{aligned}|(K_{ij} \nabla v_j, \nabla v_i)| &\leq \kappa_{ij} \|\nabla v_i\| \|\nabla v_j\|, \\ \sum_{i=1}^n \sum_{j \in \Upsilon} |(b_{ij} \cdot \nabla v_j, v_i)| &\leq \frac{\varepsilon}{2} \sum_{j \in \Upsilon} \underbrace{\left( \sum_{i=1}^n \beta_{ij} \right)}_{=: \tilde{\beta}_j} \|\nabla v_j\|^2 + \frac{1}{2\varepsilon} \sum_{i=1}^n \underbrace{\left( \sum_{j \in \Upsilon} \beta_{ij} \right)}_{=: \hat{\beta}_i} \|v_i\|^2 \\ &= \frac{\varepsilon}{2} \sum_{i \in \Upsilon} \tilde{\beta}_i \|\nabla v_i\|^2 + \frac{1}{2\varepsilon} \sum_{i=1}^n \hat{\beta}_i \|v_i\|^2,\end{aligned}$$

hold for any  $\varepsilon > 0$ . Using these inequalities, we obtain

$$\begin{aligned}\langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle &= \sum_{i \in \Upsilon} (K_{ii} \nabla v_i, \nabla v_i) + \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} (K_{ij} \nabla v_j, \nabla v_i) \\ &\quad + \sum_{i=1}^n \sum_{j \in \Upsilon} (b_{ij} \cdot \nabla v_j, v_i) + \sum_{i,j=1}^n (c_{ij} v_j, v_i) + \sum_{i=1}^n \left( \frac{\omega_i}{\mu_i}, v_i^2 \right)_{\Gamma_{Ni}}\end{aligned}$$

$$\begin{aligned} &\geq \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \tilde{\beta}_i \right) \|\nabla v_i\|^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \|\nabla v_i\| \|\nabla v_j\| - \sum_{i=1}^n \frac{\hat{\beta}_i}{2\varepsilon} \|v_i\|^2 - \sum_{i,j=1}^n \gamma_{ij} \|v_i\| \|v_j\| \\ &\geq \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \tilde{\beta}_i \right) \|\nabla v_i\|^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \|\nabla v_i\| \|\nabla v_j\| - \tilde{c}_\varepsilon \|\mathbf{v}\|_X^2, \end{aligned}$$

where  $\tilde{c}_\varepsilon$  is the  $\|\cdot\|_2$ -norm of the matrix  $\frac{1}{2\varepsilon} \text{diag}(\hat{\beta}_i) + (\gamma_{ij})$ . Since the matrix  $\kappa_{\varepsilon, \text{sym}}^1$  is positive definite, there exists a constant  $c_\varepsilon$  with

$$c_\varepsilon := \min_{\xi \in \mathbb{R}^m, \|\xi\|_2} \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \tilde{\beta}_i \right) \xi_i^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \xi_i \xi_j > 0,$$

and we obtain

$$\begin{aligned} \langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle &\geq \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \tilde{\beta}_i \right) \|\nabla v_i\|^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \|\nabla v_i\| \|\nabla v_j\| - \tilde{c}_\varepsilon \|\mathbf{v}\|_X^2 \\ &= \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \tilde{\beta}_i \right) \|\nabla v_i\|^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \frac{\kappa_{ij} + \kappa_{ji}}{2} \|\nabla v_i\| \|\nabla v_j\| - \tilde{c}_\varepsilon \|\mathbf{v}\|_X^2 \\ &\geq c_\varepsilon \sum_{i \in \Upsilon} \|\nabla v_i\|^2 - \tilde{c}_\varepsilon \|\mathbf{v}\|_X^2. \end{aligned}$$

Setting  $\lambda_\varepsilon := c_\varepsilon + \tilde{c}_\varepsilon$ , we obtain

$$\langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle + \lambda_\varepsilon \|\mathbf{v}\|_X^2 \geq c_\varepsilon \sum_{i \in \Upsilon} \|\nabla v_i\|^2 + (\lambda_\varepsilon - \tilde{c}_\varepsilon) \|\mathbf{v}\|_X^2 = c_\varepsilon \left( \sum_{i \in \Upsilon} \|\nabla v_i\|^2 + \|\mathbf{v}\|_X^2 \right) = c_\varepsilon \|\mathbf{v}\|_V^2. \quad \square$$

**Lemma 2.2.** Let there exist constants  $\underline{\kappa}_{ii} > 0$ ,  $i \in \Upsilon$ , such that

$$\xi^\top K_{ii}(x)\xi \geq \underline{\kappa}_{ii} \|\xi\|_2^2, \quad \forall \xi \in \mathbb{R}^d, \quad \forall x \in \Omega,$$

holds,  $m := |\Upsilon|$ ,  $b_{ij} \in W_\infty^1(\Omega)^d$ , and let the inequalities

$$2\mu_i^{-1}\omega_i + \nu \cdot b_{ii} \geq 0 \quad \text{on } \Gamma_{Ni}, \quad i \in \{1, \dots, n\} \text{ such that } \Gamma_{Ni} \neq \emptyset,$$

be valid. If the matrix  $\kappa_{\varepsilon, \text{sym}}^2 := \frac{\kappa_\varepsilon^2 + (\kappa_\varepsilon^2)^\top}{2}$ , where

$$\kappa_\varepsilon^2 := \begin{pmatrix} \underline{\kappa}_{11} - \frac{\varepsilon}{2} \check{\beta}_1 & -\kappa_{12} & \dots & -\kappa_{1m} \\ -\kappa_{21} & \underline{\kappa}_{22} - \frac{\varepsilon}{2} \check{\beta}_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ -\kappa_{m1} & \dots & \dots & \underline{\kappa}_{mm} - \frac{\varepsilon}{2} \check{\beta}_m \end{pmatrix} \in \mathbb{R}^{m,m},$$

is positive definite for at least one  $\varepsilon > 0$ , then the Gårding inequality (15) is valid for some positive constants  $c$  and  $\lambda$ , which may depend on  $\varepsilon$ .

**Proof.** First we note that

$$(b_{ii} \cdot \nabla v_i, v_i) = \frac{1}{2} (b_{ii}, \nabla(v_i)^2) = -\frac{1}{2} (\nabla \cdot b_{ii}, v_i^2) + \frac{1}{2} (v \cdot b_{ii}, v_i^2)_{\Gamma_{Ni}},$$

holds. Then we get

$$\begin{aligned} & (b_{ii} \cdot \nabla v_i + c_{ii} v_i, v_i) + \left( \frac{\omega_i}{\mu_i} v_i, v_i \right)_{\Gamma_{Ni}} \\ &= \left( c_{ii} - \frac{1}{2} \nabla \cdot b_{ii}, v_i^2 \right) + \left( \frac{\omega_i}{\mu_i} + \frac{1}{2} v \cdot b_{ii}, v_i^2 \right)_{\Gamma_{Ni}} \geq \underline{\gamma}_{ii} \|v_i\|^2, \end{aligned}$$

where

$$\underline{\gamma}_{ii} := \operatorname{essinf}_{x \in \Omega} \left( c_{ii} - \frac{1}{2} \nabla \cdot b_{ii} \right).$$

Moreover we have with the Cauchy–Schwarz inequality and the  $\varepsilon$ -inequality that

$$\begin{aligned} |(b_{ij} \cdot \nabla v_j, v_i)| &\leq \beta_{ij} \|\nabla v_j\| \|v_i\| \leq \frac{\beta_{ij}}{2} \left( \varepsilon \|\nabla v_j\|^2 + \frac{1}{\varepsilon} \|v_i\|^2 \right), \\ \sum_{i=1}^n \sum_{\substack{j \in \Upsilon \\ j \neq i}} |(b_{ij} \cdot \nabla v_j, v_i)| &\leq \frac{\varepsilon}{2} \sum_{j \in \Upsilon} \underbrace{\left( \sum_{\substack{i=1 \\ i \neq j}}^n \beta_{ij} \right)}_{=: \check{\beta}_j} \|\nabla v_j\|^2 + \frac{1}{2\varepsilon} \sum_{i=1}^n \underbrace{\left( \sum_{\substack{j \in \Upsilon \\ j \neq i}} \beta_{ij} \right)}_{=: \check{\beta}_i} \|v_i\|^2 \\ &= \frac{\varepsilon}{2} \sum_{i \in \Upsilon} \check{\beta}_i \|\nabla v_i\|^2 + \frac{1}{2\varepsilon} \sum_{i=1}^n \check{\beta}_i \|v_i\|^2, \end{aligned}$$

hold for  $\varepsilon > 0$ . With these inequalities we obtain

$$\begin{aligned} \langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle &= \sum_{i \in \Upsilon} [(K_{ii} \nabla v_i, \nabla v_i) + (b_{ii} \cdot \nabla v_i, v_i)] + \sum_{i=1}^n (c_{ii} v_i, v_i) + \sum_{i=1}^n \left( \frac{\omega_i}{\mu_i}, v_i^2 \right)_{\Gamma_{Ni}} \\ &\quad + \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} (K_{ij} \nabla v_j, \nabla v_i) + \sum_{i=1}^n \sum_{\substack{j \in \Upsilon \\ j \neq i}} (b_{ij} \cdot \nabla v_j, v_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^n (c_{ij} v_j, v_i) \\ &\geq \sum_{i \in \Upsilon} \underline{\kappa}_{ii} \|\nabla v_i\|^2 + \sum_{i=1}^n \underline{\gamma}_{ii} \|v_i\|^2 - \frac{\varepsilon}{2} \sum_{i \in \Upsilon} \check{\beta}_i \|\nabla v_i\|^2 \\ &\quad - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \|\nabla v_i\| \|\nabla v_j\| - \sum_{i=1}^n \frac{\check{\beta}_i}{2\varepsilon} \|v_i\|^2 - \sum_{\substack{i,j=1 \\ i \neq j}}^n \gamma_{ij} \|v_i\| \|v_j\| \\ &\geq \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \check{\beta}_i \right) \|\nabla v_i\|^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \|\nabla v_i\| \|\nabla v_j\| - \check{c}_\varepsilon \|\mathbf{v}\|_X^2, \end{aligned}$$

where  $\check{c}_\varepsilon$  is the  $\|\cdot\|_2$ -norm of the matrix

$$\frac{1}{2\varepsilon} \text{diag}(\check{\beta}_i) - \text{diag}(\underline{\gamma}_{ii}) - \text{diag}(\gamma_{ii}) + (\gamma_{ij}).$$

Since the matrix  $\kappa_{\varepsilon, \text{sym}}^2$  is positive definite, there exists a constant  $c_\varepsilon$  with

$$c_\varepsilon := \min_{\xi \in \mathbb{R}^m, \|\xi\|_2=1} \sum_{i \in \Upsilon} \left( \underline{\kappa}_{ii} - \frac{\varepsilon}{2} \check{\beta}_i \right) \xi_i^2 - \sum_{\substack{i,j \in \Upsilon \\ i \neq j}} \kappa_{ij} \xi_i \xi_j > 0,$$

and we obtain, as in the proof of Lemma 2.1,

$$\langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle \geq c_\varepsilon \sum_{i \in \Upsilon} \|\nabla v_i\|^2 - \tilde{c}_\varepsilon \|\mathbf{v}\|_X^2.$$

With  $\lambda_\varepsilon := c_\varepsilon + \tilde{c}_\varepsilon$ , this estimate implies the desired inequality.  $\square$

For the following we define the set

$$\widetilde{\Upsilon} := \left\{ i \in \{1, \dots, n\} : \sum_{j=1}^n (\kappa_{ij} + \kappa_{ji} + \beta_{ij} + \beta_{ji}) > 0 \right\} \supset \Upsilon.$$

**Lemma 2.3.** *Let there exist constants  $\underline{\kappa}_{ii} > 0$ ,  $i \in \Upsilon$ , such that*

$$\xi^\top K_{ii}(x)\xi \geq \underline{\kappa}_{ii} \|\xi\|_2^2, \quad \forall \xi \in \mathbb{R}^d, \quad \forall x \in \Omega.$$

*Assume that the conditions*

$$\mu_i^{-1} \omega_i \geq 0, \quad \text{on } \Gamma_{Ni}, \quad i \in \{1, \dots, n\} \text{ such that } \Gamma_{Ni} \neq \emptyset,$$

*are satisfied and that there exist non-negative numbers  $p_{ij}, q_{ij} \in \mathbb{R}$  such that, for  $i, j \in \widetilde{\Upsilon}$ , the inequality*

$$|(b_{ij} \cdot \nabla v_j, v_i)| \leq p_{ij} \|\nabla v_i\|_{0,2,\Omega} \|\nabla v_j\|_{0,2,\Omega} + q_{ij} \|v_i\|_{0,2,\Omega} \|v_j\|_{0,2,\Omega}$$

*is satisfied. Let the entries of the matrix  $\tilde{\beta}$  be given by*

$$\tilde{\beta}_{ij} = \begin{cases} \underline{\kappa}_{ii} - p_{ii}, & i = j, \\ -(\kappa_{ij} + p_{ij}), & i \neq j, \end{cases} \quad i, j \in \widetilde{\Upsilon},$$

*and let  $\tilde{\beta}_s := \frac{\tilde{\beta} + \tilde{\beta}^\top}{2}$  be positive definite, then the Gårding inequality (15) is satisfied for some positive  $c$  and  $\lambda$ .*

**Proof.** For  $i \in \Upsilon$  we get with  $(\frac{\omega_i}{\mu_i} v_i, v_i)_{\Gamma_{Ni}} \geq 0$

$$b^{(i,i)}(v_i, v_i) \geq (\underline{\kappa}_{ii} - p_{ii}) \|\nabla v_i\|^2 - (q_{ii} + \gamma_{ii}) \|v_i\|^2,$$

and for  $i, j \in \widetilde{\Upsilon}$

$$b^{(i,j)}(v_j, v_i) \leq (\kappa_{ij} + p_{ij}) \|\nabla v_i\| \|\nabla v_j\| + (q_{ij} + \gamma_{ij}) \|v_i\| \|v_j\|.$$

Moreover we have

$$\begin{aligned} \sum_{i \in \tilde{\Upsilon}} \sum_{j \in \tilde{\Upsilon} \setminus \{i\}} b^{(i,j)}(v_j, v_i) &\leq \sum_{i \in \tilde{\Upsilon}} \sum_{j \in \tilde{\Upsilon} \setminus \{i\}} (\kappa_{ij} + p_{ij}) \|\nabla v_i\| \|\nabla v_j\| + \sum_{i \in \tilde{\Upsilon}} \sum_{j \in \tilde{\Upsilon} \setminus \{i\}} (q_{ij} + \gamma_{ij}) \|v_i\| \|v_j\|, \\ \sum_{i \notin \tilde{\Upsilon}} \sum_{j \notin \tilde{\Upsilon} \cup \{i\}} b^{(i,j)}(v_j, v_i) &= \sum_{i \notin \tilde{\Upsilon}} \sum_{j \notin \tilde{\Upsilon} \cup \{i\}} (c_{ij} v_j, v_i) \leq \sum_{i \notin \tilde{\Upsilon}} \sum_{j \notin \tilde{\Upsilon} \cup \{i\}} \gamma_{ij} \|v_i\| \|v_j\|, \end{aligned}$$

since the cases  $\beta_{ij} > 0$  with  $i \in \tilde{\Upsilon}$  and  $j \notin (\tilde{\Upsilon} \cup \{i\})$  or  $i \notin \tilde{\Upsilon}$  and  $j \in (\tilde{\Upsilon} \setminus \{i\})$  are not possible (see definition of  $\tilde{\Upsilon}$ ). Using these inequalities and the fact that  $\beta_s$  is positive definite, i.e., there exists a constant  $c$  with

$$c := \min_{\xi \in \mathbb{R}^m, \|\xi\|_2=1} \sum_{i \in \tilde{\Upsilon}} (\kappa_{ii} - p_{ii}) \xi_i^2 - \sum_{i, j \in (\tilde{\Upsilon} \setminus \{i\})} (q_{ij} + \gamma_{ij}) \xi_i \xi_j > 0,$$

we obtain

$$\begin{aligned} \langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle &= \sum_{i \in \tilde{\Upsilon}} b^{(i,i)}(v_i, v_i) + \sum_{i \notin \tilde{\Upsilon}} b^{(i,i)}(v_i, v_i) \\ &\quad + \sum_{i \in \tilde{\Upsilon}} \sum_{j \in (\tilde{\Upsilon} \setminus \{i\})} b^{(i,j)}(v_j, v_i) + \underbrace{\sum_{i \in \tilde{\Upsilon}} \sum_{j \notin (\tilde{\Upsilon} \cup \{i\})} b^{(i,j)}(v_j, v_i)}_{=0} \\ &\quad + \underbrace{\sum_{i \notin \tilde{\Upsilon}} \sum_{j \in (\tilde{\Upsilon} \setminus \{i\})} b^{(i,j)}(v_j, v_i)}_{=0} + \sum_{i \notin \tilde{\Upsilon}} \sum_{j \notin (\tilde{\Upsilon} \cup \{i\})} b^{(i,j)}(v_j, v_i) \\ &\geq \sum_{i \in \tilde{\Upsilon}} [(\kappa_{ii} - p_{ii}) \|\nabla v_i\|^2 - (q_{ii} + \gamma_{ii}) \|v_i\|^2] - \sum_{i \notin \tilde{\Upsilon}} \gamma_{ii} \|v_i\|^2 \\ &\quad - \sum_{i \in \tilde{\Upsilon}} \sum_{j \in (\tilde{\Upsilon} \setminus \{i\})} [(\kappa_{ij} + p_{ij}) \|\nabla v_i\| \|\nabla v_j\| + (q_{ij} + \gamma_{ij}) \|v_i\| \|v_j\|] \\ &\quad - \sum_{i \notin \tilde{\Upsilon}} \sum_{j \notin (\tilde{\Upsilon} \setminus \{i\})} \gamma_{ij} \|v_i\| \|v_j\| \\ &= c \sum_{i \in \tilde{\Upsilon}} \|\nabla v_i\| \|\nabla v_i\| - \sum_{i, j=1}^n (q_{ij} + \gamma_{ij}) \|v_i\| \|v_j\|, \end{aligned}$$

where  $q_{ij} = 0$  for  $i, j \notin \tilde{\Upsilon}$ . If  $c_0$  denotes the  $\|\cdot\|_2$ -norm of the matrix  $(q_{ij} + \gamma_{ij})$ , it follows

$$\langle \mathcal{B}\mathbf{v}, \mathbf{v} \rangle + \lambda \|\mathbf{v}\|_X^2 \geq c \sum_{i \in \tilde{\Upsilon}} \|\nabla v_i\|^2 + (\lambda - c_0) \|\mathbf{v}\|_X^2.$$

With the setting  $\lambda := c + c_0$ , the lemma is proven.  $\square$

**Corollary 2.3.** *The following condition is necessary to satisfy the assumption of Lemmas 2.1–2.3: If  $\kappa_{ij} \neq 0$  and  $i \neq j$ , then  $\kappa_{ii} \neq 0$  and  $\kappa_{jj} \neq 0$ .*

**Proof.** The statement follows from fact that a matrix with a zero entry in the main diagonal cannot be positive definite.  $\square$

**Example 2.3.** We consider the PDAE (9) of Example 2.1. There we have shown that  $l_1 = 1$ ,  $l_2 = 0$ , and  $\Upsilon = \{1\}$ . So we have  $\kappa_\varepsilon^1 \in \mathbb{R}$ ,  $\tilde{\beta}_1 = \tilde{\beta}_2 = \hat{\beta}_1 = \hat{\beta}_2 = 0$ , and  $\kappa_\varepsilon^1 = 1$ . The Gårding-type inequality (15) is fulfilled since Lemma 2.1 hold with

$$c = 1, \quad \tilde{c} = \left\| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\|_2 = 1, \quad \lambda = c + \tilde{c} = 2.$$

**Example 2.4.** We consider the PDAE (10) of Example 2.2. There we have shown that  $l_1 = l_2 = 1$ ,  $l_3 = 0$ , and  $\Upsilon = \{1, 2\}$ . So we have  $\kappa_\varepsilon^1 \in \mathbb{R}^{2,2}$ ,  $\tilde{\beta}_1 = \|a\|_2 = 1$ ,  $\tilde{\beta}_2 = \|b\|_2 = 1$ ,  $\hat{\beta}_1 = 0$ ,  $\hat{\beta}_2 = \|a\|_2 + \|b\|_2 = 2$ , and

$$\kappa_\varepsilon^1 = \begin{pmatrix} 1 - \frac{\varepsilon}{2}\|b\|_2 & 0 \\ 0 & \|D\|_2 - \frac{\varepsilon}{2}\|b\|_2 \end{pmatrix} = \begin{pmatrix} 1 - \frac{\varepsilon}{2} & 0 \\ 0 & 1 - \frac{\varepsilon}{2} \end{pmatrix}.$$

With  $\varepsilon := 1$  the Gårding-type inequality (15) is fulfilled since Lemma 2.1 holds with  $c = \frac{1}{2}$ ,  $\tilde{c} = 1 + \sqrt{2}$ , and  $\lambda = c + \tilde{c} = 2 + \sqrt{2}$ .

### Remark.

- The above three lemmas cover different situations. For instance, there are situations where Lemma 2.3 is not applicable but Lemmas 2.1, 2.2 are. As an example, consider the PDAE

$$\begin{cases} \dot{u}_1 - \Delta u_1 + u'_1 + u_2 = f_1 & \text{in } J \times (0, 1), \\ -\Delta u_2 + u_1 - u_2 = f_2 & \text{in } (0, 1), \end{cases}$$

with Dirichlet boundary conditions for  $u_1, u_2$  and an initial condition for  $u_1$ . Then we have

$$k_\varepsilon^1 = \begin{pmatrix} 1 - \frac{\varepsilon}{2} & 0 \\ 0 & 1 \end{pmatrix}.$$

This matrix is positive definite for all  $\varepsilon \in (0, 2)$ . Hence Lemma 2.1 can be applied. Lemma 2.2 can be applied, too, since  $k_\varepsilon^2 = I$ . However, Lemma 2.3 cannot be applied since

$$(u'_1, u_1) \leq \|u'_1\| \|u_1\| \leq \|u'_1\|^2,$$

where the last estimate is a consequence of the Poincaré inequality, and

$$\tilde{\beta} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

- Comparing Lemma 2.1 with Lemma 2.2 we see that in Lemma 2.2 we require more smoothness of the coefficients  $b_{ij}$  than in Lemma 2.1 and different inequalities to be hold on  $\Gamma_{Ni}$ .
- In the cases where all three lemmas are applicable, different values of the constants  $c$  and  $\lambda$  result.

### 3. The perturbation index

In this section we introduce an extension of the perturbation index, which is known from the theory of DAEs, to the case of ADAEs and PDAEs. In particular, it turns out that the introduced perturbation index coincides with the classical notion in the case of DAEs. We start with the following definition.

**Definition 3.1.** We call  $\mathbf{u} \in C([0, t), V) \cap C^1(J, V)$  a weak solution of (12), (13) if  $\mathbf{u}$  satisfies (12) pointwise for all  $t \in J$  and fulfills the initial condition (13).

Concerning the problem of existence and uniqueness of weak solutions, we refer to the literature, e.g., [5,16].

We take a weak solution  $\mathbf{u}$  of the ADAE (12) which is consistent with the initial value  $\mathbf{u}_0 \in Y$ , i.e.,  $\mathcal{A}(\mathbf{u} - \mathbf{u}_0) = 0$ .

The notion of the perturbation index is based on the investigation of the sensitivity of this solution with respect to initial values, boundary values and right sides.

Starting from the ADAE written in the form (14), we introduce perturbations  $\delta_\Omega$  (of the right side),  $\delta_N$  (of the Neumann-type boundary conditions), and  $\delta_D$  (of the Dirichlet-type boundary) all acting from  $J$  to  $V^*$  and look for a solution  $\hat{\mathbf{u}}_{\text{hom}} \in V$  of the equation

$$\mathcal{A}\dot{\hat{\mathbf{u}}}_{\text{hom}} + \mathcal{B}\hat{\mathbf{u}}_{\text{hom}} = \mathbf{f}_\Omega + \mathbf{f}_D + \mathbf{f}_N + \delta_\Omega + \delta_D + \delta_N. \quad (16)$$

Subtracting (14) from (16) leads to the so-called *homogenized error equation* with respect to  $\boldsymbol{\epsilon}_{\text{hom}} \in V$

$$\mathcal{A}\dot{\boldsymbol{\epsilon}}_{\text{hom}} + \mathcal{B}\boldsymbol{\epsilon}_{\text{hom}} = \delta_\Omega + \delta_D + \delta_N =: \boldsymbol{\delta}. \quad (17)$$

Now we can define the perturbation index of an ADAE.

**Definition 3.2.** Let  $\mathcal{F}$  be a family of right sides such that, for any  $\mathbf{f} \in \mathcal{F}$ , the ADAE (12) has only one weak solution. Then the ADAE (12) has the perturbation index  $i_p$  along the solution  $\mathbf{u}$  on  $J$ , if  $i_p$  is the smallest integer such that, for all  $\hat{\mathbf{u}}$  having defects  $\delta_\Omega, \delta_D, \delta_N : J \rightarrow V^*$ , i.e.,

$$\mathcal{A}\dot{\hat{\mathbf{u}}}_{\text{hom}} + \mathcal{B}\hat{\mathbf{u}}_{\text{hom}} = \mathbf{f}_\Omega + \mathbf{f}_D + \mathbf{f}_N + \delta_\Omega + \delta_D + \delta_N,$$

there is on  $J$  an estimate of the form

$$\|\hat{\mathbf{u}}_{\text{hom}}(t) - \mathbf{u}_{\text{hom}}(t)\|_V \leq C \left( \|\hat{\mathbf{u}}_{\text{hom}}(0) - \mathbf{u}_{\text{hom}}(0)\|_V + \sum_{j=0}^{i_p-1} \sup_{\tau \in J} \left\| \frac{\partial^j \boldsymbol{\delta}(\tau)}{(\partial \tau)^j} \right\|_{V^*} \right). \quad (18)$$

### Remark.

- (i) In the definition it is implicitly assumed that Eq. (16) is solvable in  $J$  for the perturbations  $\boldsymbol{\delta}$  under consideration.
- (ii) Recall that the norms of the spaces  $V_i$ ,  $V$ ,  $V_i^*$  and  $V^*$  are defined as

$$\|v_i\|_{V_i}^2 := \|v_i\|_{l_i, 2, \Omega}^2 := \sum_{|\alpha| \leq l_i} \|\partial^\alpha v_i\|_{0, 2, \Omega}^2, \quad v_i \in V_i, \quad (19)$$

$$\|\mathbf{v}\|_V^2 := \sum_{i=1}^n \|v_i\|_{V_i}^2 := \sum_{i=1}^n \sum_{|\alpha| \leq l_i} \|\partial^\alpha v_i\|_{0, 2, \Omega}^2, \quad \mathbf{v} \in V, \quad (20)$$

$$\|\delta_i\|_{V_i^*} := \sup_{v_i \in V_i \setminus \{0\}} \frac{|\langle \delta_i, v_i \rangle|}{\|v_i\|_{V_i}}, \quad \delta_i \in V_i^*,$$

$$\|\boldsymbol{\delta}\|_{V^*} := \sum_{i=1}^n \|\delta_i\|_{V_i^*}, \quad \boldsymbol{\delta} \in V^*. \quad (21)$$

In the following we transform the perturbation index of an ADAE into a perturbation index of a PDAE. In [12] we find a perturbation index for the one-dimensional linear PDAE

$$A\dot{u} + Bu'' + Cu = f,$$

with respect to the  $L_\infty$ -norm in time and to the  $L_2$ -norm in space. Including an additive convection term, Campbell and Marszalek define in [4] a similar index which was generalized for so-called “hyperbolic” problems in [8].

In analogy to the above treatment of Dirichlet boundary data we assume that there exists an element  $\boldsymbol{\varepsilon}_D : J \times \Omega \rightarrow \mathbb{R}^n$  such that

(i) the perturbations  $\boldsymbol{\varepsilon}_{\Gamma i}$  of the Dirichlet data on  $\Gamma_i$  are the restrictions of  $\boldsymbol{\varepsilon}_{Di}$  to  $\Gamma_{Di}$ :

$$\boldsymbol{\varepsilon}_{\Gamma i} = \boldsymbol{\varepsilon}_{Di} \quad \text{on } \Gamma_{Di}, \quad i = 1, \dots, n,$$

(ii)  $\dot{\boldsymbol{\varepsilon}}_D : J \rightarrow Y$  exists,

(iii) the following extension properties hold:

$$\|\boldsymbol{\varepsilon}_D\|_V \leq C_1 \|\boldsymbol{\varepsilon}_\Gamma\|_{0,2,\Gamma_D}, \quad \|\dot{\boldsymbol{\varepsilon}}_D\|_{0,2,\Omega} \leq C_2 \|\dot{\boldsymbol{\varepsilon}}_\Gamma\|_{0,2,\Gamma_D},$$

where  $C_1, C_2 > 0$  do not depend on  $\boldsymbol{\varepsilon}_\Gamma$ .

Then  $\delta_D = -\mathcal{A}\dot{\boldsymbol{\varepsilon}}_D - \mathcal{B}\boldsymbol{\varepsilon}_D$ .

The perturbation  $\delta_N$  is given by means of the perturbations  $\boldsymbol{\varepsilon}_{\Gamma i} \in L_2(\Gamma_{Ni})$  of the Neumann data on  $\Gamma_{Ni}$ , i.e.,

$$\delta_{Ni}(v_i) = \int_{\Gamma_{Ni}} \frac{1}{\mu_i} \boldsymbol{\varepsilon}_{\Gamma i} v_i \, ds, \quad v_i \in V_i, \quad i = 1, \dots, n.$$

Since  $\delta_{\Omega i}$ ,  $i = 1, \dots, n$ , can be written in the form

$$\delta_{\Omega i}(v_i) = \sum_{|\alpha| \leq l_i} \int_{\Omega} \partial^\alpha \phi_{\Omega i}(t, x) \partial^\alpha v_i(x) \, dx, \quad v_i \in V_i,$$

with certain  $\phi_{\Omega i}(t, \cdot) \in V_i$  and since the inequalities

$$\begin{aligned} |a^{(i,j)}(\dot{\boldsymbol{\varepsilon}}_{Dj}, v_i)| &\leq \|a_{ij}\|_{0,\infty,\Omega} \|\dot{\boldsymbol{\varepsilon}}_{Dj}\|_{0,2,\Omega} \|v_i\|_{0,2,\Omega}, \\ |b^{(i,j)}(\boldsymbol{\varepsilon}_{Dj}, v_i)| &\leq \kappa_{ij} |\boldsymbol{\varepsilon}_{Dj}|_{1,2,\Omega} |v_i|_{1,2,\Omega} + \beta_{ij} |\boldsymbol{\varepsilon}_{Dj}|_{1,2,\Omega} \|v_i\|_{0,2,\Omega} \\ &\quad + \gamma_{ij} \|\boldsymbol{\varepsilon}_{Dj}\|_{0,2,\Omega} \|v_i\|_{0,2,\Omega} + \delta_{ij} C \|\boldsymbol{\varepsilon}_{Di}\|_{0,2,\Gamma_{Ni}} \|v_i\|_{0,2,\Gamma_{Ni}}, \end{aligned}$$

hold, we get the following estimates:

$$\begin{aligned} |\delta_{\Omega i}(v_i)| &\leq \|\phi_{\Omega i}\|_{V_i} \|v_i\|_{V_i}, \\ |\delta_{Ni}(v_i)| &= \left| \int_{\Gamma_{Ni}} \frac{1}{\mu_i} \boldsymbol{\varepsilon}_{\Gamma i} v_i \, ds \right| \leq C \|\boldsymbol{\varepsilon}_{\Gamma i}\|_{0,2,\Gamma_{Ni}} \|v_i\|_{0,2,\Gamma_{Ni}}, \\ |\delta_{Di}(v_i)| &= \left| - \sum_{j=1}^n [a^{(i,j)}(\dot{\boldsymbol{\varepsilon}}_{Dj}, v_i) - b^{(i,j)}(\boldsymbol{\varepsilon}_{Dj}, v_i)] \right| \leq C \sum_{j=1}^n (\|\dot{\boldsymbol{\varepsilon}}_{Dj}\|_{0,2,\Omega} + \|\boldsymbol{\varepsilon}_{Dj}\|_{V_j}) \|v_i\|_{V_i} \\ &= C (\|\dot{\boldsymbol{\varepsilon}}_D\|_{0,2,\Omega} + \|\boldsymbol{\varepsilon}_D\|_V) \|v_i\|_{V_i}. \end{aligned}$$

It follows

$$\begin{aligned}\|\delta_{\Omega i}\|_{V_i^*} &\leq \|\phi_{\Omega i}\|_{V_i}, & \|\delta_{N i}\|_{V_i^*} &\leq C\|\varepsilon_{\Gamma i}\|_{0,2,\Gamma_{Ni}}, \\ \|\delta_{D i}\|_{V_i^*} &\leq C(\|\dot{\varepsilon}_{\Gamma}\|_{0,2,\Gamma_D} + \|\varepsilon_{\Gamma}\|_{0,2,\Gamma_D})\end{aligned}$$

holds. Hence, by (21), we have

$$\|\delta\|_{V^*} \leq C \sum_{i=1}^n [\|\phi_{\Omega i}\|_{V_i} + \|\varepsilon_{\Gamma i}\|_{0,2,\partial\Omega} + \|\dot{\varepsilon}_{\Gamma i}\|_{0,2,\Gamma_{Di}}],$$

and differentiating  $j$ -times with respect to  $t$ , leads to

$$\|\partial_t^j \delta\|_{V^*} \leq C \sum_{i=1}^n \left[ \sum_{|\alpha| \leq l_i} \|\partial_t^j \partial^\alpha \phi_{\Omega i}\|_{0,2,\Omega} + \|\partial_t^j \varepsilon_{\Gamma i}\|_{0,2,\partial\Omega} + \|\partial_t^{j+1} \varepsilon_{\Gamma i}\|_{0,2,\Gamma_{Di}} \right].$$

Inserting into (18) leads to

$$\begin{aligned}\|\hat{\mathbf{u}}_{\text{hom}}(t) - \mathbf{u}_{\text{hom}}(t)\|_V &\leq C \sum_{i=1}^n \left\{ \sum_{|\alpha| \leq l_i} \|\partial^\alpha (\hat{\mathbf{u}}_{\text{hom}}(0) - \mathbf{u}_{\text{hom}}(0))\|_{0,2,\Omega} + \sum_{j=0}^{i_p-1} \sum_{|\alpha| \leq l_i} \sup_{\tau \in J} \|\partial_t^j \partial^\alpha \phi_{\Omega i}\|_{0,2,\Omega} \right. \\ &\quad \left. + \sum_{j=0}^{i_p-1} [\|\partial_t^j \varepsilon_{\Gamma i}\|_{0,2,\partial\Omega} + \|\partial_t^{j+1} \varepsilon_{\Gamma i}\|_{0,2,\Gamma_{Di}}] \right\}.\end{aligned}$$

To motivate a perturbation index for PDAEs we consider two examples.

**Example 3.1.** Let  $\Omega = (0, 1)$ ,  $u - u_D \in V := H_0^1(\Omega)$ , and

$$\begin{cases} \Delta u = f & \text{in } \Omega, \\ u(0) = g_1, & u(1) = g_2. \end{cases}$$

Then  $u_D := g_2x + g_1(1-x)$  and the solution is given by

$$u(x) = F(x) + (g_2 - g_1 - F(1))x + g_1 = F(x) - F(1)x + u_D,$$

where

$$F(x) := \int_0^x \int_0^y f(z) dz dy.$$

Since  $u_{\text{hom}} = F(x) - F(1)x$ , we have

$$u'_{\text{hom}} = F'(x) - F(1) = \int_0^x f(y) dy - \int_0^1 \int_0^x f(y) dy dx.$$

Using the triangle inequality and Hölder inequality, we get the estimate

$$|u'_{\text{hom}}| \leq 2\|f\|_{0,1,\Omega} \leq 2\|f\|_{0,2,\Omega}.$$

It follows

$$\|u'_{\text{hom}}\|_{0,2,\Omega} \leq 2\|f\|_{0,2,\Omega}$$

and with the Poincaré inequality we obtain

$$\|u_{\text{hom}}\|_{0,2,\Omega} \leq \|u'_{\text{hom}}\|_{0,2,\Omega} \leq 2\|f\|_{0,2,\Omega}.$$

In summary we have  $\|u_{\text{hom}}\|_{0,2,\Omega} \leq 2\|f\|_{0,2,\Omega}$  and  $\|u_{\text{hom}}\|_V \leq 2\sqrt{2}\|f\|_{0,2,\Omega}$ . For  $u_D$  we get after simple calculations

$$\begin{aligned} \|u_D\|_{0,2,\Omega}^2 &= \int_0^1 [g_2 x + g_1(1-x)]^2 dx \leq \frac{1}{3}(|g_1| + |g_2|)^2, \\ \|u_D\|_{0,2,\Omega} &= \frac{\sqrt{3}}{3}[|g_1| + |g_2|], \quad \|u'_D\|_{0,2,\Omega} = |g_2 - g_1|. \end{aligned}$$

Finally we have the estimates

$$\|u\|_{0,2,\Omega} \leq 2(\|f\|_{0,2,\Omega} + |g_1| + |g_2|), \quad \|u\|_Y \leq 2\sqrt{2}(\|f\|_{0,2,\Omega} + |g_1| + |g_2|).$$

Let us now consider the weak formulation of the homogeneous problem

$$-(u'_{\text{hom}}, v') = (f, v) + (u'_D, v'), \quad \forall v \in V.$$

Using the Poincaré inequality to estimate the last bound in

$$\begin{aligned} \|u'_{\text{hom}}\|_{0,2,\Omega}^2 &= (u'_{\text{hom}}, u'_{\text{hom}}) = (-f, u'_{\text{hom}}) - (u'_D, u'_{\text{hom}}) \\ &\leq [\|f\|_{0,2,\Omega} + |g_2 - g_1|] \|u'_{\text{hom}}\|_{0,2,\Omega}, \end{aligned}$$

we get the estimates

$$\|u'_{\text{hom}}\|_{0,2,\Omega} \leq \|f\|_{0,2,\Omega} + |g_2 - g_1|, \quad \|u_{\text{hom}}\|_V \leq \sqrt{2}[\|f\|_{0,2,\Omega} + |g_2 - g_1|].$$

A simple calculation shows that the strong solution  $u$  solves the weak problem, too.

**Example 3.2.** We consider the problem of Example 2.3:

$$\begin{cases} \dot{u}_2 - \Delta u_1 = f_1 & \text{in } J \times \Omega, \\ u_1 = f_2 & \text{in } J \times \Omega, \\ u_2(0, x) = u_{20}(x) & x \in \Omega. \end{cases}$$

The exact solution is given by  $u_1 = f_2$ ,  $u_2 = L_t(f_1 + \Delta f_2) + u_{20}$ , where (see [4])

$$L_t f(t) := \int_0^t f(\tau) d\tau.$$

Obviously, we need no boundary conditions to obtain a unique solution. In order to fit the PDAE with the conditions (C1)–(C5) we add formally the Neumann boundary condition

$$\nu \cdot \nabla u_1 = u_{\Gamma 1} := \nu \cdot \nabla f_1.$$

Since  $\mathbf{u} = \mathbf{u}_{\text{hom}}$  we get only estimates for  $\mathbf{u}$ :

$$\begin{aligned}\|u_1\|_{0,2,\Omega} &\leq \|f_2\|_{0,2,\Omega}, \\ \|u_2\|_{0,2,\Omega} &\leq C \left\{ \sup_{t \in J} t \|f_1(t, x)\|_{0,2,\Omega} + \sup_{t \in J} t \|\Delta f_2(t, x)\|_{0,2,\Omega} + \|u_{20}\|_{0,2,\Omega} \right\}.\end{aligned}$$

According to the discussion in Example 2.4, we consider for  $u_2 \in L_2(\Omega)$  and  $u_1 \in H^1(\Omega)$  the problem

$$(\dot{u}_2, v_1) - (\nabla u_1, \nabla v_1) = (f_1, v_1), \quad (u_1, v_2) = (f_2, v_2),$$

where  $v_2 \in L_2(\Omega)$  and  $v_1 \in H^1(\Omega)$  are arbitrary. The second equation imply that  $u_1 = f_2$  in the  $L_2(\Omega)$ -sense. Provided that  $u_1$  is sufficiently smooth, we substitute the expression for  $u_1$  into the first equation. Furthermore, if partial integration is possible, we obtain

$$(\dot{u}_2 - \Delta f_2, v_1) = (f_1, v_1),$$

i.e.,

$$\dot{u}_2 = f_1 + \Delta f_2$$

in the  $L_2(\Omega)$ -sense. Finally the weak solution is equal to the strong solution.

These examples motivate the following definition.

**Definition 3.3.** Let  $\mathbf{f}: J \times \Omega \rightarrow \mathbb{R}^n$  be a sufficiently smooth function. Let us assume that the PDAE (1) has only one weak solution. Then PDAE (1) has the perturbation index triple  $[i_{X,p,t}, i_{X,p,\Omega}, i_{X,p,B}]$  and  $[i_{V,p,t}, i_{V,p,\Omega}, i_{V,p,B}]$ , respectively along a solution  $\mathbf{u}$  on  $J$ , if  $i_{X,p,t}, i_{X,p,\Omega}, i_{X,p,B}$  and  $i_{V,p,t}, i_{V,p,\Omega}, i_{V,p,B}$ , respectively, are all the smallest integer such that, for all  $\hat{\mathbf{u}}$  having a defect  $\delta: J \rightarrow V^*$ , i.e.,

$$A\dot{\hat{\mathbf{u}}} + \mathcal{L}\hat{\mathbf{u}} = \mathbf{f} + \delta,$$

there is on  $J \times \Omega$  an  $L_2(\Omega)$ -estimate of the form

$$\begin{aligned}\|\hat{\mathbf{u}}(t) - \mathbf{u}(t)\|_{0,2,\Omega} &\leq C \left[ \sum_{|\alpha| \leq i_{X,p,\Omega}} \|\partial^\alpha(\hat{\mathbf{u}}_0 - \mathbf{u}_0)\|_{0,2,\Omega} + \sum_{i=1}^n \sum_{j=0}^{i_{X,p,t}-1} \sum_{|\alpha| \leq i_{X,p,\Omega}} \sup_{\tau \in J} \|\partial_t^j \partial^\alpha \phi_{\Omega i}\|_{0,2,\Omega} \right. \\ &\quad \left. + \sum_{i=1}^n \sum_{j=0}^{i_{X,p,t}-1} \sum_{|\alpha| \leq i_{X,p,B}} \sup_{\tau \in J} \|\partial_t^j \partial^\alpha (\hat{u}_{\Gamma i} - u_{\Gamma i})\|_{0,2,\partial\Omega} \right],\end{aligned}$$

and a  $V$ -estimate of the form

$$\begin{aligned}\|\hat{\mathbf{u}}(t) - \mathbf{u}(t)\|_Y &\leq C \left[ \sum_{|\alpha| \leq i_{V,p,\Omega}} \|\partial_x^\alpha(\hat{\mathbf{u}}_0 - \mathbf{u}_0)\|_{0,2,\Omega} + \sum_{i=1}^n \sum_{j=0}^{i_{V,p,t}-1} \sum_{|\alpha| \leq i_{V,p,\Omega}} \sup_{\tau \in J} \|\partial_t^j \partial^\alpha \phi_{\Omega i}\|_{0,2,\Omega} \right. \\ &\quad \left. + \sum_{i=1}^n \sum_{j=0}^{i_{V,p,t}-1} \sum_{|\alpha| \leq i_{V,p,B}} \sup_{\tau \in J} \|\partial_t^j \partial^\alpha (\hat{u}_{\Gamma i} - u_{\Gamma i})\|_{0,2,\partial\Omega} \right].\end{aligned}$$

**Corollary 3.4.** It follows from Definitions 3.2 and 3.3 that  $i_{p,t} = i_p$ .

## 4. Examples

Next we consider some examples which show that our definitions make sense. In the following we give only the boundary and initial conditions which are needed to make the solution unique. The remaining conditions are taken from the exact solution. In the following we assume that integration with respect to  $t$  is possible.

**Example 4.1.** The PDE of Example 3.1 has the perturbation index triple  $[i_{X,p,t}, i_{X,p,\Omega}, i_{X,p,B}] = [1, 0, 0]$  and  $[i_{V,p,t}, i_{V,p,\Omega}, i_{V,p,B}] = [1, 0, 0]$ , respectively.

**Example 4.2.** The PDE of Example 3.2 has the perturbation index triple  $[i_{X,p,t}, i_{X,p,\Omega}, i_{X,p,B}] = [1, 2, 0]$  and  $[i_{V,p,t}, i_{V,p,\Omega}, i_{V,p,B}] = [1, 3, 0]$ , respectively.

**Example 4.3.** Consider the PDAE of Example 2.4

$$\begin{cases} -\Delta u_1 + u_2 = f_1 & \text{in } J \times \Omega, \\ \dot{u}_2 + a \cdot \nabla u_1 - \nabla \cdot (D \nabla u_2) + b \cdot \nabla u_2 + u_1 - u_2 = f_2 & \text{in } J \times \Omega, \\ \mathbf{u} = \mathbf{u}_D & \text{on } J \times \Gamma, \\ u_1(0, x) = u_{10}(x) & x \in \Omega, \end{cases}$$

where  $a, b \in \mathbb{R}^d$  and  $D \in \mathbb{R}^{d,d}$  are constant with  $\|a\|_2 = \|b\|_2 = \|D\|_2 = 1$ . First we apply the Gårding-type inequality and the  $\varepsilon$ -inequality with  $\varepsilon := \frac{c}{2}$

$$(\dot{w}_2, w_2) + c\|\mathbf{w}\|_V^2 - \lambda\|\mathbf{w}\|_X^2 \leq \frac{c}{2}\|\mathbf{w}\|_V^2 + \frac{1}{2c}\|\delta\|_{V^*}^2.$$

It follows

$$\frac{1}{2}\frac{d}{dt}\|w_2\|^2 + \frac{c}{2}\|\mathbf{w}\|_V^2 - \lambda\|\mathbf{w}\|_X^2 \leq \frac{1}{2c}\|\delta\|_{V^*}^2,$$

and since  $c > 0$

$$\frac{1}{2}\frac{d}{dt}\|w_2\|^2 - \lambda\|\mathbf{w}\|_X^2 \leq \frac{1}{2c}\|\delta\|_{V^*}^2. \quad (22)$$

Next we consider the first equation of the perturbed problem

$$(\nabla w_1, \nabla w_1) + (w_2, w_1) = \langle \delta_1, w_1 \rangle.$$

We get

$$\|\nabla w_1\|^2 \leq \|w_1\|_{V_1}\|\delta_1\|_{V_1^*} + \|w_1\|\|w_2\|.$$

Applying the  $\varepsilon$ -inequality and the Poincaré inequality twice leads to

$$\begin{aligned} \|\nabla w_1\|^2 &\leq \varepsilon_1\|w_1\|_V^2 + \frac{1}{4\varepsilon_1}\|\delta_1\|_{V_1^*}^2 + \varepsilon_2\|w_1\|^2 + \frac{1}{4\varepsilon_2}\|w_2\|^2 \\ &\leq \varepsilon_1(C_P^2 + 1)\|\nabla w_1\|^2 + \frac{1}{4\varepsilon_1}\|\delta_1\|_{V_1^*}^2 + \varepsilon_2\|w_1\|^2 + \frac{1}{4\varepsilon_2}\|w_2\|^2, \end{aligned}$$

for all  $\varepsilon_1, \varepsilon_2 > 0$ . Setting  $\varepsilon_1 := \frac{1}{2(C_P^2 + 1)}$  leads to

$$\|\nabla w_1\|^2 \leq (C_P^2 + 1)\|\delta_1\|_{V_1^*}^2 + 2\varepsilon_2\|w_1\|^2 + \frac{1}{2\varepsilon_2}\|w_2\|^2.$$

Applying the Poincaré inequality again, it follows that

$$\|w_1\|^2 \leq C_P^2(C_P^2 + 1)\|\delta_1\|_{V_1^*}^2 + 2\varepsilon_2\|w_1\|^2 + C_P^2 \frac{1}{2\varepsilon_2}\|w_2\|^2.$$

With  $\varepsilon_2 := \frac{1}{4C_P^2}$  we get

$$\|w_1\|^2 \leq 2C_P^2(C_P^2 + 1)\|\delta_1\|_{V_1^*}^2 + 4C_P^4\|w_2\|^2. \quad (23)$$

Since  $\|\mathbf{w}\|_X^2 = \|w_1\|^2 + \|w_2\|^2$  we get with (22)

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|w_2\|^2 - \lambda \|w_2\|^2 &\leq \frac{1}{2c} \|\boldsymbol{\delta}\|_{V^*}^2 + \lambda \|w_1\|^2 \\ &\leq \frac{1}{2c} \|\boldsymbol{\delta}\|_{V^*}^2 + 2C_P^2(C_P^2 + 1)\|\delta_1\|_{V_1^*}^2 + 4C_P^4\|w_2\|^2. \end{aligned}$$

Setting  $\alpha := \lambda + 4C_P^4$  and  $C_1 := \frac{1}{2c} + 2C_P^2(C_P^2 + 1)$ , we obtain the estimate

$$\frac{d}{dt} \|w_2\|^2 - 2\alpha \|w_2\|^2 \leq 2C_1 \|\boldsymbol{\delta}\|_{V^*}^2.$$

Multiplying with  $e^{-2\alpha t}$ , we have

$$\frac{d}{dt}(e^{-2\alpha t} \|w_2\|^2) = e^{-2\alpha t} \frac{d}{dt} \|w_2\|^2 - 2\alpha e^{-2\alpha t} \|w_2\|^2 \leq 2C_1 e^{-2\alpha t} \|\boldsymbol{\delta}\|_{V^*}^2.$$

Integration over  $(0, t)$  yields

$$e^{-2\alpha t} \|w_2\|^2 - \|w_{20}\|^2 \leq \int_0^t 2C_1 e^{-2\alpha s} \|\boldsymbol{\delta}\|_{V^*}^2 ds,$$

for all  $t \in (0, \bar{t})$ . Multiplying by  $e^{-2\alpha t}$  and using the initial condition we obtain

$$\|w_2\|^2 \leq e^{2\alpha t} \|w_{20}\|^2 + \int_0^t 2C_1 e^{2\alpha(t-s)} \|\boldsymbol{\delta}\|_{V^*}^2 ds. \quad (24)$$

Using this estimate in (23), it follows that

$$\|w_1\|^2 \leq 2C_P^2(C_P^2 + 1)\|\delta_1\|_{V_1^*}^2 + 4C_P^4 \left[ e^{2\alpha t} \|w_{20}\|^2 + \int_0^t 2C_1 e^{2\alpha(t-s)} \|\boldsymbol{\delta}\|_{V^*}^2 ds \right].$$

Together with (24) and  $\|\delta_1\|_{V_1^*}^2 \leq \|\boldsymbol{\delta}\|_{V^*}^2$ , this estimate implies that

$$\|\mathbf{w}\|_X^2 \leq 2C_P^2(C_P^2 + 1)\|\boldsymbol{\delta}\|_{V^*}^2 + (1 + 4C_P^2)C_P^2 \left[ e^{2\alpha t} \|w_{20}\|^2 + \int_0^t 2C_1 e^{2\alpha(t-s)} \|\boldsymbol{\delta}\|_{V^*}^2 ds \right].$$

Hence the problem has the perturbation index 1.

**Example 4.4.** Consider the PDAE

$$\begin{cases} -\Delta u_1 = f_1 & \text{in } J \times \Omega, \\ \dot{u}_1 - \Delta u_2 = f_2 & \text{in } J \times \Omega, \\ \mathbf{u} = \mathbf{u}_D & \text{on } J \times \Gamma. \end{cases}$$

The first equation implies  $\|u_1\|^2 \leq C\|\delta_1\|_{V_1^*}^2$ ,  $\|\dot{u}_1\|^2 \leq C\|\dot{\delta}_1\|_{V_1^*}^2$ . Moreover we have by the Poincaré inequality

$$\|\nabla u_2\|^2 \leq \|\delta_2\|_{V_2^*}\|u_2\| + \|\dot{\delta}_1\|_{V_1^*}\|u_2\|, \quad \|u_2\|^2 \leq C[\|\dot{\delta}_1\|_{V_1^*} + \|\delta_2\|_{V_2^*}].$$

An analogous estimate can be found in the  $V$ -norm. Hence the problem has the perturbation index 2.

## References

- [1] K.E. Brenan, S.L. Campbell, L.R. Petzold, Numerical Solution of Initial-Value Problems in DAEs, Classics in Applied Mathematics, vol. 14, SIAM, Philadelphia, PA, 1996.
- [2] S.L. Campbell, C.W. Gear, The index of general nonlinear DAEs, *Numer. Math.* 72 (1995) 173–196.
- [3] S.L. Campbell, W. Marszalek, ODE/DAE integrators and MOL problems, *Z. Angew. Math. Mech.* 76 (Suppl. 1) (1996) 251–254.
- [4] S.L. Campbell, W. Marszalek, The index of an infinite dimensional implicit system, *Math. Comput. Model. Dyn. Syst.* 5 (1) (1999) 18–42.
- [5] A. Favini, A. Yagi, Degenerate Differential Equations in Banach Spaces, Marcel Dekker, New York, 1999.
- [6] E. Griepentrog, M. Hanke, R. März, Toward a better understanding of differential-algebraic equations (Introductory survey), Seminarberichte Nr. 92-1, Humboldt-Universität zu Berlin, Fachbereich Mathematik, Berlin, 1992.
- [7] E. Griepentrog, R. März, Differential-Algebraic Equations and Their Numerical Treatment, Teubner-Texte zur Mathematik, vol. 88, Teubner, Leipzig, 1986.
- [8] M. Günther, Y. Wagner, Index concepts for linear mixed systems of differential-algebraic and hyperbolic-type equations, *SIAM J. Sci. Comput.* 22 (5) (2000) 1610–1629.
- [9] E. Hairer, C. Lubich, M. Roche, The Numerical Solution of Differential Algebraic Systems by Runge–Kutta Methods, Lecture Notes in Mathematics, vol. 1409, Springer, Berlin, 1989.
- [10] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, second ed., Springer Series in Computational Mathematics, vol. 14, Springer, Berlin, 1996.
- [11] W. Lucht, K. Strehmel, Discretization based indices for semilinear partial differential algebraic equations, *Appl. Numer. Math.* 28 (1998) 371–386.
- [12] W. Lucht, K. Strehmel, Sensitivity of solutions of linear PDAE's with respect to perturbations, Technical Report 02, Martin-Luther-Universität Halle-Wittenberg, Fachbereich Mathematik und Informatik, Halle, 2000.
- [13] W. Lucht, K. Strehmel, C. Eichler-Liebenow, Indexes and special discretization methods for linear partial differential algebraic equations, *BIT* 39 (3) (1999) 484–512.
- [14] R. März, Characterizing differential algebraic equations without the use of derivative arrays, Preprint Nr. 02-8, Humboldt-Universität zu Berlin, Fachbereich Mathematik, Berlin, 1992.
- [15] W.S. Martinson, P.I. Barton, A differentiation index for partial differential equations, *SIAM J. Sci. Comput.* 21 (6) (2000) 2295–2315.
- [16] R.E. Showalter, Monotone Operators in Banach Spaces and Nonlinear Partial Differential Equations, American Mathematical Society, Providence, RI, 1997.
- [17] J. Weickert, Applications of the Theory of Differential-Algebraic Equations to Partial Differential Equations of Fluid Dynamics, Ph.D. Thesis, Technische Universität Chemnitz-Zwickau, Chemnitz, 1997.