Stability Analysis for Electromagnetic Waveguides. Part 1: Acoustic and Homogeneous Electromagnetic Waveguides

Jens M. Melenk[a], Leszek Demkowicz[b], Stefan Henneking[b]

[a]Technical University of Vienna
[b]Oden Institute, The University of Texas at Austin

Abstract

In a time-harmonic setting, we show for heterogeneous acoustic and homogeneous electromagnetic waveguides stability estimates with the stability constant depending linearly on the length L of the waveguide. These stability estimates are used for the analysis of the (ideal) ultraweak (UW) variant of the Discontinuous Petrov Galerkin (DPG) method. For this UW DPG, we show that the stability deterioration with L can be countered by suitably scaling the test norm of the method. The stability analysis applies also to the UW DPG method based on the "full envelope approximation", based on an exponential ansatz function that allows for a treating long waveguides.

Acknowledgments

J.M. Melenk was supported by the JTO fellowship and the Austrian Science Fund (FWF) under grant F65 "Taming complexity in partial differential systems". L. Demkowicz and S. Henneking were supported with AFOSR grant FA9550-19-1-0237 and NSF award 2103524.

1 Introduction

Motivation. Acoustic and electromagnetic (EM) waveguide problems have many important applications and are therefore discussed widely in the literature [22, 34, 1, 11, 19]. For many applications of interest, such as optical fibers, the propagating wave has high frequency and the waveguide length, denoted by L throughout this work, is very large compared to the wavelength. It is therefore challenging to approximate the solution numerically; to obtain an accurate solution, one must sufficiently resolve the wavelength scale and additionally counter the effect of numerical pollution to overcome stability issues of the discretization [2, 8].

While the stability of Finite Element (FE) discretizations of Helmholtz and time-harmonic Maxwell problems has been analyzed in fixed domains for increasing wave frequency ω [28, 5, 29], there is to the best of our knowledge no such corresponding analysis for the waveguide problem where ω is fixed and the waveguide length L increases. In practical applications, this is of great relevance as the available computational tools become more powerful thereby enabling numerical solution of waveguide models of realistic length scales. We discuss the present work in the context of modeling optical fiber amplifiers but emphasize that the main results of this work are relevant to FE discretization of acoustic and EM waveguide problems with the Discontinuous Petrov–Galerkin (DPG) Method [4] *in general.*

Optical amplifiers. Optical fiber amplifiers can produce highly coherent laser outputs with great efficiency, which has enabled advances in many engineering applications [20]. However, at high-power operation, these fiber laser systems are susceptible to the onset of various nonlinear effects that are adverse to the beam quality of the laser [1, 23, 33]. One particular challenge is mitigating the effects of heating of the silica-glass fiber. Under sufficient heat load, the fiber amplifier experiences a thermally-induced nonlinear effect called the transverse mode instability (TMI) [7, 21]. TMI is characterized by a sudden reduction of the beam coherence above a certain power threshold. This instability is a major limitation for the average power scaling of fiber laser systems [21]. While a scientific consensus on the thermal origins of TMI has developed over the past years, finding effective mitigation strategies that do not incite other power limiting nonlinearities remains an active field of research in fiber optics.

In the context of studying TMI and other nonlinear effects in fibers, numerical simulations play an important role. Typically, a model needs to account for two fields in the fiber amplifier: 1) the *signal laser*, a highly coherent light source that is seeded into the fiber core; and 2) the *pump field*, which provides the energy for amplification of the signal and is typically injected into the fiber cladding. A variety of different models are employed in fiber amplifier simulations (e.g., [32, 35, 30, 10] and references therein). These models are usually derived from the time-harmonic Maxwell equations, and by making additional modeling assumptions they become easier to discretize and compute than a vectorial Maxwell problem.

Vectorial Maxwell fiber amplifier model. This work is part of a continued effort to build reliable, high-fidelity FE models for investigating TMI in optical amplifiers [31, 12, 13, 15, 14]. The model consists of a system of two nonlinear time-harmonic Maxwell equations (one for the signal field and one for the pump field) coupled with each other and with the transient heat equation to account for thermal effects [31, 13]. Modeling of a 1–10 m long fiber segment involves the solution with $\mathcal{O}(1\text{–}10\text{ M})$ wavelengths. Solving such a problem with a direct FE discretization is infeasible, even on state-of-the-art supercomputers. Hence, our initial efforts focused on so-called *equivalent short fiber models*, which artificially scale physical parameters of the model, involving first $\mathcal{O}(100)$ wavelengths (using OpenMP parallelization) [31, 13], and then, more realistic models of (a tiny segment of) the actual fiber with up to $\mathcal{O}(10{,}000)$ wavelengths [12, 15, 14] (using MPI+OpenMP parallelization and up to 512 manycore compute nodes).

Because the laser light in a meter-long optical fiber has millions of wavelengths, it is extremely difficult to resolve the wavelength scale of the propagating light for the full length. For this reason, even simplified models typically resolve a longer length scale. In the context of TMI studies, it is common to resolve only the length scale of the *mode beat* between the fundamental mode and higher-order modes since the mode instabilities occur at that scale. In a typical weakly-guiding, large-mode-area fiber amplifier, the mode beat length is on order of $\mathcal{O}(1{,}000)$ wavelengths.

This brought forth the idea of the *full envelope approximation*, the solution to an alternative formulation of our vectorial Maxwell model in which the field is less oscillatory in the (longitudinal) z-direction. Consider the linear time-harmonic Maxwell equations in a non-magnetic and dielectric medium, in the absence of free charges:

$$\nabla \times \boldsymbol{E} = -i\omega\mu_0 \boldsymbol{H}\,, \tag{1.1}$$

$$\nabla \times \boldsymbol{H} = i\omega\epsilon \boldsymbol{E}\,, \tag{1.2}$$

where $i = \sqrt{-1}$, ω is the angular frequency of the light, μ_0 denotes permeability in vacuum, and $\epsilon = \epsilon(x, y, z)$ denotes permittivity. The idea of the full envelope approximation is very simple: Instead of solving for the original EM field $(\boldsymbol{E}, \boldsymbol{H})$ using (1.1)–(1.2), we seek the solution in the form:

$$\boldsymbol{E}(x, y, z) = \mathsf{E}(x, y, z)e^{-ikz}\,, \tag{1.3}$$

$$\boldsymbol{H}(x, y, z) = \mathsf{H}(x, y, z)e^{-ikz}\,, \tag{1.4}$$

where the envelope wavenumber k corresponds to a dominant frequency in the waveguide z-direction. If the effective wavenumber of the propagating fields $(\boldsymbol{E}, \boldsymbol{H})$ in the z-direction is indeed close to k, then the approximation of the new fields (E, H) requires orders of magnitude fewer elements along the fiber than the approximation of the original fields $(\boldsymbol{E}, \boldsymbol{H})$. Upon substituting the ansatz into Maxwell's equations and factoring out the exponential, we obtain the modified Maxwell model

$$\nabla \times \mathsf{E} - ik\boldsymbol{e}_z \times \mathsf{E} = -i\omega\mu_0 \mathsf{H}\,, \tag{1.5a}$$

$$\nabla \times \mathsf{H} - ik\boldsymbol{e}_z \times \mathsf{H} = i\omega\epsilon \mathsf{E}\,, \tag{1.5b}$$

where $\boldsymbol{e}_z = (0, 0, 1)^T$ is the unit vector in z-direction. This idea has turned out to be very successful enabling modeling of TMI in fiber amplifiers with several million wavelengths [16].

Scope of this work. The presented paper provides theoretical foundations for the full envelope model and the convergence of the ultraweak DPG method for acoustic and EM waveguide problems with increasing waveguide length L. Numerical studies to this extent were carried out in [12], however the convergence of the method was not analyzed therein.

We begin by introducing in Section 2 two model problems of interest: a (possibly non-homogeneous) acoustic waveguide, and a homogeneous EM waveguide. In the same section, we provide a quick overview of DPG analysis and demonstrate that the modified Maxwell model resulting from the full envelope ansatz shares the boundedness below constant with the original Maxwell operator. Section 3 analyses the Helmholtz problem (main result: Theorem 1), and Section 4 analyses the Maxwell problem (main result: Theorem 2). Section 5 presents numerical results and concludes with a discussion on the non-homogeneous EM waveguide problem.

Notation. We use $(\cdot,\cdot)_{L^2}$ and $\|\cdot\|_{L^2}$ to denote the standard sesquilinear form (antilinear in the second argument) and associated norm on the Hilbert space L^2. If clear from the context, we write $(\cdot,\cdot)$ and $\|\cdot\|$ respectively. For Sobolev spaces on Lipschitz domains $\omega \subset \mathbb{R}^d$, we follow the conventions of [24], i.e., for $s \geq 0$ we denote by $H^s(\omega)$ the closure of $C^\infty(\overline{\omega})$ under the norm

$\|\cdot\|_{H^s(\omega)}$, $\widetilde{H}^s(\omega)$ is the closure of $C_0^\infty(\omega)$ under the norm $\|\cdot\|_{H^s(\mathbb{R}^d)}$, $H^{-s}(\omega)$ is the dual of $\widetilde{H}^s(\omega)$ and $\widetilde{H}^{-s}(\omega)$ the dual of $H^s(\omega)$. The constant $C > 0$ may be different in each occurrence but does not depend on critical parameters such as the length L of the waveguide. The expression $A \lesssim B$ indicates the existence of $C > 0$ such that $A \leq CB$ with the implied constant C independent of critical parameters.

2 Model problems and DPG formulation

We present two model problems whose stability we analyze in the following section. Throughout this work, $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is a bounded Lipschitz domain, $L > 0$ and we set $\Omega = D \times (0, L)$. Throughout, we assume $\omega > 0$ given.

2.1 Helmholtz problem

On D let $a : D \to \mathrm{GL}(\mathbb{R}^{d\times d})$ be a pointwise symmetric positive definite matrix with $0 < \lambda_{min} \leq a \leq \lambda_{max} < \infty$ uniformly in $x \in D$. Set

$$\mathfrak{a} := \begin{pmatrix} a(x) & 0 \\ 0 & 1 \end{pmatrix}.$$

On $\Omega = D \times (0, L) \subset \mathbb{R}^{d+1}$ we consider

$$\begin{aligned} i\omega\boldsymbol{u} + \mathfrak{a}\nabla p &= f \quad \text{in } \Omega && (2.1a)\\ i\omega p + \operatorname{div}\boldsymbol{u} &= \boldsymbol{g} \quad \text{in } \Omega && (2.1b)\\ p &= 0 \quad \text{on } \Gamma_{\text{in}} := D \times \{0\} && (2.1c)\\ \boldsymbol{u}\cdot\boldsymbol{n} &= 0 \quad \text{on } \Gamma_{\text{lat}} := \partial D \times \{0, L\} && (2.1d)\\ i\omega\boldsymbol{u}\cdot\boldsymbol{n} - \mathrm{DtN}\, p &= 0 \quad \text{on } \Gamma_{\text{out}} := D \times \{L\}. && (2.1e) \end{aligned}$$

Here, $\boldsymbol{n}$ denotes the outer normal vector on $\partial\Omega$. We describe the operator DtN in Section 3.2. This operator ensures that waves are "outgoing" on Γ_{out}. That is, if one considers instead an infinite waveguide $D \times (0, \infty)$ such that the right-hand sides f, $\boldsymbol{g}$ vanish on $D \times (L, \infty)$, one requires that waves be going to the right.

We define the operator

$$A^{\text{helm}} \begin{pmatrix} p \\ \boldsymbol{u} \end{pmatrix} := \begin{pmatrix} i\omega\boldsymbol{u} + \mathfrak{a}\nabla p \\ i\omega p + \operatorname{div}\boldsymbol{u} \end{pmatrix} \tag{2.2}$$

with domain $D(A^{\text{helm}})$ that includes the boundary conditions and is given by

$$\begin{aligned} D(A^{\text{helm}}) = \{(p, \boldsymbol{u}) \in H^1(\Omega) \times \boldsymbol{H}(\operatorname{div}, \Omega) \mid\ & p|_{\Gamma_{\text{in}}} = 0,\ \boldsymbol{u}\cdot\boldsymbol{n} = 0 \text{ in } H^{-1/2}(\Gamma_{\text{lat}}), \\ & i\omega\boldsymbol{u}\cdot\boldsymbol{n} - \mathrm{DtN}\, p = 0 \text{ in } \widetilde{H}^{-1/2}(\Gamma_{\text{out}})\} \end{aligned} \tag{2.3}$$

2.2 Maxwell problem

Let $d = 2$. We consider the homogeneous Maxwell system for the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{H}$ that satisfy

$$\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{E} - i\omega \boldsymbol{H} &= \boldsymbol{f} && \text{in } \Omega = D \times (0, L) \subset \mathbb{R}^3, && \text{(2.4a)}\\
\boldsymbol{\nabla} \times \boldsymbol{H} + i\omega \boldsymbol{E} &= \boldsymbol{g} && \text{in } \Omega \subset \mathbb{R}^3, && \text{(2.4b)}\\
\gamma_t \boldsymbol{E} &= 0 && \text{on } \Gamma_{\text{in}}, && \text{(2.4c)}\\
\gamma_t \boldsymbol{E} &= 0 && \text{on } \Gamma_{\text{lat}}, && \text{(2.4d)}\\
\boldsymbol{E}, \boldsymbol{H} & \text{ outgoing on } \Gamma_{\text{out}}. && && \text{(2.4e)}
\end{aligned}$$

Here, $\gamma_t \boldsymbol{E} = \boldsymbol{E} \times \boldsymbol{n}$ is the tangential trace on $\partial\Omega$. We will discuss the condition to be "outgoing" in more detail in Section 4.

Problem (2.4) can be written in operator form

$$A^{\text{maxwell}} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{g} \end{pmatrix}, \qquad A^{\text{maxwell}} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} := \begin{pmatrix} \nabla \times \boldsymbol{E} - i\omega \mathbf{H} \\ \nabla \times \boldsymbol{H} + i\omega \boldsymbol{E} \end{pmatrix}. \tag{2.5}$$

2.3 Ultraweak DPG formulation and DPG essentials

Suppose we are given an injective closed operator representing a system of first-order linear Partial Differential Equations (PDEs),

$$A \,:\, L^2(\Omega) \supset D(A) \to L^2(\Omega)\,,$$

where $D(A)$ is the domain of the operator incorporating (homogeneous) Boundary Conditions (BCs). We want to solve the problem,

$$\begin{cases} u \in D(A) \\ Au = f\,. \end{cases}$$

The problem is trivially equivalent to the so-called *strong variational formulation:*

$$\begin{cases} u \in H_A(\Omega) \\ (Au, v) = (f, v) \quad \forall\, v \in L^2(\Omega)\,, \end{cases}$$

where $H_A(\Omega) := D(A)$ is a Hilbert space equipped with the graph norm:

$$\|u\|^2_{H_A} := \|u\|^2 + \|Au\|^2\,.$$

The ultraweak (UW) variational formulation is obtained by integrating by parts and passing *all* derivatives to the test function. It reads as follows:

$$\begin{cases} u \in L^2(\Omega) \\ (u, A^* v) = (f, v) \quad \forall\, v \in H_{A^*}(\Omega)\,, \end{cases} \tag{2.6}$$

where

$$A^* \,:\, L^2(\Omega) \supset D(A^*) \to L^2(\Omega)$$

is the L^2-adjoint of the original operator in the sense of closed operator theory, assumed to be injective as well, and the test space $H_{A^*}(\Omega) := D(A^*)$ has been equipped with the scaled adjoint graph norm:

$$\|v\|_{H_{A^*}}^2 := \beta^2 \|v\|^2 + \|A^* v\|^2$$

with $\beta \geq 0$. As the test functions are now more regular (and the solution less regular), the right-hand side may be upgraded to an arbitrary continuous and antilinear functional $l(\cdot)$ on $H_{A^*}(\Omega)$.

Lemma 1

Assume that that the operator A is bounded below with a (boundedness below) constant $\alpha > 0$:

$$\alpha \|u\| \leq \|Au\| \qquad \forall u \in D(A)\,.$$

The UW formulation (2.6) is then well-posed with the inf-sup constant

$$\gamma \geq \left[1 + \left(\frac{\beta}{\alpha}\right)^2\right]^{-\frac{1}{2}}.$$

▮

Proof: The Closed Range Theorem implies that adjoint A^* is bounded below with the same constant α. Take an arbitrary $u \in L^2(\Omega)$ and consider the corresponding solution v_u of the adjoint problem,

$$\begin{cases} v_u \in D(A^*) \\ A^* v_u = u\,. \end{cases}$$

We have:

$$\|v_u\| \leq \alpha^{-1}\|u\| \quad \Rightarrow \quad \|v_u\|_{H_{A^*}}^2 = \|A^* v\|^2 + \beta^2 \|v\|^2 \leq [1 + \left(\frac{\beta}{\alpha}\right)^2]\|u\|^2\,,$$

and, in turn,

$$\sup_{v \in D(A^*)} \frac{|(u, A^* v)|}{\|v\|_{A^*}} \geq \frac{|(u,u)|}{\|v_u\|_{A^*}} \geq [1 + \left(\frac{\beta}{\alpha}\right)^2]^{-\frac{1}{2}} \frac{|(u,u)|}{\|u\|} = [1 + \left(\frac{\beta}{\alpha}\right)^2]^{-\frac{1}{2}} \|u\|\,.$$

Finally, the conjugate operator coincides with the strong operator A and its injectivity implies the existence of the solution for any right-hand side $l \in H_{A^*}(\Omega)'$. ▮

Note that, for $\beta = 0$, the inf-sup constant is one. The DPG Method with Optimal Test Functions is based on extending the UW formulation to a larger class of *broken* test functions $H_{A^*}(\Omega_h)$ [3]. Unfortunately, for $\beta = 0$, the form $b(u, v) = (u, A^* v)$ is not *localizable*, i.e. the adjoint test norm cannot be extended to the broken test space and *we must* use a positive scaling constant $\beta > 0$. In principle though, by employing a sufficiently small constant β, we can make the inf-sup constant as close to one as we wish. It is not intuitive at all that the UW formulation allows for *improving* the stability of the original operator A in the strong setting.

Finally, the stability of the UW formulation *is inherited* by the *broken formulation* [3]. In the *ideal DPG* method, the optimal test functions are assumed to be computed exactly; the discrete inf-sup constant is bounded from below by the continuous inf-sup constant which indicates the relevance of understanding the continuous variational problem. The ideal DPG method is merely a logical construct on the way to analyze the *practical DPG* method where the optimal test functions are computed approximately. The additional error resulting from the approximation of the test functions is analyzed with the help of appropriate Fortin operators [4].

2.4 The full envelope approximation (1.5)

The analysis of the stability of the full envelope approximation method (1.5) turns out to be surprisingly simple.

Lemma 2
Let $\widetilde{A}$ be the operator corresponding to the full envelope ansatz, i.e.,

$$\widetilde{A}\widetilde{u} := e^{ikz} A(e^{-ikz}\widetilde{u}) \,,$$

where $A \in \{A^{\mathrm{helm}}, A^{\mathrm{maxwell}}\}$ denotes the operator corresponding to the acoustic or EM waveguide problem. Then, the operator $\widetilde{A}$ is bounded below if and only if operator A is bounded below, and the corresponding boundedness below constants are identical:

$$\|Au\| \geq \alpha \|u\| \quad \Leftrightarrow \quad \|\widetilde{A}\widetilde{u}\| \geq \alpha \|\widetilde{u}\| \,.$$

▮

Proof: We observe

$$\|\widetilde{A}\widetilde{u}\| = \|e^{ikz} A(e^{-ikz}\widetilde{u})\| = \|A(e^{-ikz}\widetilde{u})\| \geq \alpha \|e^{-ikz}\widetilde{u}\| = \alpha \|\widetilde{u}\| \,.$$

▮

3 Stability

Section 2.3 shows that at the heart of the analysis of the ultraweak DPG method is the understanding of the stability properties of the operator A. In this section, we provide the stability of operator A^{helm} of the Helmholtz problem (2.1) making the dependence on the length L of the waveguide explicit. In view of Lemma 1, this analysis then provides a guideline for selecting the parameter β in the DPG method.

3.1 Analysis of the 1D Helmholtz equation

Since the domain $\Omega = D \times (0, L)$ has product structure and the coefficients of the Helmholtz equation are independent of the longitudinal variable, a decoupling into transversal modes is possible and leads to a stability analysis analysis of 1D equations. In the present section, we present the necessary 1D stability results.

Lemma 3 ([27, Thm. 4.3])
Let $\hat{I} := (0,1)$. *Let* $\hat{\kappa} \in \mathbb{C}_{\geq 0} := \{z \in \mathbb{C} \colon \operatorname{Re} \geq 0\}$ *satisfy* $|\hat{\kappa}| \geq 1$. *Introduce* $H^1_{(0}(\hat{I}) := \{u \in H^1(\hat{I}) \colon u(0) = 0\}$ *and the sesquilinear form*

$$a^{1D}_{\hat{\kappa},\hat{I}}(u,v) := (u',v')_{L^2(0,1)} + \hat{\kappa}^2 (u,v)_{L^2(0,1)} + \hat{\kappa} u(1)\overline{v}(1).$$

Introduce on $H^1(\hat{I})$ *the norm* $\|\cdot\|_{1,|\kappa|,\hat{I}}$ *by* $\|u\|^2_{1,|\hat{\kappa}|} := \|u'\|^2_{L^2(\hat{I})} + |\hat{\kappa}|^2 \|u\|^2_{L^2(\hat{I})}$. *Then there is* $c > 0$ *independent of* κ *such that for both choices* $V = H^1(\hat{I})$ *and* $V = H^1_{(0}(\hat{I})$ *there holds*

$$\hat{\gamma}_{\hat{\kappa}} := \inf_{0 \neq u \in V} \sup_{0 \neq V} \frac{\operatorname{Re} a^{1D}_{\hat{\kappa},\hat{I}}(u,v)}{\|u\|_{1,|\hat{\kappa}|,\hat{I}} \|v\|_{1,|\hat{\kappa}|,\hat{I}}} \geq \frac{1}{1 + c \frac{\operatorname{Im} \hat{\kappa}}{1 + \operatorname{Re} \hat{\kappa}}} \tag{3.1}$$

$\blacksquare$

Proof: The proof follows essentially from [27, Thm. 4.3], where a similar sesquilinear form in 2D and 3D is considered. Details are given in Appendix A. $\blacksquare$

It is convenient to introduce for $\kappa \in \mathbb{C}$ and an interval $I = (0,L)$ the sesquilinear form a^{1D}_κ and the norm $\|\cdot\||_{1,|\kappa|}$ by

$$a^{1D}_\kappa(u,v) := (u',v')_{L^2(I)} + \kappa^2 (u,v)_{L^2(I)} + \kappa u(L)\overline{v}(L), \tag{3.2}$$

$$\|q\|^2_{1,|\kappa|} := \|q'\|^2_{L^2(I)} + |\kappa|^2 \|q\|^2_{L^2(I)}\,. \tag{3.3}$$

Lemma 4
Let $I = (0,L)$ *and* $V = H^1(I)$ *or* $V = H^1_{(0}(I) := \{w \in H^1(I) \colon w(0) = 0\}$. *Let* $\kappa \in \mathbb{C}_{\geq 0}$ *with* $|\kappa| \geq 1$. *Consider the two problems: Find* q_1, $q_2 \in V$ *such that*

$$a^{1D}_\kappa(q_1,w) = (f,w)_{L^2(I)} \quad \forall w \in V\,,$$
$$a^{1D}_\kappa(q_1,w) = (f,w')_{L^2(I)} \quad \forall w \in V\,.$$

Then the following holds:

(i) There are C, $c > 0$ *depending only on a lower bound for* $L|\kappa|$ *such that with*

$$\gamma := \frac{1}{1 + c \frac{|\operatorname{Im}(L\kappa)|}{1 + \operatorname{Re}(L\kappa)}} \tag{3.4}$$

there holds

$$\|q_1\|_{1,|\kappa|} \leq C \frac{L}{\gamma |\kappa|} \|f\|_{L^2(I)}\,, \qquad \|q_2\|_{1,|\kappa|} \leq C \gamma^{-1} \|f\|_{L^2(I)}\,.$$

(ii) If $\kappa \in i\mathbb{R}$, then we have for a constant $C > 0$ depending only on a lower bound for $L|\kappa|$,

$$\|q_1\|_{1,|\kappa|} \leq CL\|f\|_{L^2(I)}\,, \qquad \|q_2\|_{1,|\kappa|} \leq CL|\kappa|\|f\|_{L^2(I)}\,,$$

(iii) If $\kappa > 0$, then we have for a constant $C > 0$ depending only on a lower bound for $L\kappa$

$$\|q_1\|_{1,\kappa} \leq C\kappa^{-1}\|f\|_{L^2(I)}\,, \qquad \|q_2\|_{1,\kappa} \leq C\|f\|_{L^2(I)}\,.$$

▮

Proof: Rescaling the equation posed on $(0, L)$ to an equation posed on $(0, 1) =: \hat{I}$, we get by denoting the rescaled functions with a $\hat{\ }$ and $\hat{\kappa} := L\kappa$ and spaces $\hat{V} = H^1(\hat{I})$ if $V = H^1(I)$ and $\hat{V} = H^1_{(0}(\hat{I})$ if $V = H^1_{(0}(I)$

$$\begin{aligned}(\hat{q}_1', w')_{L^2(0,1)} + \hat{\kappa}^2(\hat{q}_1, w)_{L^2(0,1)} + \hat{\kappa}\hat{q}_1(1)\overline{w(1)} &= L^2(\hat{f}, w)_{L^2(I)} \quad \forall w \in \hat{V},\\ (\hat{q}_2', w')_{L^2(I)} + \hat{\kappa}^2(\hat{q}_2, w)_{L^2(I)} + \hat{\kappa}\hat{q}_2(1)\overline{w(1)} &= L(\hat{f}, w')_{L^2(I)} \quad \forall w \in \hat{V}.\end{aligned}$$

In terms of the inf-sup constant $\hat{\gamma}_{\hat{\kappa}}$ of (3.1), i.e.,

$$\gamma := \inf_{q\in\hat{V}} \sup_{w\in\hat{V}} \frac{|a^{1D}_{\hat{\kappa}}(q, w)|}{\|q\|_{1,|\hat{\kappa}|}\|w\|_{1,|\hat{\kappa}|}}$$

we have

$$\begin{aligned}\|\hat{q}_1\|_{1,|\hat{\kappa}|} &\leq C\gamma^{-1}|\hat{\kappa}|^{-1}\|L^2\hat{f}\|_{L^2(\hat{I})},\\ \|\hat{q}_2\|_{1,|\hat{\kappa}|} &\leq C\gamma^{-1}\|L\hat{f}\|_{L^2(\hat{I})}\end{aligned}$$

with $C > 0$ depending only on a lower bound for $|\hat{\kappa}|$. Scaling back to $(0, L)$ yields

$$\begin{aligned}\|q_1\|_{1,|\kappa|} &\sim L^{-1/2}\|\hat{q}_1\|_{1,|\hat{\kappa}|} \leq CL^2L^{-1/2}\gamma^{-1}|\hat{\kappa}|^{-1}\|\hat{f}\|_{L^2(\hat{I})} \leq CL\gamma^{-1}|\hat{\kappa}|^{-1}\|f\|_{L^2(I)},\\ \|q_2\|_{1,|\kappa|} &\sim L^{-1/2}\|\hat{q}_2\|_{1,|\hat{\kappa}|} \leq CL^{1/2}\gamma^{-1}\|\hat{f}\|_{L^2(\hat{I})} \leq C\gamma^{-1}\|f\|_{L^2(I)}.\end{aligned}$$

Proof of (i): By Lemma 3, γ has the form (3.4). The result follows.

Proof of (ii): For $\kappa \in i\mathbb{R}$, the inf-sup constant γ of (3.4) satisfies $\gamma \sim |\hat{\kappa}|^{-1}$ with a implied constants depending on a lower bound for $|\hat{\kappa}|$. The result follows.

Proof of (iii): For $\kappa > 0$, the inf-sup constant γ of (3.4) satisfies $\gamma \sim 1$ with implied constants depending only on a lower bound for $|\hat{\kappa}|$. The result follows. ▮

3.2 The DtN operator

We describe the DtN-operator at the outflow boundary Γ_{out}. This is achieved in terms of a modal decomposition obtained by a suitable eigenvalue problem in transverse direction.

Let $(\varphi_n, \lambda_n)_{n\in\mathbb{N}} \subset H^1(D) \setminus \{0\} \times \mathbb{R}^+$ be the eigenpairs of the operator $p \mapsto -\operatorname{div}(a\nabla u)$, i.e.,

$$-\operatorname{div}(a\nabla\varphi_n) = \lambda_n\varphi_n \quad \text{in } D, \tag{3.5a}$$

$$\boldsymbol{n}\cdot a\nabla\varphi_n = 0 \qquad \text{on } \partial D \tag{3.5b}$$

with the normalization $\|\varphi_n\|_{L^2(D)} = 1$ for all $n \in \mathbb{N}$. Here, the normal vector $\boldsymbol{n}$ is the outer normal on ∂D. We have the orthogonalities

$$(\varphi_n, \varphi_m)_{L^2(D)} = \delta_{nm}, \qquad (a\nabla\varphi_n, \nabla\varphi_m)_{L^2(D)} = \delta_{nm}\lambda_n. \tag{3.6}$$

The positive values λ_n are assumed sorted in ascending order and listed according to their multiplicity. We collect them in the spectrum $\sigma = \{\lambda_n \colon n \in \mathbb{N}\}$. It is convenient to derive the operator DtN in (2.1d) using the second-order formulation for $(f, \boldsymbol{g}) = (0, 0)$, i.e., consider

$$-\operatorname{div}(\mathfrak{a}\nabla p) - \omega^2 p = 0 \quad \text{on } D \times (L, \infty).$$

Making the ansatz $p(x,z) = \sum_n p_n(z)\varphi_n(x)$ yields the 1D equations

$$-p_n''(z) + (\lambda_n - \omega^2)p_n = 0 \quad \text{on } (L, \infty)$$

with the fundamental solutions $e^{\pm\kappa_n z}$, where

$$\kappa_n := \sqrt{\lambda_n - \omega^2} \tag{3.7}$$

and the square root is taken to be the principal branch. Writing $p|_{\Gamma_{\text{out}}} = \sum_n p_n(L)\varphi_n(x)$, the "outgoing" solution is defined on (L, ∞) to be

$$p(x,z) = \sum_n p_n(L)\varphi_n(x)e^{-\kappa_n(z-L)}.$$

The operator DtN is given by $p|_{\Gamma_{\text{out}}} \mapsto (\partial_z p(x,z))|_{z=L}$, viz.,

$$\operatorname{DtN} p := -\sum_n p_n\kappa_n\varphi_n(x), \qquad p_n := (p, \varphi_n)_{L^2(D)}.$$

The indices $n \in \mathbb{N}$ with $\kappa_n > 0$ are called *evanescent modes*, those with $\kappa_n \in i\mathbb{R} \setminus \{0\}$ are the *propagating modes*. We assume throughout that

$$\kappa_n \neq 0 \qquad \forall n \in \mathbb{N} \tag{3.8}$$

so that we can write $\mathbb{N} = I_{\text{prop}} \dot\cup I_{\text{eva}}$ with the finite set I_{prop} of propagating modes and the infinite set I_{eva} of evanescent modes:

$$I_{\text{prop}} := \{n \in \mathbb{N} \,|\, \kappa_n \in i\mathbb{R}\}, \qquad I_{\text{eva}} := \{n \in \mathbb{N} \,|\, \kappa_n > 0\}. \tag{3.9}$$

Concerning the mapping properties of DtN, we have

$$\operatorname{DtN} : H^{1/2}(\Gamma_{\text{out}}) \to \widetilde{H}^{-1/2}(\Gamma_{\text{out}}) := (H^{1/2}(\Gamma_{\text{out}}))'$$

is bounded, linear, which follows from the representation of DtN: Since $H^{1/2}(\Gamma_{\text{out}})$ is the interpolation space between $L^2(\Gamma_{\text{out}})$ and $H^1(\Gamma_{\text{out}})$ (see, e.g., [24]), we we have the norm equivalence $\|p\|_{H^{1/2}(\Gamma_{\text{out}})} \sim \sum_n |p_n|^2\sqrt{\lambda_n}$, where $p = \sum_n p_n\varphi_n$ with $p_n = (p, \varphi_n)_{L^2(D)}$. Hence, writing p, $q \in H^{1/2}(\Gamma_{\text{out}})$ in the form $p = \sum_n p_n\varphi_n$ and $q = \sum_n q_n\varphi_n$, we estimate

$$|\langle \operatorname{DtN} p, q\rangle| = \left|-\sum_n \kappa_n p_n \overline{q}_n\right| \lesssim \|p\|_{H^{1/2}(\Gamma_{\text{out}})}\|q\|_{H^{1/2}(\Gamma_{\text{out}})}.$$

3.3 Stability estimates for A^{helm}

Introduce $H^1_{\Gamma_{\text{in}}} := \{v \in H^1(\Omega) \,|\, v|_{\Gamma_{\text{in}}} = 0\}$. In the ensuing analysis, we will need the following observation about the sequences $\{\lambda_n\}_n$ and $\{\kappa_n\}_n$: Noting that there are only finitely many propagating modes and that we assumed (3.8) we have, for a constant that depends on ω,

$$\max_{n \in I_{\text{prop}}} \left(|\kappa_n|^{-1} + |\kappa_n| \right) \leq C, \tag{3.10a}$$

$$\max_{n \in I_{\text{prop}}} |\sqrt{\lambda_n}| \leq C, \tag{3.10b}$$

$$\max_{n \in \mathbb{N}} \frac{|\sqrt{\lambda_n}|}{|\kappa_n|} \leq C \tag{3.10c}$$

Lemma 5
Let $f \in L^2(\Omega)$. The solution $p \in H^1_{\Gamma_{in}}$ of

$$(\mathfrak{a}\nabla p, \nabla v)_{L^2(\Omega)} - \omega^2 (p, v)_{L^2(\Omega)} - \langle \text{DtN}\, p, v \rangle = (f, v)_{L^2(\Omega)} \quad \forall v \in H^1_{\Gamma_{in}} \tag{3.11}$$

satisfies

$$\|p\|_{H^1(\Omega)} \leq CL \|f\|_{L^2(\Omega)}.$$

$\blacksquare$

Proof: We make the ansatz $p(x, z) = \sum_n p_n(z) \varphi_n(x)$ and set $f_n(z) := (f(\cdot, z), \varphi_n)_{L^2(D)}$. By orthogonality properties of the functions $\{\varphi_n\}_n$ we have

$$\|p\|^2_{L^2(\Omega)} = \sum_n \|p_n\|^2_{L^2(0,L)},$$
$$\|\sqrt{a} \nabla_x p\|^2_{L^2(\Omega)} = \sum_n \lambda_n^2 \|p_n\|^2_{L^2(0,L)},$$
$$\|\partial_z p\|^2_{L^2(\Omega)} = \sum_n \|p_n'\|^2_{L^2(0,L)},$$
$$\|f\|^2_{L^2(\Omega)} = \sum_n \|f_n\|^2_{L^2(0,L)}.$$

Testing (3.11) with $v(x, z) = v_n(z) \varphi_n(x)$ with arbitrary $v_n \in H^1_{(0}(0, L)$ (cf. Lemma 4) gives due to the orthogonalities satisfied by the functions $\{\varphi_n\}_n$

$$(p_n', v_n')_{L^2(0,L)} + \kappa_n^2 (p_n, v_n)_{L^2(0,L)} - p_n(L) \kappa_n \overline{v}_n(L) = (f_n, v_n)_{L^2(0,L)}.$$

From Lemma 4, we conclude

$$\|p_n\|_{1,|\kappa_n|} \leq C \begin{cases} L \|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{prop}}, \\ \kappa_n^{-1} \|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{eva}}. \end{cases} \tag{3.12}$$

We arrive at

$$\|\sqrt{a}\nabla_x p\|^2_{L^2(\Omega)} = \sum_n \lambda_n^2 \|p_n\|^2_{L^2(0,L)} \overset{(3.12)}{\lesssim} L^2 \sum_{n\in I_{\text{prop}}} \frac{\lambda_n^2}{|\kappa_n|^2}\|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} \frac{\lambda_n^2}{|\kappa_n|^4}\|f_n\|^2_{L^2(0,L)}$$

$$\overset{(3.10)}{\lesssim} L^2\|f\|^2_{L^2(\Omega)},$$

$$\|\partial_z p\|^2_{L^2(\Omega)} = \sum_n \|p_n'\|^2_{L^2(0,L)} \overset{(3.12)}{\lesssim} L^2 \sum_{n\in I_{\text{prop}}} \|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} \lesssim L^2\|f\|^2_{L^2(\Omega)},$$

$$\|p\|^2_{L^2(\Omega)} = \sum_n \|p_n\|^2_{L^2(0,L)} \overset{(3.12)}{\lesssim} L^2 \sum_{n\in I_{\text{prop}}} |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{prop}}} |\kappa_n|^{-4}\|f_n\|^2_{L^2(0,L)}$$

$$\lesssim L^2\|f\|^2_{L^2(\Omega)} + \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-4}\|f_n\|^2_{L^2(0,L)} \lesssim L^2\|f\|^2_{L^2(\Omega)}.$$

∎

Lemma 6
Let $\boldsymbol{f} \in L^2(\Omega)$. *The solution* $q \in H^1_{\Gamma_{in}}$ *of*

$$(\mathfrak{a}\nabla p, \nabla v)_{L^2(\Omega)} - \omega^2(p,v)_{L^2(\Omega)} - \langle \mathrm{DtN}\, p, v\rangle = (\boldsymbol{f}, \nabla v)_{L^2(\Omega)} \quad \forall v \in H^1_{\Gamma_{in}} \tag{3.13}$$

satisfies

$$\|p\|_{H^1(\Omega)} \leq CL\|\boldsymbol{f}\|_{L^2(\Omega)}.$$

∎

Proof: We write the vector $\boldsymbol{f}$ as $\boldsymbol{f} = (\boldsymbol{f}_x, f_z)^\top$ with a vector-valued function $\boldsymbol{f}_x$ and a scalar function f_z. By linearity of the problem, we may consider the cases $(\boldsymbol{f}_x, 0)^\top$ and $(0, f_z)^\top$ as right-hand sides separately. For $\boldsymbol{f}_x = 0$, we proceed as in Lemma 5 by writing $f_z = \sum_n f_n(z)\varphi_n(x)$ and get with Lemma 4 for the corresponding functions p_n

$$\|p_n\|_{1,|\kappa_n|} \leq C \begin{cases} L|\kappa_n|\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{prop}}, \\ \|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{eva}} \end{cases}$$
$$\leq CL\|f_n\|_{L^2(0,L)}$$

since $\max_{n\in I_{\text{prop}}} |\kappa_n| \leq C$. We may repeat the calculations performed in Lemma 5 to establish

$$\|\sqrt{a}\nabla_x p\|^2_{L^2(\Omega)} = \sum_n \lambda_n \|p_n\|^2_{L^2(0,L)} \leq CL^2 \sum_n \frac{\lambda_n}{|\kappa_n|^2}\|f_n\|^2_{L^2(0,L)} \leq C\|f_z\|^2_{L^2(\Omega)},$$

$$\|\partial_z p\|^2_{L^2(\Omega)} = \sum_n \|p_n'\|^2_{L^2(0,L)} \leq CL^2 \sum_n \|f_n\|^2_{L^2(0,L)} \leq CL^2\|f_z\|^2_{L^2(\Omega)},$$

$$\|p\|^2_{L^2(\Omega)} = \sum_n \|p_n\|^2_{L^2(0,L)} \leq CL^2 \sum_n |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} \leq CL^2\|f_z\|^2_{L^2(\Omega)}.$$

For the case of the right-hand side $\boldsymbol{f} = (\boldsymbol{f}_x, 0)^\top$, we define

$$f_n(z) := (\boldsymbol{f}_x(\cdot, z), \nabla\varphi_n)_{L^2(D)} = (a^{-1}\boldsymbol{f}_x(\cdot, z), a\nabla\varphi_n)_{L^2(D)}$$

and note by the fact that the functions $\{\|\sqrt{a}\nabla\varphi_n\|^{-1}_{L^2(D)}\nabla\varphi_n\}_n$ are an orthonormal (with respect to $(a\cdot,\cdot)_{L^2(D)}$) basis of its span that

$$\sum_n \lambda_n^{-1}\|f_n\|^2_{L^2(0,L)} = \int_0^L \sum_n \frac{1}{\|\sqrt{a}\nabla\varphi_n\|^2_{L^2(D)}}|(a^{-1}\boldsymbol{f}_x(\cdot,z), a\nabla\varphi_n)_{L^2(D)}|^2 \leq \|a^{-1}\boldsymbol{f}_x\|^2_{L^2(\Omega)}.$$

We expand the solution p as $p = \sum_n p_n(z)\varphi_n(x)$. Testing the equation with functions of the form $v_n(z)\varphi_n(x)$ yields again an equation for the coefficients q_n:

$$-\kappa_n^2(p_n, v_n)_{L^2(0,L)} + (p_n', v_n')_{L^2(0,L)} \pm \kappa_n p_n(L)\overline{v}_n(L) = (f_n, v_n)_{L^2(0,L)} \qquad \forall v_n \in H^1_{(0}(0,L).$$

By Lemma 4 we get

$$\|p_n\|_{1,|\kappa_n|} \leq C \begin{cases} L\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{prop}}\ , \\ |\kappa_n|^{-1}\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{eva}}\ . \end{cases}$$

Hence,

$$\|\sqrt{a}\nabla_x p\|^2_{L^2(\Omega)} = \sum_n \lambda_n\|p_n\|^2_{L^2(0,L)} \leq CL^2 \sum_{n\in I_{\text{prop}}} \lambda_n\|f_n\|^2_{L^2(0,L)} + C\sum_{n\in I_{\text{eva}}} \frac{\lambda_n}{|\kappa_n|^4}\|f_n\|^2_{L^2(0,L)} \leq CL\|\boldsymbol{f}_x\|^2_{L^2(\Omega)},$$

$$\|\partial_z p\|^2_{L^2(\Omega)} = \sum_n \|p_n'\|^2_{L^2(0,L)} \leq CL^2 \sum_{n\in I_{\text{prop}}} \|f_n\|^2_{L^2(0,L)} + C\sum_{n\in I_{\text{eva}}} |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} \leq C\|\boldsymbol{f}_x\|^2_{L^2(\Omega)}$$

$$\|p\|^2_{L^2(\Omega)} = \sum_n \|p_n\|^2_{L^2(0,L)} \leq CL^2 \sum_{n\in I_{\text{prop}}} |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-4}\|f_n\|^2_{L^2(0,L)} \leq C\|\boldsymbol{f}_x\|^2_{L^2(\Omega)}.$$

Putting together the above results proves the claim. ∎

THEOREM 1

Assume (3.8). There is a constant $C > 0$ (depending on $\mathfrak{a}$ and ω but independent of L) such that for all $(f, \boldsymbol{g}) \in L^2(\Omega)$ the problem (2.1) has a unique solution $(p, \boldsymbol{u}) \in D(A^{\text{helm}})$ with

$$\|\boldsymbol{u}\|_{\boldsymbol{H}(\text{div},\Omega)} + \|p\|_{H^1(\Omega)} \leq CL\left[\|f\|_{L^2(\Omega)} + \|\boldsymbol{g}\|_{L^2(\Omega)}\right]. \tag{3.14}$$

*In particular, [**todo:** check this!] for all $(\boldsymbol{u}, p) \in D(A^{\text{helm}})$*

$$\|A^{\text{helm}}(p, \boldsymbol{u})^\top\|_{L^2(\Omega)} \geq CL^{-1}\|(\boldsymbol{u}, p)\|_{L^2(\Omega)}. \tag{3.15}$$

∎

Proof: *Proof of (3.15):* First, we note that $A^{\text{helm}} : D(A^{\text{helm}}) \to L^2(\Omega)$ is injective. Indeed, $A^{\text{helm}}(p, \boldsymbol{u})^\top = 0$ implies $\boldsymbol{u} = -i\omega\nabla p$ and therefore $p \in H^1(\Omega)$ satisfies a homogeneous second-order equation. Together with the boundary condition, one checks that $p = 0$ so that also $\boldsymbol{u} = 0$.

Abbreviate for the two components of $A^{\mathrm{helm}}(p,\boldsymbol{u})^\top$

$$\boldsymbol{g} := i\omega\boldsymbol{u} + \mathfrak{a}\nabla p \in L^2(\Omega), \qquad f := \operatorname{div}\boldsymbol{u} + i\omega p \in L^2(\Omega).$$

Hence, $(p,\boldsymbol{u})$ satisfy for smooth $(\widetilde{p},\widetilde{\boldsymbol{u}})$

$$\begin{aligned}(i\omega\boldsymbol{u},\widetilde{\boldsymbol{u}})_{L^2(\Omega)} + (\mathfrak{a}\nabla p,\widetilde{\boldsymbol{u}})_{L^2(\Omega)} &= (\boldsymbol{g},\widetilde{\boldsymbol{u}})_{L^2(\Omega)}\,,\\ (\operatorname{div}\boldsymbol{u},\widetilde{p})_{L^2(\Omega)} + (i\omega p,\widetilde{p})_{L^2(\Omega)} &= (f,\widetilde{p})_{L^2(\omega)}\,.\end{aligned}$$

Considering $\widetilde{p}$ with $\widetilde{p}|_{\Gamma_{\mathrm{in}}} = 0$ and using the boundary conditions satisfied by $(p,\boldsymbol{u})$ (i.e., $(p,\boldsymbol{u}) \in D(A^{\mathrm{helm}})$) yields after an integration by parts

$$\begin{aligned}(i\omega\boldsymbol{u},\widetilde{\boldsymbol{u}})_{L^2(\Omega)} + (\mathfrak{a}\nabla p,\widetilde{\boldsymbol{u}})_{L^2(\Omega)} &= (\boldsymbol{g},\widetilde{\boldsymbol{u}})_{L^2(\Omega)}\,,\\ -(\boldsymbol{u},\nabla\widetilde{p})_{L^2(\Omega)} + (i\omega p,\widetilde{p})_{L^2(\Omega)} + \frac{1}{i\omega}\langle \operatorname{DtN} p,\widetilde{p}\rangle_{\Gamma_{\mathrm{out}}} &= (f,\widetilde{p})_{L^2(\Omega)}\,.\end{aligned}$$

Selecting $\widetilde{\boldsymbol{u}} = -\frac{1}{i\omega}\nabla\widetilde{p}$ yields

$$\begin{aligned}(\boldsymbol{u},\nabla\widetilde{p})_{L^2(\Omega)} + \frac{1}{i\omega}(\mathfrak{a}\nabla p,\nabla\widetilde{p})_{L^2(\Omega)} &= \frac{1}{i\omega}(\boldsymbol{g},\nabla\widetilde{p})_{L^2(\Omega)}\,,\\ -(\boldsymbol{u},\nabla\widetilde{p})_{L^2(\Omega)} + (i\omega p,\widetilde{p})_{L^2(\Omega)} + \frac{1}{i\omega}\langle \operatorname{DtN} p,\widetilde{p}\rangle_{\Gamma_{\mathrm{out}}} &= (f,\widetilde{p})_{L^2(\Omega)}\,,\end{aligned}$$

so that, by adding these two equations and multiplying by $i\omega$, we arrive at

$$(\mathfrak{a}\nabla p,\nabla\widetilde{p})_{L^2(\Omega)} - \omega^2(p,\widetilde{p})_{L^2(\Omega)} + \langle \operatorname{DtN} p,\widetilde{p}\rangle_{\Gamma_{\mathrm{out}}} = (\boldsymbol{g},\nabla\widetilde{p})_{L^2(\Omega)} + i\omega(f,\widetilde{p})_{L^2(\Omega)} \quad \forall \widetilde{p} \in H^1_{\Gamma_{\mathrm{in}}}\,. \tag{3.16}$$

From Lemmas 5, 6 we infer

$$\|p\|_{H^1(\Omega)} \le CL\left[\|\boldsymbol{g}\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right],$$

which in turn yields

$$\|(p,\boldsymbol{u})\|^2_{L^2(\Omega)} \le \|p\|^2_{L^2(\Omega)} + \|\boldsymbol{u}\|^2_{L^2(\Omega)} \le \|p\|^2_{L^2(\Omega)} + 2\omega^{-2}\|\nabla p\|^2_{L^2(\Omega)} + 2\omega^{-2}\|\boldsymbol{g}\|^2_{L^2(\Omega)}\,.$$

In total, we arrive at $\|(p,\boldsymbol{u})\|_{L^2(\Omega)} \le CL\left[\|\boldsymbol{g}\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right]$, i.e., $\|(p,\boldsymbol{u})\|_{L^2(\Omega)} \le CL\|A^{\mathrm{helm}}(p,\boldsymbol{u})^\top\|_{L^2(\Omega)}$, which is (3.15).

Proof of (3.14): To see solvability for any $(f,\boldsymbol{g}) \in L^2(\Omega)$, we reverse the above arguments. Lemmas 5, 6 imply solvability of (3.16) for $p \in H^1_{\Gamma_{\mathrm{in}}}$. Next, setting $\boldsymbol{u} := (i\omega)^{-1}(\boldsymbol{g} - \mathfrak{a}\nabla p)$, one infers from (3.16) that $\boldsymbol{u} \in \boldsymbol{H}(\operatorname{div},\Omega)$ with $\operatorname{div}\boldsymbol{u} = f \in L^2(\Omega)$. This implies the estimate (3.14). To see that $(p,\boldsymbol{u}) \in D(A^{\mathrm{helm}})$, we infer, using the fact that $\operatorname{div}\boldsymbol{u} = f$, that $\boldsymbol{u}\cdot\boldsymbol{n} = 0$ in $H^{-1/2}(\Gamma_{\mathrm{lat}})$ and $\boldsymbol{u}\cdot\boldsymbol{n} - \operatorname{DtN} p = 0$ in $H^{-1/2}(\Gamma_{\mathrm{out}})$. To see that in fact $\boldsymbol{u}\cdot\boldsymbol{n} = 0$ in $\widetilde{H}^{-1/2}(\Gamma_{\mathrm{lat}})$, we note that $\boldsymbol{u}\cdot\boldsymbol{n} \in H^{-1/2}(\partial\Omega\setminus\overline{\Gamma_{\mathrm{in}}})$ with $\operatorname{supp}\boldsymbol{u}\cdot\boldsymbol{n} \subset \overline{\Gamma_{\mathrm{out}}}$. Hence, by [24, Thm. 3.29], we have $\boldsymbol{u}\cdot\boldsymbol{n} \in \widetilde{H}^{-1/2}(\Gamma_{\mathrm{out}})$. Since by construction $\operatorname{DtN} p \in \widetilde{H}^{-1/2}(\Gamma_{\mathrm{out}})$, the difference $z := \boldsymbol{u}\cdot\boldsymbol{n} - \operatorname{DtN} p \in \widetilde{H}^{-1/2}(\Gamma_{\mathrm{out}})$. Since $z|_{\Gamma_{\mathrm{out}}} = 0$, we conclude that the zero extension $\widetilde{z}$ to $H^{-1/2}(\partial\Omega\setminus\overline{\Gamma_{\mathrm{in}}})$ is a distribution on $\partial\Omega\setminus\overline{\Gamma_{\mathrm{in}}}$ with $\operatorname{supp} z \subset \partial\Gamma_{\mathrm{out}}$. However, since $H^{1/2}(\partial\Omega\setminus\overline{\Gamma_{\mathrm{in}}})$-functions do not admit traces on lower-dimensional manifolds, the support of the distribution $z \in H^{-1/2}(\partial\Omega\setminus\overline{\Gamma_{\mathrm{in}}})$ must be empty. Hence, $z = 0$ in $H^{-1/2}(\partial\Omega\setminus\overline{\Gamma_{\mathrm{in}}})$, i.e., $z = 0$ in $\widetilde{H}^{-1/2}(\Gamma_{\mathrm{out}})$. This shows that $(p,\boldsymbol{u}) \in D(A^{\mathrm{helm}})$. ∎

4 Stability for Maxwell's equations

The stability analysis of the Maxwell system (2.4) proceeds similarly to the case of the Helmholtz equation in that the use of a suitable system of functions in the transverse direction reduces the stability analysis to one for a decoupled ordinary differential equation (ODE) system.

4.1 Motivation: the waveguide eigenproblems

We motivate our choice of transversal expansion system by deriving appropriate transversal eigenvalue problems for the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{H}$. We write the electric field $\boldsymbol{E} = \boldsymbol{E}(x,z)$, $\boldsymbol{H} = \boldsymbol{H}(x,z)$ with $x \in D$, $z \in (0,L)$ as

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{E}_t \\ E_3 \end{pmatrix}, \quad \boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_t \\ H_3 \end{pmatrix}$$

with the transversal components $\boldsymbol{E}_t$, $\boldsymbol{H}_t : D \times (0,L) \to \mathbb{R}^2$ and the longitudinal components E_3, $H_3 : D \times (0,L) \to \mathbb{R}$. The 2D curl operators $\boldsymbol{\nabla}\times$ and curl are defined by $\boldsymbol{\nabla} \times \psi = (\partial_2\psi, -\partial_1\psi)^\top$ for scalar functions ψ and $\operatorname{curl} \boldsymbol{\psi} = \partial_1 \boldsymbol{\psi}_2 - \partial_2 \boldsymbol{\psi}_1$ for vector-valued functions $\boldsymbol{\psi}$. The vector $\boldsymbol{e}_z$ is the unit vector in z-direction and the notation $\boldsymbol{e}_z \times \boldsymbol{\psi}$ for a 2D-vector field is understood by viewing $\boldsymbol{\psi}$ as 2D-vector field with vanishing third component.

We will be using the 2D identities:

$$\begin{aligned} &\boldsymbol{e}_z \times (\boldsymbol{e}_z \times \boldsymbol{E}_t) = -\boldsymbol{E}_t \\ &\boldsymbol{e}_z \times (\boldsymbol{\nabla} \times E_3) = \boldsymbol{\nabla} E_3 \qquad \boldsymbol{e}_z \times \boldsymbol{\nabla} E_3 = -\boldsymbol{\nabla} \times E_3 \\ &\operatorname{curl}(\boldsymbol{e}_z \times \boldsymbol{E}_t) = \operatorname{div} \boldsymbol{E}_t \qquad \operatorname{div}(\boldsymbol{e}_z \times \boldsymbol{E}_t) = -\operatorname{curl} \boldsymbol{E}_t\,. \end{aligned} \tag{4.1}$$

The original system of equations (2.4) translates into:

$$\left\{ \begin{aligned} \boldsymbol{\nabla} \times E_3 + \boldsymbol{e}_z \times \tfrac{\partial}{\partial z}\boldsymbol{E}_t - i\omega \boldsymbol{H}_t &= \boldsymbol{f}_t \\ \operatorname{curl} \boldsymbol{E}_t - i\omega H_3 &= f_3 \\ \boldsymbol{\nabla} \times H_3 + \boldsymbol{e}_z \times \tfrac{\partial}{\partial z}\boldsymbol{H}_t + i\omega \boldsymbol{E}_t &= \boldsymbol{g}_t \\ \operatorname{curl} \boldsymbol{H}_t + i\omega E_3 &= g_3\,. \end{aligned} \right. \tag{4.2}$$

Multiplying $(4.2)_1$ and $(4.2)_3$ by $i\omega\, \boldsymbol{e}_z\times$, we obtain:

$$\left\{ \begin{aligned} \boldsymbol{\nabla} i\omega E_3 - \tfrac{\partial}{\partial z} i\omega \boldsymbol{E}_t + \omega^2\, \boldsymbol{e}_z \times \boldsymbol{H}_t &= i\omega\, \boldsymbol{e}_z \times \boldsymbol{f}_t \\ \operatorname{curl} \boldsymbol{E}_t - i\omega H_3 &= f_3 \\ \boldsymbol{\nabla} i\omega H_3 - \tfrac{\partial}{\partial z} i\omega \boldsymbol{H}_t - \omega^2\, \boldsymbol{e}_z \times \boldsymbol{E}_t &= i\omega\, \boldsymbol{e}_z \times \boldsymbol{g}_t \\ \operatorname{curl} \boldsymbol{H}_t + i\omega E_3 &= g_3\,. \end{aligned} \right. \tag{4.3}$$

The eigensystem corresponding to the first order system operator and $e^{i\beta z}$ ansatz in the variable z

is as follows:

$$
\begin{cases}
\boldsymbol{E}_t \in H_0(\text{curl}, D),\ E_3 \in H_0^1(D) \\
\boldsymbol{H}_t \in H(\text{curl}, D),\ H_3 \in H^1(D) \\
\qquad i\omega \boldsymbol{\nabla} E_3 + \omega^2 \boldsymbol{e}_z \times \boldsymbol{H}_t &= -\omega\beta \boldsymbol{E}_t \\
\qquad \text{curl}\, \boldsymbol{E}_t - i\omega H_3 &= 0 \\
\qquad i\omega \boldsymbol{\nabla} H_3 - \omega^2 \boldsymbol{e}_z \times E_t &= -\omega\beta \boldsymbol{H}_t \\
\qquad \text{curl}\, \boldsymbol{H}_t + i\omega E_3 &= 0\,.
\end{cases}
\tag{4.4}
$$

Eliminating E_3 and H_3 from the system (4.4), we obtain a simplified but second order system for $\boldsymbol{E}_t$, $\boldsymbol{H}_t$ only:

$$
\begin{cases}
\boldsymbol{E}_t \in H_0(\text{curl}, D),\ \text{curl}\, \boldsymbol{E}_t \in H^1(D) \\
\boldsymbol{H}_t \in H(\text{curl}, D),\ \text{curl}\, \boldsymbol{H}_t \in H_0^1(D) \\
\qquad -\boldsymbol{\nabla}(\text{curl}\, \boldsymbol{H}_t) + \omega^2 \boldsymbol{e}_z \times \boldsymbol{H}_t &= -\omega\beta \boldsymbol{E}_t \\
\qquad \boldsymbol{\nabla}(\text{curl}\, \boldsymbol{E}_t) - \omega^2 \boldsymbol{e}_z \times \boldsymbol{E}_t &= -\omega\beta \boldsymbol{H}_t\,.
\end{cases}
\tag{4.5}
$$

4.2 Reduction to single variable eigensystems.

Assume $\beta \neq 0$. Solving $(4.5)_2$ for $\boldsymbol{H}_t$,

$$
\begin{aligned}
\boldsymbol{H}_t &= -\frac{1}{\omega\beta}[\boldsymbol{\nabla}\, \text{curl}\, \boldsymbol{E}_t - \omega^2 \boldsymbol{e}_z \times \boldsymbol{E}_t] \\
\text{curl}\, \boldsymbol{H}_t &= \frac{\omega}{\beta} \text{curl}(\boldsymbol{e}_z \times \boldsymbol{E}_t) = \frac{\omega}{\beta} \text{div}\, \boldsymbol{E}_t
\end{aligned}
\tag{4.6}
$$

and substituting it into $(4.5)_1$, we obtain an eigenvalue problem for $\boldsymbol{E}_t$ alone:

$$
\begin{cases}
\boldsymbol{E}_t \in H_0(\text{curl}, D),\ \text{curl}\, \boldsymbol{E}_t \in H^1(D),\ \text{div}\, \boldsymbol{E}_t \in H_0^1(D) \\
\boldsymbol{\nabla} \times \text{curl}\, \boldsymbol{E}_t - \omega^2 \boldsymbol{E}_t - \boldsymbol{\nabla}\, \text{div}\, \boldsymbol{E}_t = -\beta^2 \boldsymbol{E}_t\,.
\end{cases}
\tag{4.7}
$$

Similarly, solving $(4.5)_1$ for $\boldsymbol{E}_t$,

$$
\begin{aligned}
\boldsymbol{E}_t &= -\frac{1}{\omega\beta}[-\boldsymbol{\nabla}\, \text{curl}\, \boldsymbol{H}_t + \omega^2 \boldsymbol{e}_z \times \boldsymbol{H}_t] \\
\text{curl}\, \boldsymbol{E}_t &= -\frac{\omega}{\beta} \text{curl}(\boldsymbol{e}_z \times \boldsymbol{H}_t) = -\frac{\omega}{\beta} \text{div}\, \boldsymbol{H}_t
\end{aligned}
\tag{4.8}
$$

and substituting it into $(4.5)_2$, we obtain an eigenvalue problem for H_t alone.

$$
\begin{cases}
\boldsymbol{H}_t \in H(\text{curl}, D) \cap H_0(\text{div}, D),\ \text{curl}\, \boldsymbol{H}_t \in H_0^1(D),\ \text{div}\, \boldsymbol{H}_t \in H^1(D) \\
\boldsymbol{\nabla} \times \text{curl}\, \boldsymbol{H}_t - \omega^2 \boldsymbol{H}_t - \boldsymbol{\nabla}(\text{div}\, \boldsymbol{H}_t) = -\beta^2 \boldsymbol{H}_t\,.
\end{cases}
\tag{4.9}
$$

Note that the BC $\boldsymbol{n} \times \boldsymbol{E}_t = 0$ implies the BC $\boldsymbol{n} \cdot \boldsymbol{H}_t = 0$.

4.3 Structure of the single variable eigenproblems

Lemma 7 (Helmholtz decompositions)

Let $D \subset \mathbb{R}^2$ be a simply connected domain. For every $\boldsymbol{E} \in L^2(D)^2$ there exist a unique $\phi \in H^1_0(D)$ and a unique $\psi \in H^1(D), \int_D \psi = 0$, such that

$$\boldsymbol{E} = \boldsymbol{\nabla}\phi + \boldsymbol{\nabla} \times \psi \,. \tag{4.10}$$

Similarly, for every $\boldsymbol{H} \in L^2(D)^2$ there exist a unique $\phi \in H^1_0(D)$ and a unique $\psi \in H^1(D), \int_D \psi = 0$, such that

$$\boldsymbol{H} = \boldsymbol{\nabla} \times \phi + \boldsymbol{\nabla}\psi \,. \tag{4.11}$$

$\blacksquare$

Proof: Standard. $\blacksquare$

Consider now the eigenvalue problem (4.7) and Helmholtz decomposition of $\boldsymbol{E}$. The boundary condition $\boldsymbol{E}_t = 0$ implies that $\frac{\partial\psi}{\partial n} = 0$ on ∂D. Substituting (4.10) into (4.7), we obtain:

$$\boldsymbol{\nabla} \times (\underbrace{-\Delta\psi + (\beta^2 - \omega^2)\psi}_{=:\Psi}) + \boldsymbol{\nabla}(\underbrace{-\Delta\phi + (\beta^2 - \omega^2)\phi}_{=:\Phi}) = 0 \,. \tag{4.12}$$

The equation above represents the Helmholtz decomposition of zero function. Indeed, $\operatorname{div} \boldsymbol{E} = \delta\phi$ is zero on the boundary and, therefore, Φ is zero on the boundary as well. Multiplying (4.12) by $\boldsymbol{\nabla} \times \Psi$ and integrating the second term by parts, we learn that:

$$\Psi = -\Delta\psi + (\beta^2 - \omega^2)\psi = \text{const} \,.$$

Consequently,

$$(\Psi, 1) = (\boldsymbol{\nabla}\psi, \underbrace{\boldsymbol{\nabla}1}_{=0}) - \langle \underbrace{\frac{\partial\psi}{\partial n}}_{=0}, 1\rangle + (\beta^2 - \omega^2)\underbrace{(\psi, 1)}_{=0} = 0 \,.$$

Uniqueness of ϕ and ψ in the Helmholtz decomposition implies now that $\Phi = \Psi = 0$. Let (λ_i, ϕ_i) and $(\mu_j \psi_j)$ be the Dirichlet and Neumann eigenpairs of the Laplacian in domain D. Vanishing of Φ and Ψ implies that there exist i, j such that

$$\phi = \phi_i,\, \omega^2 - \beta^2 = \lambda_i \quad \text{and} \quad \psi = \psi_j, \omega^2 - \beta^2 = \mu_j \,.$$

If the Dirichlet and Neumann eigenvalues are distinct, the eigenvector $\boldsymbol{E}$ must reduce to either gradient or curl. This is the case, e.g., for a circular domain D. In the case of a common Dirichlet and Neumann eigenvalue, $\lambda_i = \mu_j$, we obtain a multiple eigenvalue $\beta^2 = \omega^2 - \lambda_i = \omega^2 - \mu_j$, with the eigenspace consisting of vectors:

$$\boldsymbol{E} = A\boldsymbol{\nabla} \times \psi_j + B\boldsymbol{\nabla}\phi_i \,, \quad A, B \in \mathbb{C} \,.$$

Lemma 8

Let (λ_i, ϕ_i) and $(\mu_j \psi_j)$ denote the Dirichlet and Neumann eigenpairs of the Laplacian in the domain D. The eigenvalues β_i^2 for (4.7) are classified into the following three families.

(a) $\beta^2 = \omega^2 - \mu_j$ *with* μ_j *distinct from all* λ_i. *The corresponding eigenvectors are curls:*

$$\boldsymbol{E} = \boldsymbol{\nabla} \times \psi_j \,,$$

with multiplicity of β^2 *equal to the multiplicity of* μ_j.

(a) $\beta^2 = \omega^2 - \lambda_i$ *with* λ_i *distinct from all* μ_j. *The corresponding eigenvectors are gradients:*

$$\boldsymbol{E} = \boldsymbol{\nabla} \phi_i \,,$$

with multiplicity of β^2 *equal to the multiplicity of* λ_i.

(c) $\beta^2 = \omega^2 - \mu_j = \omega^2 - \lambda_i$ *for* $\mu_j = \lambda_i$. *The corresponding eigenvectors are linear combinations of curls and gradients:*

$$\boldsymbol{E} = A \boldsymbol{\nabla} \times \psi_j + B \boldsymbol{\nabla} \phi_i \,, \quad A, B \in \mathbb{C} \,,$$

with multiplicity of β^2 *equal to the sum of multiplicities of* μ_j *and* λ_i.

∎

In the same way, we prove the analogous result for eigenproblem (4.9).

Lemma 9

Let (λ_i, ϕ_i) and $(\mu_j \psi_j)$ denote the Dirichlet and Neumann eigenpairs of the Laplacian in the domain D. The eigenvalues β_i^2 for (4.9) are classified into the following three families.

(a) $\beta^2 = \omega^2 - \mu_j$ *with* μ_j *distinct from all* λ_i. *The corresponding eigenvectors are gradients:*

$$\boldsymbol{H} = \boldsymbol{\nabla} \psi_j \,,$$

with multiplicity of β^2 *equal to the multiplicity of* μ_j.

(a) $\beta^2 = \omega^2 - \lambda_i$ *with* λ_i *distinct from all* μ_j. *The corresponding eigenvectors are curls:*

$$\boldsymbol{H} = \boldsymbol{\nabla} \times \phi_i \,,$$

with multiplicity of β^2 *equal to the multiplicity of* λ_i.

(c) $\beta^2 = \omega^2 - \mu_j = \omega^2 - \lambda_i$ *for* $\mu_j = \lambda_i$. *The corresponding eigenvectors are linear combinations of gradients and curls:*

$$\boldsymbol{H} = A \boldsymbol{\nabla} \psi_j + B \boldsymbol{\nabla} \times \phi_i \,, \quad A, B \in \mathbb{C} \,,$$

with multiplicity of β^2 *equal to the sum of multiplicities of* μ_j *and* λ_i.

∎

The following example shows that multiple eigenvalues can arise.

Example 1 (Cylindrical waveguide)

Consider the Dirichlet or Neumann Laplace eigenvalue problem in a unit circle,

$$-\Delta u = \lambda u, \quad \lambda = \nu^2 \, .$$

For the Dirichlet problem the operator is positive definite, so $\nu > 0$, for the Neumann problem, $u = \text{const}$ corresponds to zero eigenvalue, all other eigenvalues are positive as well. Rewriting the operator in polar coordinates r, θ,

$$-\frac{1}{r}\frac{\partial}{\partial r}(r\frac{\partial u}{\partial r}) - \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = \nu^2 u$$

and separating variables, $u = R(r)\Theta(\theta)$, we get

$$-\frac{1}{r}(rR')'\Theta - \frac{1}{r^2}R\Theta'' = \nu^2 R\Theta$$

or,

$$\frac{r(rR')'}{R} + \nu^2 r^2 = -\frac{\Theta''}{\Theta} = k^2$$

where k^2 is a real and positive separation constant. We obtain,

$$\Theta = A\cos k\theta + B\sin k\theta$$

and periodic BCs on u and, therefore, Θ, imply that $k = 0, 1, 2, \ldots$. This leads to the Bessel equation in r,

$$r(rR')' + (\nu^2 r^2 - k^2)R = 0$$

with solution:

$$R = CJ_k(\nu r) + DY_k(\nu r) \, .$$

The requirement of finite energy eliminates the second term, $D = 0$.

Dirichlet BC: *$R(1) = 0$ leads to ν being a root of the Bessel function $J_k(\nu) = 0$. We have a family of roots (and, therefore Dirichlet Laplace eigenvalues ν^2) : $\nu = \nu_{k,m}$, $k = 0, 1, 2, \ldots, m = 1, 2, \ldots$. For $k = 0$, the roots are simple, with corresponding eigenvectors given by:*

$$u = J_0(\nu_{0,m} r) \, .$$

For $k > 0$, we have double eigenvectors with eigenspaces given by:

$$u = J_0(\nu_{k,m} r)(A\cos k\theta + B\sin k\theta) \, .$$

Neumann BC: *The situation is similar except that we are dealing now with the roots of the derivative of Bessel functions: $J'_k(\lambda) = 0$, $\lambda = \lambda_{k,m}$, $k = 0, 1, 2, \ldots, m = 1, 2, \ldots$.*

∎

4.4 Stability analysis

In this section, we show

THEOREM 2
There is $C > 0$ independent of L such that the solution $(\boldsymbol{E}, \boldsymbol{H})$ of (2.4) satisfies

$$\|(\boldsymbol{E}, \boldsymbol{H})\|_{L^2(\Omega)} \le CL\|(\boldsymbol{f}, \boldsymbol{g})\|_{L^2(\Omega)}.$$

∎

Let (μ_i, ψ_i) and (λ_j, ϕ_j) denote the Neumann and Dirichlet eigenpairs for the Laplace equation in D. We normalize the eigenvectors as follows:

$$\|\boldsymbol{\nabla}\psi_i\|^2 = \|\boldsymbol{\nabla} \times \psi_i\|^2 = 1, \quad \|\boldsymbol{\nabla}\phi_j\|^2 = \|\boldsymbol{\nabla} \times \phi_j\|^2 = 1 \quad \Rightarrow \quad \|\psi_i\|^2 = \mu_i^{-1}, \quad \|\phi_j\|^2 = \lambda_j^{-1}.$$

Consistently with Lemma 8 and Lemma 9, we make the following ansatz for $\boldsymbol{E}$ and $\boldsymbol{H}$:

$$\begin{aligned} \boldsymbol{E} &= \textstyle\sum_i \boldsymbol{\nabla} \times \psi_i\alpha_i(z) + \sum_j \boldsymbol{\nabla}\phi_j\beta_j(z) + \sum_j \boldsymbol{e}_z\phi_j\gamma_j(z) \\ \boldsymbol{H} &= \textstyle\sum_i \boldsymbol{\nabla}\psi_i\delta_i(z) + \sum_j \boldsymbol{\nabla} \times \phi_j\eta_j(z) + \sum_i \boldsymbol{e}_z\psi_i\zeta_i(z)\,. \end{aligned} \tag{4.13}$$

Substituting into the system (4.2) and multiplying (in the sense of L^2-product) the first equation with $\boldsymbol{\nabla}\psi_i, \boldsymbol{\nabla} \times \phi_j, \boldsymbol{e}_z\psi_i$, and the second equation with $\boldsymbol{\nabla} \times \psi_i, \boldsymbol{\nabla}\phi_j, \boldsymbol{e}_z\phi_j$, we obtain a system of six ODEs:

$$\alpha_i' - i\omega\delta_i = (\boldsymbol{f}, \boldsymbol{\nabla}\psi_i)_{L^2(D)} \quad =: f_1, \tag{4.14a}$$
$$-\beta_j' + \gamma_j - i\omega\eta_j = (\boldsymbol{f}, \boldsymbol{\nabla} \times \phi_j)_{L^2(D)} \quad =: f_2, \tag{4.14b}$$
$$\alpha_i - i\omega\mu_i^{-1}\zeta_i = (\boldsymbol{f}, \boldsymbol{e}_z\psi_i)_{L^2(D)} \quad =: f_3, \tag{4.14c}$$
$$-\delta_i' + \zeta_i + i\omega\alpha_i = (\boldsymbol{g}, \boldsymbol{\nabla} \times \psi_i)_{L^2(D)} \quad =: g_1, \tag{4.14d}$$
$$\eta_j' + i\omega\beta_j = (\boldsymbol{g}, \boldsymbol{\nabla}\phi_j)_{L^2(D)} \quad =: g_2, \tag{4.14e}$$
$$\eta_j + i\omega\lambda_j^{-1}\gamma_j = (\boldsymbol{g}, \boldsymbol{e}_z\phi_j)_{L^2(D)} \quad =: g_3. \tag{4.14f}$$

Note that the system decouples into two subsystems consisting of two first-order ODEs and one algebraic equation. The ODE system is complemented with boundary conditions. The condition (2.4b) requires the conditions

$$\alpha_i(0) = 0 \quad \forall i \in \mathbb{N}, \tag{4.15a}$$
$$\beta_j(0) = 0 \quad \forall j \in \mathbb{N}. \tag{4.15b}$$

The condition (2.4d) that waves be outgoing takes the following form. Assuming $\boldsymbol{f} = 0$ and $\boldsymbol{g} = 0$ on (L, ∞), one obtains for α_i and β_j the second-order ODEs (see the proofs of Lemmas 10, 11 for some more details)

$$-\alpha_i'' - \widetilde{\mu}_i^2\alpha_i = 0, \qquad \widetilde{\mu}_i^2 = \mu_i - \omega^2, \tag{4.16}$$
$$-\beta_j'' - \widetilde{\lambda}_j^2\beta_j = 0, \qquad \widetilde{\lambda}_j^2 = \lambda_j - \omega^2 \tag{4.17}$$

with fundamental solutions $e^{\pm\widetilde{\mu}_i z}$ and $e^{\pm\widetilde{\lambda}_j z}$. For solutions to be outgoing, we select the minus sign and realize this with the boundary conditions at $z = L$ given by

$$\alpha_i'(L) + \widetilde{\mu}_i \alpha_i(L) = 0, \tag{4.18a}$$

$$\beta_j'(L) + \widetilde{\lambda}_j \beta_j(L) = 0. \tag{4.18b}$$

REMARK 1 From (4.14) with $\boldsymbol{f} = 0 = \boldsymbol{g}$, one can infer alternative conditions, e.g., $i\omega\delta_i(L) = -\widetilde{\mu}_i\alpha(L)$ and $\widetilde{\lambda}_j\eta_j(L) = i\omega\beta_j(L)$. ▮

Representing $\boldsymbol{E}$ using the ansatz $(4.13)_1$, we obtain:

$$\int_D |\boldsymbol{E}(x,y,z)|^2 \, \mathrm{d}x \, \mathrm{d}y = \sum_i \left[|\alpha_i(z)|^2 + |\beta_i(z)|^2 + \lambda_i^{-1} |\gamma_i(z)|^2 \right] .$$

Consequently,

$$\|\boldsymbol{E}\|^2 = \int_0^L \int_D |E(x,y,z)|^2 \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \sum_i \|(\alpha_i, \beta_i, \lambda_i^{-\frac{1}{2}} \gamma_i)\|^2_{L^2(0,L)} .$$

Similarly,

$$\|\boldsymbol{H}\|^2 = \sum_i \|(\delta_i, \eta_i, \mu_i^{-\frac{1}{2}} \zeta_i)\|^2_{L^2(0,L)} .$$

Now, let $\boldsymbol{f}$ denote the right-hand side in the first equation. Using the mutual orthogonality of $\boldsymbol{\nabla}\psi_i$, $\boldsymbol{\nabla} \times \phi_i$, $\boldsymbol{e}_z\psi_i$ in $L^2(D)$, we can represent f in the form:

$$\boldsymbol{f} = \sum_i \left[(\boldsymbol{f}, \boldsymbol{\nabla}\psi_i)_{L^2(D)} \boldsymbol{\nabla}\psi_i + (\boldsymbol{f}, \boldsymbol{\nabla} \times \phi_i)_{L^2(D)} \boldsymbol{\nabla} \times \phi_i + (\boldsymbol{f}, \boldsymbol{e}_z \mu_i^{\frac{1}{2}} \psi_i)_{L^2(D)} \boldsymbol{e}_z \mu_i^{\frac{1}{2}} \phi_i \right]$$

Consequently,

$$\int_D |\boldsymbol{f}(x,y,z)|^2 \, \mathrm{d}x \, \mathrm{d}y = \sum_i \left[|(\boldsymbol{f}, \boldsymbol{\nabla}\psi_i)_{L^2(D)}(z)|^2 + |(\boldsymbol{f}, \boldsymbol{\nabla} \times \phi_i)_{L^2(D)}(z)|^2 + |(\boldsymbol{f}, \boldsymbol{e}_z\psi_i)_{L^2(D)}(z)|^2 \mu_i \right]$$

and

$$\|\boldsymbol{f}\|^2 = \int_0^L \int_D |\boldsymbol{f}(x,y,z)|^2 \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \sum_i \|((\boldsymbol{f}, \boldsymbol{\nabla}\psi_i), (\boldsymbol{f}, \boldsymbol{\nabla} \times \phi_i), \mu_i^{\frac{1}{2}} (\boldsymbol{f}, \boldsymbol{e}_z\psi_i))\|^2_{L^2(0,L)} .$$

Similarly, for the right-hand side $\boldsymbol{g}$ in the second equation in (2.4) we have

$$\|\boldsymbol{g}\|^2 = \sum_i \|((\boldsymbol{g}, \boldsymbol{\nabla} \times \psi_i), (\boldsymbol{g}, \boldsymbol{\nabla}\phi_i), \lambda_i^{\frac{1}{2}} (\boldsymbol{g}, \boldsymbol{e}_z\phi_i))\|^2_{L^2(0,L)} .$$

Now, the formulas for $\|\boldsymbol{E}\|^2, \|\boldsymbol{H}\|^2$ and $\|\boldsymbol{f}\|^2, \|\boldsymbol{g}\|^2$, and the system of ODEs (4.14) imply that sufficient (and necessary) for the boundedness below are the L^2-estimates for the subsystems of ODEs:

$$\|(\alpha_i, \delta_i, \mu_i^{-\frac{1}{2}} \zeta_i)\|^2_{L^2(0,L)} \le C \|((\boldsymbol{f}, \boldsymbol{\nabla}\psi_i), (\boldsymbol{g}, \boldsymbol{\nabla} \times \psi_i), (\boldsymbol{f}, \boldsymbol{e}_z \mu_i^{\frac{1}{2}} \psi_i))\|^2_{L^2(0,L)}, \tag{4.19a}$$

$$\|(\beta_i, \eta_i, \lambda_i^{-\frac{1}{2}} \gamma_i)\|^2_{L^2(0,L)} \le C \|((\boldsymbol{f}, \boldsymbol{\nabla} \times \psi_i), (\boldsymbol{g}, \boldsymbol{\nabla}\psi_i), (\boldsymbol{g}, \boldsymbol{e}_z \lambda_i^{\frac{1}{2}} \phi_i))\|^2_{L^2(0,L)} \tag{4.19b}$$

with some constant C independent of i. We have from Lemma 4:

Lemma 10

Let $(\alpha_i, \delta_i, \zeta_i)$ solve (4.14a), (4.14c), (4.14d) together with the boundary conditions (4.15a), (4.18a). Assume that $\inf_i |\widetilde{\mu}_i| > 0$*. Then the following holds:*

(i) (evanescent modes) There is $C > 0$ independent of i such that for all i with $\widetilde{\mu}_i > 0$

$$\|\alpha_i'\| + \sqrt{\mu_i}\|\alpha_i\| \leq C\left[\|f_1\| + \sqrt{\mu_i}\|f_3\| + \|g_1\|\right], \tag{4.20}$$

$$\|\delta_i\| \leq C\left[\|f_1\| + \sqrt{\mu_i}\|f_3\| + \|g_1\|\right], \tag{4.21}$$

$$\mu_i^{-1/2}\|\zeta_i\| \leq C\left[\|f_1\| + \sqrt{\mu_i}\|f_3\| + \|g_1\|\right]. \tag{4.22}$$

(ii) (propagating modes) For all i with $\widetilde{\mu}_i \in i\mathbb{R}$ the estimates (4.20)–(4.22) hold with μ_i replaced by 1 and an additional factor L on the right-hand side.

∎

Proof: The component α_i satisfies the following weak form: For all $v \in H^1_{(0}(0, L)$ there holds

$$(\alpha_i', v')_{L^2(0,L)} + \widetilde{\mu}_i^2(\alpha_i, v)_{L^2(0,L)} + \widetilde{\mu}_i\alpha_i(L)\overline{v}(L) = (f_1, v')_{L^2(0,L)} + i\omega(g_1, v)_{L^2(0,L)} + \mu_i(f_3, v)_{L^2(0,L)}. \tag{4.23}$$

This is obtained by the following steps (see Appendix B.1) for details): first one eliminates the variable ζ_i; second, one multiplies (4.14a) (in the form obtained after removing ζ_i) by v' and integrates over $(0, L)$; third, one multiplies (4.14d) by v and integrates over $(0, L)$; fourth, in the thus obtained equation the term $(\delta', v)_{L^2(0,L)}$ is integrated by parts and the condition $i\omega\delta_i(L) = -\widetilde{\mu}_i\alpha_i(L)$ from Remark 1 is used.

Proof of (i): We note $\widetilde{\mu}_i \sim \sqrt{\mu_i}$ with implied constant independent of i. Lemma 4 together with the condition $\alpha_i(0) = 0$ from (4.15a) then readily implies (4.20). Combining (4.20) with (4.14a) provides (4.21). Finally, (4.20) and (4.14c) yield (4.22).

We remark that the estimate for α_i is a better by a factor $\sqrt{\mu_i}$ than required.

Proof of (ii): This case is shown in the same way as (i) noting that for the finitely many propagating modes one has $\mu_i \sim 1$ so that $|\widetilde{\mu}_i| \sim 1$. ∎

Analogously, we have for the subsystem involving $(\beta_j, \eta_j, \gamma_j)$:

Lemma 11

Let $(\beta_j, \eta_j, \gamma_j)$ solve (4.14b), (4.14e), (4.14f) together with the boundary conditions (4.15b), (4.18b). Assume that $\inf_j |\widetilde{\lambda}_j| > 0$*. Then the following holds:*

(i) (evanescent modes) There is $C > 0$ independent of j such that for all j with $\widetilde{\lambda}_j > 0$

$$\|\beta_j'\| + \sqrt{\lambda_i}\|\beta_j\| \leq C\left[\|f_2\| + \lambda_j\|g_3\| + \sqrt{\lambda_j}\|g_2\|\right], \tag{4.24}$$

$$\lambda_j\|\eta_j\| \leq C\left[\|f_2\| + \lambda_j\|g_3\| + \sqrt{\lambda_j}\|g_1\|\right], \tag{4.25}$$

$$\lambda_j^{-1/2}\|\gamma_j\| \leq C\left[\lambda_j^{-1/2}\|f_2\| + \sqrt{\lambda_j}\|g_3\| + \|g_2\|\right]. \tag{4.26}$$

(ii) (propagating modes) For all j with $\widetilde{\lambda}_j \in i\mathbb{R}$ the estimates (4.24)–(4.26) hold with λ_j replaced by 1 and an additional factor L on the right-hand side.

∎

Proof: The component β_j satisfies the following weak form: For all $v \in H^1_{(0}(0,L)$ there holds

$$(\beta_j', v')_{L^2(0,L)} + \widetilde{\lambda}_j^2(\beta_j, v)_{L^2(0,L)} + \widetilde{\lambda}_j\beta_j(L)\overline{v}(L) = -(f_2, v')_{L^2(0,L)} + \frac{\lambda_j}{i\omega}(g_3, v')_{L^2(0,L)} + \frac{\widetilde{\lambda}_j^2}{i\omega}(g_2, v)_{L^2(0,L)}. \tag{4.27}$$

This is obtained by the following steps (see Appendix B.2) for details): first one eliminates the variable γ_j, which yields the additional equation

$$-i\omega\beta_j' - \widetilde{\lambda}_j^2\eta_j = i\omega f_2 - \lambda_j g_3. \tag{4.28}$$

Second, one multiplies (4.28) by v' and integrates over $(0,L)$; third, one multiplies (4.14e) by v and integrates over $(0,L)$; fourth, in the thus obtained equation the term $(\eta_j', v)_{L^2(0,L)}$ is integrated by parts and the condition $i\omega\widetilde{\lambda}_j\beta_j(L) = -\widetilde{\lambda}_j^2\beta_j(L)$ from Remark 1 is used.

Proof of (i): We note $\widetilde{\lambda}_j \sim \sqrt{\lambda_j}$ with implied constant independent of j. Lemma 4 together with the condition $\beta_j(0) = 0$ from (4.15b) then readily implies (4.24). Combining (4.24) with (4.28) provides (4.25). Finally, (4.24) and (4.14f) yield (4.22).

Proof of (ii): This case is shown in the same way as (i) noting that for the finitely many propagating modes one has $\lambda_j \sim 1$ so that $|\widetilde{\lambda}_j| \sim 1$. ∎

Proof of Theorem 2: The stability bound follows from Lemmas 10 and 11 together with the observation that the estimates (4.19) imply Theorem 2. □

5 Numerical results and conclusions

To test the dependence of the ultraweak DPG Maxwell discretization on the waveguide length L and the scaling constant β in the test norm, we solve the linear time-harmonic Maxwell equations in a homogeneous rectangular waveguide with transverse domain $D = (0.0, 1.0) \times (0.0, 0.5)$. At Γ_{in}, the waveguide is excited with the lowest-order transverse electric (TE) mode by prescribing the corresponding tangential electric field. The cross-section is modeled with two hexahedral elements, which is justified by the simple transverse mode profile. In the longitudinal direction, we use a "fixed discretization" of two elements per wavelength; that is, as the waveguide length L increases, the number of elements per wavelength remains the same. The DPG discretization uses uniform polynomial order $p = 5$ with enriched test functions of order $p + 1$. All of the computations were done in *hp*3D [18, 17].

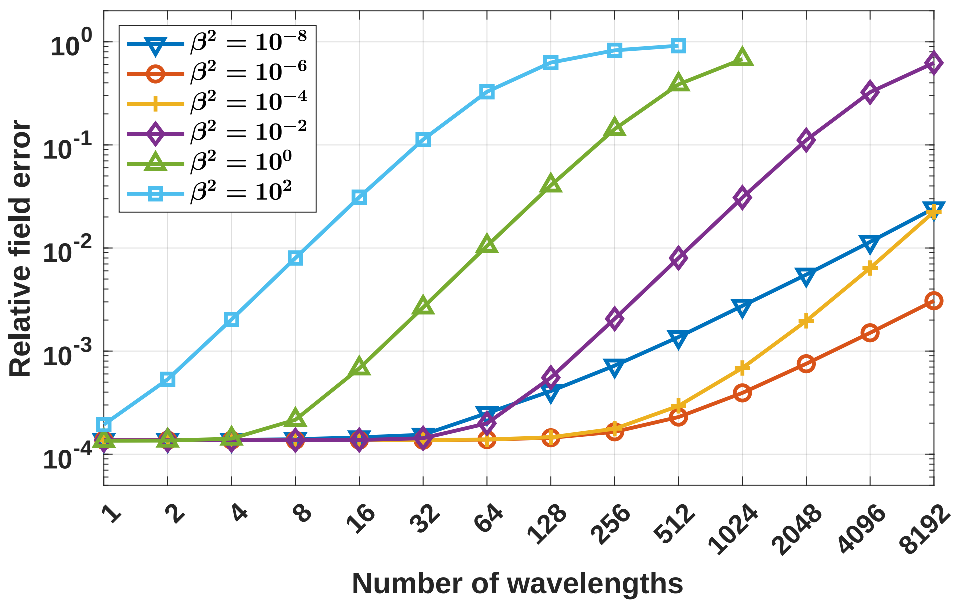

Figure 1: Relative field error of the ultraweak DPG Maxwell solution for the fundamental TE mode in a homogeneous 3D rectangular waveguide. The discretization is uniformly fifth-order with a fixed number of elements per wavelength. Choosing a small scaling constant β in the DPG test norm improves the stability of the method significantly, but choosing β too small results in round-off errors that adversely affect the solution.

We recall the scaled adjoint graph norm from Section 2:

$$\|v\|^2_{H_{A^*}(\Omega_h)} = \|A^*_h v\|^2 + \beta^2 \|v\|^2.$$

The analysis of Lemma 1 and Theorem 2 suggests scaling β with L^{-1} to maintain stability of the method as the waveguide length L increases. Figure 1 shows the relative L^2-error of the electric field $\boldsymbol{E}$ for various choices of β and increasing waveguide length. The numerical results confirm that the choice of β indeed significantly affects the stability of the discretization. In practice, of course, scaling β is limited due to rounding errors as β becomes very small (or very large). In our experiments, the best results were obtained for $\beta^2 = 10^{-6}$. In other words, by choosing β sufficiently small, ultraweak DPG can compensate (to a certain extent) the loss of stability due to the L^{-1} dependence of the boundedness-below constant.

Conclusions. Non-homogeneous waveguide. The convergence analysis for the DPG method for the modified Maxwell model resulting from a full envelope ansatz led us to the stability analysis for EM waveguides. We started with a simpler, acoustic waveguide, and then proceeded to the homogeneous Maxwell model. For both models, we have obtained the same result: the stability constant depends linearly on the length of the waveguide, i.e.,

$$\|\boldsymbol{E}\| + \|\boldsymbol{H}\| \le CL\left(\|\boldsymbol{\nabla} \times \boldsymbol{E} - i\omega \boldsymbol{H}\| + \|\boldsymbol{\nabla} \times \boldsymbol{H} + i\omega \boldsymbol{E}\|\right), \qquad C > 0\,.$$

For the non-homogeneous optical waveguide, $\mu = 1$, $\epsilon = 1 + \delta\epsilon$, the triangle inequality implies:

$$\begin{aligned}\|\boldsymbol{E}\| + \|\boldsymbol{H}\| &\le CL\left(\|\boldsymbol{\nabla} \times \boldsymbol{E} - i\omega \boldsymbol{H}\| + \|\boldsymbol{\nabla} \times \boldsymbol{H} + i\omega(1 + \delta\epsilon - \delta\epsilon)\boldsymbol{E}\|\right)\\ &\le CL\left(\|\boldsymbol{\nabla} \times \boldsymbol{E} - i\omega \boldsymbol{H}\| + \|\boldsymbol{\nabla} \times \boldsymbol{H} + i\omega(1 + \delta\epsilon)\boldsymbol{E}\| + \omega\|\delta\epsilon\|_{L^\infty}\|\boldsymbol{E}\|\right).\end{aligned}$$

Consequently,

$$(1 - CL\omega\|\delta\epsilon\|_{L^\infty}\omega)(\|\boldsymbol{E}\| + \|\boldsymbol{H}\|) \le CL\left(\|\boldsymbol{\nabla} \times \boldsymbol{E} - i\omega \boldsymbol{H}\| + \|\boldsymbol{\nabla} \times \boldsymbol{H} + i\omega(1 + \delta\epsilon)\boldsymbol{E}\|\right),$$

which proves that, for a sufficiently small perturbation $\delta\epsilon$, the operator remains bounded below. However, the information about the linear dependence of the stability constant upon L is lost.[1] In fact, the larger L, the smaller $\delta\epsilon$ is allowed. In the second part of this work [6], we generalize our results to the case of non-homogeneous waveguides using perturbation theory.

References

[1] G. Agrawal. *Nonlinear fiber optics*. 5th ed. Academic Press, 2012.

[2] I. Babuška and S. Sauter. "Is the Pollution Effect of the FEM Avoidable for the Helmholtz's Equation Considering High Wave Numbers?" In: *SIAM J. Numer. Anal* 6 (1997), pp. 2392–2423.

[3] C. Carstensen, L. Demkowicz, and J. Gopalakrishnan. "Breaking spaces and forms for the DPG method and applications including Maxwell equations". In: *Comput. Math. Appl.* 72.3 (2016), pp. 494–522.

[4] L. Demkowicz and J. Gopalakrishnan. "Encyclopedia of Computational Mechanics, Second Edition". In: Eds. Erwin Stein, René de Borst, Thomas J. R. Hughes, see also ICES Report 2015/20. Wiley, 2018. Chap. Discontinuous Petrov-Galerkin (DPG) Method.

[5] L. Demkowicz, J. Gopalakrishnan, I. Muga, and J. Zitelli. "Wavenumber Explicit Analysis for a DPG Method for the Multidimensional Helmholtz Equation". In: *Comput. Methods Appl. Mech. Engrg.* 213-216 (2012), pp. 126–138.

[6] L. Demkowicz, M. Melenk, S. Henneking, and J. Badger. *Stability Analysis for Acoustic and Electromagnetic Waveguides. Part 2: Non-homogeneous Waveguides.* Tech. rep. 3. The University of Texas at Austin, Austin, TX 78712: Oden Institute, 2023.

[7] T. Eidam et al. "Experimental observations of the threshold-like onset of mode instabilities in high power fiber amplifiers". In: *Opt. Express* 19.14 (2011), pp. 13218–13224.

[8] O. Ernst and M. J. Gander. "Why it is difficult to solve Helmholtz problems with classical iterative methods". In: *Numerical Analysis of Multiscale Problems. Lect. Notes Comput. Sci. Eng.* Vol. 83. Springer, 2012, pp. 325–363.

[1] $CL/(1 - CL\omega\,\delta\epsilon) \approx CL(1 + CL\omega\,\delta\epsilon)$.

[9] M. J. Gander, I. G. Graham, and E. A. Spence. "Applying GMRES to the Helmholtz equation with shifted Laplacian preconditioning: what is the largest shift for which wavenumber-independent convergence is guaranteed?" In: *Numer. Math.* 131.3 (2015), pp. 567–614.

[10] T. Goswami, J. Grosek, and J. Gopalakrishnan. "Simulations of single-and two-tone Tm-doped optical fiber laser amplifiers". In: *Opt. Express* 29.8 (2021), pp. 12599–12615.

[11] D. Griffiths. *Introduction to electrodynamics*. 3rd ed. Prentice Hall, 1999.

[12] S. Henneking and L. Demkowicz. "A numerical study of the pollution error and DPG adaptivity for long waveguide simulations". In: *Comp. and Math. Appl.* 95 (2021), pp. 85–100.

[13] S. Henneking, J. Grosek, and L. Demkowicz. "Model and computational advancements to full vectorial Maxwell model for studying fiber amplifiers". In: *Comp. and Math. Appl.* 85 (2021), pp. 30–41.

[14] S. Henneking, J. Grosek, and L. Demkowicz. "Parallel simulations of high-power optical fiber amplifiers". In: *Lect. Notes Comput. Sci. Eng. (accepted)* (2022).

[15] S. Henneking. "A scalable *hp*-adaptive finite element software with applications in fiber optics". PhD thesis. The University of Texas at Austin, 2021.

[16] S. Henneking, J. Badger, J. Grosek, and L. Demkowicz. *Numerical simulation of transverse mode instability with a full envelope DPG Maxwell model*. In preparation. 2023.

[17] S. Henneking and L. Demkowicz. *Computing with hp Finite Elements. III. Parallel hp3D Code*. In preparation, 2023.

[18] S. Henneking and L. Demkowicz. "*hp*3D User Manual". In: *arXiv preprint arXiv:2207.12211* (2022).

[19] J. Jackson. *Classical electrodynamics*. 3rd ed. John Wiley & Sons, 1999.

[20] C. Jauregui, J. Limpert, and A. Tünnermann. "High-power fibre lasers". In: *Nat. Photonics* 7.11 (2013), p. 861.

[21] C. Jauregui, C. Stihler, and J. Limpert. "Transverse mode instability". In: *Adv. Opt. Photonics* 12.2 (2020), pp. 429–484.

[22] K. Kawano and T. Kitoh. *Introduction to Optical Waveguide Analysis: Solving Maxwell's Equation and the Schrödinger Equation*. John Wiley & Sons, 2004.

[23] A. Kobyakov, M. Sauer, and D. Chowdhury. "Stimulated Brillouin scattering in optical fibers". In: *Adv. Opt. Photonics* 2.1 (2010), pp. 1–59.

[24] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000, pp. xiv+357.

[25] J. M. Melenk. "On Generalized Finite Element Methods". PhD thesis. University of Maryland at College Park, 1995.

[26] J. M. Melenk, S. A. Sauter, and C. Torres. *Wavenumber explicit analysis for Galerkin discretizations of lossy Helmholtz problems*. arXiv:1904.00207v1. 2019.

[27] J. M. Melenk, S. A. Sauter, and C. Torres. "Wavenumber explicit analysis for Galerkin discretizations of lossy Helmholtz problems". In: *SIAM J. Numer. Anal.* 58.4 (2020), pp. 2119–2143.

[28] J. M. Melenk and S. A. Sauter. "Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation". In: *SIAM J. Numer. Anal.* 49.3 (2011), pp. 1210–1243.

[29] J. M. Melenk and S. A. Sauter. “Wavenumber-explicit hp-FEM analysis for Maxwell’s equations with transparent boundary conditions”. In: *Found. Comput. Math.* (2020), pp. 1–117.

[30] S. Naderi, I. Dajani, T. Madden, and C. Robin. “Investigations of modal instabilities in fiber amplifiers through detailed numerical simulations”. In: *Opt. Express* 21.13 (2013), pp. 16111–16129.

[31] S. Nagaraj et al. “A 3D DPG Maxwell approach to nonlinear Raman gain in fiber laser amplifiers”. In: *J. Comp. Phys. X 2* (2019), p. 100002.

[32] K. Saitoh and M. Koshiba. “Full-vectorial finite element beam propagation method with perfectly matched layers for anisotropic optical waveguides”. In: *J. Light. Technol.* 19.3 (2001), pp. 405–413.

[33] R. Smith. “Optical power handling capacity of low loss optical fibers as determined by stimulated Raman and Brillouin scattering”. In: *Appl. Opt.* 11.11 (1972), pp. 2489–2494.

[34] A. W. Snyder, J. D. Love, et al. *Optical waveguide theory.* Vol. 175. Chapman and hall London, 1983.

[35] B. Ward. “Modeling of transient modal instability in fiber amplifiers”. In: *Opt. Express* 21.10 (2013), pp. 12053–12067.

A Proof of Lemma 3

For the convenience of the reader and to clarify some arguments from [27] we provide some details of the proof of Lemma 3. To simply the notation, we write in the present section

$$a_\kappa^{1D}(u,v) = (u',v')_{L^2(0,1)} + \kappa^2 (u,v)_{L^2(0,1)} + \kappa u(1)\overline{v}(1),$$
$$\|u\|^2_{1,|\kappa|} := \|u'\|^2_{L^2(0,1)} + |\kappa|^2 \|u\|^2_{L^2(0,1)}$$

Lemma 12 ([27, Lemma 4.1])
For any $\kappa \in \mathbb{C}_{\geq 0} \setminus \{0\}$, the sesquilinear form a_κ^{1D} satisfies

$$\inf_{u\in V\setminus\{0\}} \sup_{v\in V\setminus\{0\}} \frac{\operatorname{Re} a_\kappa^{1D}(u,v)}{\|u\|_{1,|\kappa|}\|v\|_{1,|\kappa|}} \geq \frac{\operatorname{Re}\kappa}{|\kappa|}.$$

▮

Proof: Given $u \in V$ take $v := \frac{\kappa}{|\kappa|} u \in V$ and compute

$$\operatorname{Re} a_\kappa^{1D}(u,v) = \operatorname{Re}\left(\frac{\overline{\kappa}}{|\kappa|}\|u'\|^2_{L^2(0,1)} + \kappa|\kappa|\|u\|^2_{L^2(0,1)} + |\kappa||u(1)|^2 \right) \geq \frac{\operatorname{Re}\kappa}{|\kappa|}\|u\|^2_{1,|\kappa|} = \frac{\operatorname{Re}\kappa}{|\kappa|}\|u\|_{1,|\kappa|}\|v\|_{1,|\kappa|}.$$

▮

The following is a corrected version of [27, Lemma 4.2] although the statement of [27, Lemma 4.2] is correct (cf. Cor. 1 below).

Lemma 13 ([27, Lemma 4.2])
Let $V = H^1(0,1)$ or $V = H^1_{(0}(0,1)$. There is $C > 0$ such that for all $\kappa \in \mathbb{C}_{\geq 0}$ with $|\kappa| \geq 1$ and all $f \in L^2(0,1)$, $g \in \mathbb{C}$ the solution $u \in V$ of

$$a_\kappa^{1D}(u,v) = (f,v)_{L^2(0,1)} + g\overline{v}(1) \quad \forall v \in V$$

satisfies the stability bound

$$\|u\|_{1,|\kappa|} \leq C \left[\|f\|_{L^2(0,1)} + |g|\right].$$

$\blacksquare$

Proof: This is taken from [27, Lemma 4.2], where the multidimensional case is covered. We repeat the arguments, which are based on the multiplier technique worked out in detail in [26, Lemma 4.2], [9, Cor. 2.11], [25, Prop. 8.1.3].

Case a: For a parameter $\beta > 0$, consider the case $|\operatorname{Im}\kappa| \leq \beta \operatorname{Re}\kappa$. In view of $|\kappa| \geq 1$, this implies

$$\operatorname{Re}\kappa \geq \frac{1}{\sqrt{1+\beta^2}}|\kappa| \geq \frac{1}{\sqrt{1+\beta^2}}.$$

By the inf-sup condition satisfied by a_κ^{1D} in view of Lemma 12 so that the solution u satisfies by the variational formulation

$$\|u\|^2_{1,|\kappa|} \leq C_\beta \left[\|f\|_{L^2(0,1)}\|u\|_{L^2(0,1)} + |g||u(1)|\right] \leq C_\beta \left[\frac{1}{|\kappa|}\|f\|_{L^2(0,1)} + \frac{1}{\sqrt{|\kappa|}}|g|\right] \|u\|_{1,|\kappa|}.$$

In view of $\operatorname{Re}\kappa \geq 1/\sqrt{(1+\beta^2)}$ and $|\kappa| \geq \operatorname{Re}\kappa$, this is a stronger result than needed (cf. Cor. 1 ahead).

Case b: We consider the case $|\operatorname{Im}\kappa| \geq \beta \operatorname{Re}\kappa$ so that

$$|\operatorname{Im}\kappa| \leq |\kappa| \leq \left(1+\beta^{-1}\right)|\operatorname{Im}\kappa|.$$

Taking as the test function $v = u$ gives by taking the real part

$$\|u'\|^2_{L^2(0,1)} + \operatorname{Re}\kappa^2\|u\|^2_{L^2(0,1)} + \operatorname{Re}\kappa|u(1)|^2 \lesssim \|f\|_{L^2(0,1)}\|u\|_{L^2(0,1)} + |g||u(1)|. \tag{A.1}$$

Taking as the test function $v = (\operatorname{Im}\kappa)u$ gives by taking the imaginary part

$$2\underbrace{|\operatorname{Im}\kappa|\operatorname{Re}\kappa}_{\geq 0}\|u\|^2_{L^2(0,1)} + |\operatorname{Im}\kappa||u(1)|^2 \lesssim \|f\|_{L^2(0,1)}\|u\|_{L^2(0,1)} + |g||u(1)|. \tag{A.2}$$

The multiplier technique amounts to taking as a test function $v = xu'$. Using the relations

$$\operatorname{Re} u'(x\overline{u}')' = \frac{1}{2}x\left(|u'|^2\right)' + |u'|^2, \qquad \operatorname{Re} ux\overline{u}' = \frac{1}{2}x\left(|u|^2\right)', \qquad \operatorname{Im} ux\overline{u} = 0$$

we get for the test function $v = xu'$

$$\begin{aligned}\operatorname{Re} a_\kappa^{1D}(u,xu') &= \operatorname{Re}(u',(xu')')_{L^2(0,1)} + \operatorname{Re}\kappa^2(u,xu)_{L^2(0,1)} + \operatorname{Re}\kappa u(1)\overline{u}'(1)\\ &= \|u'\|^2_{L^2(0,1)} - \frac{1}{2}\|u'\|^2_{L^2(0,1)} + \frac{1}{2}|u'(1)|^2 - \operatorname{Re}\kappa^2\frac{1}{2}\|u\|^2_{L^2(0,1)} + \operatorname{Re}\kappa^2\frac{1}{2}|u(1)|^2 + \operatorname{Re}\kappa u(1)\overline{u}'(1),\end{aligned}$$

so that we arrive from the weak formulation of u at

$$\|u'\|^2_{L^2(0,1)} - \operatorname{Re}\kappa^2\|u\|^2_{L^2(0,1)} + |u'(1)|^2 \leq 2\left(\|f\|_{L^2(0,1)}\|xu'\|_{L^2(0,1)} + |g||u'(1)| + |\operatorname{Re}\kappa||u(1)||u'(1)|\right)$$

We note that for sufficiently large β we have $\operatorname{Re}\kappa^2 = (\operatorname{Re}\kappa)^2 - (\operatorname{Im}\kappa)^2 \sim -|\kappa|^2$. Hence, appropriate Young inequalities yield

$$\|u'\|^2_{L^2(0,1)}|\kappa|^2\|u\|^2_{L^2(0,1)} + |u'(1)|^2 \lesssim \|f\|^2_{L^2(0,1)} + |g|^2 + (\operatorname{Re}\kappa)^2|u(1)|^2.$$

Young's inequality applied to (A.2) allows us to remove the term $(\operatorname{Re}\kappa)^2|u(1)|^2$. ▮

We are now in position to prove Lemma 3:

Proof of Lemma 3: We follow [27, Thm. 4.3]. For $\operatorname{Re}\kappa \geq 1/\sqrt{2}$ we have by Lemma 12

$$\inf_{u\in V}\sup_{v\in V}\frac{\operatorname{Re}a_\kappa(u,v)}{\|u\|_{1,|\kappa|}\|v\|_{1,|\kappa|}} \geq \frac{\operatorname{Re}\kappa}{|\kappa|} = \frac{\operatorname{Re}\kappa}{\sqrt{(\operatorname{Re}\kappa)^2 + (\operatorname{Im}\kappa)^2}} \geq \frac{\operatorname{Re}\kappa}{\operatorname{Re}\kappa + |\operatorname{Im}\kappa|} \geq \frac{1}{1 + c\frac{|\operatorname{Im}\kappa|}{1+\operatorname{Re}\kappa}}, \qquad c = 1+\sqrt{2}.$$

Let $\operatorname{Re}\kappa \leq 1/\sqrt{2}$. Given $u \in V$, we seek v in the form $v = u + z$ with z solving

$$a_{\overline{\kappa}}(z,w) = \alpha^2(u,w)_{L^2(0,1)} \quad \forall w \in V, \qquad \alpha^2 := |\kappa|^2 - \kappa^2 = -2i(\operatorname{Im}\kappa)\kappa.$$

By Lemma 13, we get

$$\|z\|_{1,|\kappa|} \leq C|\alpha|^2\|u\|_{L^2(0,1)} \leq C|\operatorname{Im}\kappa||\kappa|\|u\|_{L^2(0,1)}$$

and note

$$\begin{aligned}
\operatorname{Re}a_\kappa(u,u+z) &= \|u'\|^2_{L^2(0,1)} + \operatorname{Re}\kappa^2\|u\|^2_{L^2(0,1)} + \operatorname{Re}\kappa|u(1)|^2 + \operatorname{Re}a_\kappa(u,z)\\
&= \|u'\|^2_{L^2(0,1)} + \operatorname{Re}\kappa^2\|u\|^2_{L^2(0,1)} + \operatorname{Re}\kappa|u(1)|^2 + \operatorname{Re}a_{\overline{\kappa}}(z,u)\\
&= \|u'\|^2_{L^2(0,1)} + \operatorname{Re}\kappa^2\|u\|^2_{L^2(0,1)} + \operatorname{Re}\kappa|u(1)|^2 + \operatorname{Re}\alpha^2\|u\|^2_{L^2(0,1)}\\
&= \|u\|^2_{1,|\kappa|} + \underbrace{\operatorname{Re}\kappa}_{\geq 0}|u(1)|^2 \geq \|u\|^2_{1,|\kappa|}
\end{aligned}$$

as well as

$$\|u+z\|_{1,|\kappa|} \leq \|u\|_{1,|\kappa|} + \|z\|_{1,|\kappa|} \leq (1 + C|\operatorname{Im}\kappa|)|\kappa|\|u\|_{L^2(0,1)} \leq (1 + C|\operatorname{Im}\kappa|)\|u\|_{1,|\kappa|} \tag{A.3}$$

so that we arrive at

$$\operatorname{Re}a^{1D}_\kappa(u,v) = \operatorname{Re}a^{1D}_\kappa(u,u+z) = \|u\|^2_{1,|\kappa|} \geq \frac{\|u\|_{1,|\kappa|}\|u+z\|_{1,|\kappa|}}{1 + C|\operatorname{Im}\kappa|} \geq \frac{\|u\|_{1,|\kappa|}\|u+z\|_{1,|\kappa|}}{1 + C\frac{|\operatorname{Im}\kappa|}{1+\operatorname{Re}\kappa}}.$$

□

From the inf-sup condition, we can actually infer the statement of [27, Lemma 4.2]:

Corollary 1
Let $V = H^1(0,1)$ or $V = H^1_{(0}(0,1)$. There is $C > 0$ such that for all $\kappa \in \mathbb{C}_{\geq 0}$ with $|\kappa| \geq 1$ and all $f \in L^2(0,1)$, $g \in \mathbb{C}$ the solution $u \in V$ of

$$a^{1D}_\kappa(u,v) = (f,v)_{L^2(0,1)} + g\overline{v}(1) \quad \forall v \in V$$

satisfies the stability bound

$$\|u\|_{1,|\kappa|} \leq C\left[\frac{1}{1+\operatorname{Re}\kappa}\|f\|_{L^2(0,1)} + \frac{1}{\sqrt{1+\operatorname{Re}\kappa}}|g|\right].$$

∎

Proof: With the inf-sup constant γ_κ of Lemma 3 in hand, we get

$$\|u\|_{1,|\kappa|} \leq \frac{C}{\gamma_\kappa}\left[\frac{1}{|\kappa|}\|f\|_{L^2(0,1)} + \frac{1}{\sqrt{|\kappa|}}|g|\right].$$

We have

$$\frac{1}{\gamma_\kappa}\frac{1}{|\kappa|} \sim \frac{1}{1+C\frac{|\operatorname{Im}\kappa|}{1+\operatorname{Re}\kappa}}\frac{1}{|\operatorname{Im}\kappa| + \operatorname{Re}\kappa}$$

and now distinguish between the cases $|\operatorname{Im}\kappa| \leq 1/2$ (which implies $\operatorname{Re}\kappa \geq \sqrt{3/4}$ by the requirement $|\kappa| \geq 1$) and the case $|\operatorname{Im}\kappa| \geq 1/2$. In the first case, we have $|\operatorname{Im}\kappa|/\operatorname{Re}\kappa \leq c$ for some $c > 0$ and can simply

$$\frac{1}{1+C\frac{|\operatorname{Im}\kappa|}{1+\operatorname{Re}\kappa}}\frac{1}{|\operatorname{Im}\kappa| + \operatorname{Re}\kappa} \sim \frac{1}{\operatorname{Re}\kappa}\frac{1}{1+\frac{|\operatorname{Im}\kappa|}{\operatorname{Re}\kappa}}\frac{1}{1+\frac{|\operatorname{Im}\kappa|}{\operatorname{Re}\kappa}} \sim \frac{1}{\operatorname{Re}\kappa}.$$

In the second case, we have $|\operatorname{Im}\kappa|/\operatorname{Re}\kappa \geq c > 0$ and can estimate

$$\frac{1}{1+C\frac{|\operatorname{Im}\kappa|}{1+\operatorname{Re}\kappa}}\frac{1}{|\operatorname{Im}\kappa| + \operatorname{Re}\kappa} = \frac{1}{\operatorname{Re}\kappa}\frac{1}{1+C\frac{|\operatorname{Im}\kappa|}{1+\operatorname{Re}\kappa}}\frac{1}{\frac{|\operatorname{Im}\kappa|}{\operatorname{Re}\kappa}} \leq \frac{1}{\operatorname{Re}\kappa}\frac{1}{c}.$$

We conclude that the prefactor of the term $\|f\|_{L^2(0,1)}$ has the desired form. The estimate for the prefactor of $|g|$ is analogous. ∎

B Reduction of the first-order system to second equations for Maxwell's equations

B.1 The system for α_i, δ_i, ζ_i

We consider (4.14a), (4.14c), (4.14d), viz., (skipping the subscript i)

$$\alpha' - i\omega\delta = f_1, \tag{B.1}$$

$$\alpha - i\omega\mu^{-1}\zeta = f_3, \tag{B.2}$$

$$-\delta' + \zeta + i\omega\alpha = g_1. \tag{B.3}$$

Multiplying (B.2) by $i\omega$ and eliminating ζ yields

$$\alpha' - i\omega\delta = f_1, \tag{B.4}$$

$$-i\omega\delta' + \underbrace{(\mu - \omega^2)}_{=:\widetilde{\mu}^2}\alpha = i\omega g_1 + \mu f_3. \tag{B.5}$$

On (L, ∞), this system corresponds to the right-hand sides $f_1 = f_3 = g_1 = 0$. By eliminating δ we get

$$\alpha'' - \widetilde{\mu}^2\alpha = 0$$

with fundamental system $e^{\pm\widetilde{\mu}z}$. The requirement of outgoing waves implies that we seek solutions of the homogeneous system as multiples of $e^{-\widetilde{\mu}z}$, which provides us with the boundary condition at $z = L$

$$\alpha'(L) + \widetilde{\mu}\alpha(L) = 0. \tag{B.6}$$

In view of (B.1) on (L, ∞) with $f_1 = 0$ we obtain

$$0 \overset{(B.1)}{=} \alpha'(L) - i\omega\delta(L) \overset{(B.6)}{=} -\widetilde{\mu}\alpha(L) - i\omega\delta(L). \tag{B.7}$$

We also have the boundary condition

$$\alpha(0) = 0 \tag{B.8}$$

We now turn to the weak formulation. For v with $v(0) = 0$ we get by multiplying (B.4) by v' and integrating over $(0, L)$ and by multiplying (B.5) by v and integrating over $(0, L)$

$$(\alpha', v')_{L^2(0,L)} - i\omega(\delta, v')_{L^2(0,L)} = (f_1, v')_{L^2(0,L)}, \tag{B.9}$$

$$-i\omega(\delta', v)_{L^2(0,L)} + \widetilde{\mu}^2(\alpha, v)_{L^2(0,L)} = i\omega(g_1, v)_{L^2(0,L)} + \mu(f_3, v)_{L^2(0,L)}. \tag{B.10}$$

Integrating the term $(\delta', v)_{L^2(0,L)}$ by parts and using $v(0) = 0$ as well as $i\omega\delta(L) = \widetilde{\mu}\alpha(L)$ by (B.7) gives

$$(\alpha', v')_{L^2(0,L)} - i\omega(\delta, v')_{L^2(0,L)} = (f_1, v')_{L^2(0,L)}, \tag{B.11}$$

$$i\omega(\delta, v')_{L^2(0,L)} - \widetilde{\mu}\alpha(L)\overline{v}(L) + \widetilde{\mu}^2(\alpha, v)_{L^2(0,L)} = i\omega(g_1, v)_{L^2(0,L)} + \mu(f_3, v)_{L^2(0,L)}. \tag{B.12}$$

Adding these two equations yields

$$(\alpha', v')_{L^2(0,L)} + \widetilde{\mu}^2(\alpha, v)_{L^2(0,L)} + \widetilde{\mu}\alpha(L)\overline{v}(L) = (f_1, v')_{L^2(0,L)} + i\omega(g_1, v)_{L^2(0,L)} + \mu(f_3, v)_{L^2(0,L)}. \tag{B.13}$$

B.2 The system for β_j, γ_j, η_j

We consider (4.14b), (4.14e), (4.14f), viz., (skipping the subscript j)

$$-\beta' + \gamma - i\omega\eta = f_2 \tag{B.14}$$

$$\eta' + i\omega\beta = g_2 \tag{B.15}$$

$$\eta + i\omega\lambda^{-1}\gamma = g_3 \tag{B.16}$$

Eliminating γ in (B.14) and multiplying (B.15) by $\widetilde{\lambda}^2 := \lambda - \omega^2$ yields

$$-i\omega\beta' - \widetilde{\lambda}^2\eta = i\omega f_2 - \lambda g_3, \tag{B.17}$$

$$\widetilde{\lambda}^2\eta' + i\omega\widetilde{\lambda}^2\beta = \widetilde{\lambda}^2 g_2, \tag{B.18}$$

On (L,∞), this system corresponds to the right-hand sides $f_2 = g_2 = g_3 = 0$. By eliminating η we get

$$-\beta'' + \widetilde{\lambda}^2\beta = 0$$

with fundamental system $e^{\pm\widetilde{\lambda}z}$. The requirement of outgoing waves implies that we seek solutions of the homogeneous system as multiples of $e^{-\widetilde{\lambda}z}$, which provides us with the boundary condition at $z = L$

$$\beta'(L) + \widetilde{\lambda}\beta(L) = 0. \tag{B.19}$$

In view of (B.17) on (L,∞) with $f_2 = g_3 = 0$ we obtain

$$0 \overset{(B.17)}{=} -i\omega\beta'(L) - \widetilde{\lambda}^2\eta(L) \overset{(B.19)}{=} i\omega\widetilde{\lambda}\beta(L) - \widetilde{\lambda}^2\eta(L). \tag{B.20}$$

We also have the boundary condition

$$\beta(0) = 0. \tag{B.21}$$

We now turn to the weak formulation. For v with $v(0) = 0$ we get by multiplying (B.17) by v' and integrating over $(0,L)$ and by multiplying (B.18) by v and integrating over $(0,L)$

$$-i\omega(\beta',v')_{L^2(0,L)} - \widetilde{\lambda}^2(\eta,v')_{L^2(0,L)} = i\omega(f_2,v')_{L^2(0,L)} - \lambda(g_3,v'), \tag{B.22}$$

$$\widetilde{\lambda}^2(\eta',v)_{L^2(0,L)} + i\omega\widetilde{\lambda}^2(\beta,v)_{L^2(0,L)} = \widetilde{\lambda}^2(g_2,v)_{L^2(0,L)}. \tag{B.23}$$

Integrating the term $(\eta',v)_{L^2(0,L)}$ by parts and using $v(0) = 0$ as well as $\widetilde{\lambda}^2\eta(L) = i\omega\widetilde{\lambda}\beta(L)$ by (B.20) gives

$$-i\omega(\beta',v')_{L^2(0,L)} - \widetilde{\lambda}^2(\eta,v')_{L^2(0,L)} = i\omega(f_2,v')_{L^2(0,L)} - \lambda(g_3,v'), \tag{B.24}$$

$$-\widetilde{\lambda}^2(\eta,v')_{L^2(0,L)} + i\omega\widetilde{\lambda}\beta(L)\overline{v}(L) + i\omega\widetilde{\lambda}^2(\beta,v)_{L^2(0,L)} = \widetilde{\lambda}^2(g_2,v)_{L^2(0,L)}. \tag{B.25}$$

Subtracting these two equations yields

$$-i\omega(\beta',v')_{L^2(0,L)} - i\omega\widetilde{\lambda}^2(\beta,v)_{L^2(0,L)} - i\omega\widetilde{\lambda}\beta(L)\overline{v}(L) = i\omega(f_2,v')_{L^2(0,L)} - \lambda(g_3,v') - \widetilde{\lambda}^2(g_2,v)_{L^2(0,L)}, \tag{B.26}$$

i.e.,

$$(\beta',v')_{L^2(0,L)} + \widetilde{\lambda}^2(\beta,v)_{L^2(0,L)} + \widetilde{\lambda}\beta(L)\overline{v}(L) = -(f_2,v')_{L^2(0,L)} + \frac{\lambda}{i\omega}(g_3,v') + \widetilde{\lambda}^2\frac{1}{i\omega}(g_2,v)_{L^2(0,L)}. \tag{B.27}$$

C Stability analysis of the adjoint operator for the Helmholtz problem (2.1)

Our goal is the proof of Theorem 3, which parallels Theorem 1.

C.1 The ultraweak formulation of (2.1)

The operator A is given by

$$A\begin{pmatrix} p \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} i\omega\boldsymbol{u} + \mathfrak{a}\nabla p \\ i\omega p + \operatorname{div}\boldsymbol{u} \end{pmatrix}$$

where $D(A)$ includes the boundary conditions

$$\begin{aligned} D(A) = \{&(p,\boldsymbol{u}) \in H^1(\Omega)\times \boldsymbol{H}(\operatorname{div},\Omega)\colon \\ &p|_{\Gamma_{\text{in}}} = 0, \quad \boldsymbol{u}\cdot\boldsymbol{n} = 0 \text{ in } H^{-1/2}(\Gamma_{\text{lat}}), \quad i\omega\boldsymbol{u}\cdot\boldsymbol{n} - \operatorname{DtN} p = 0 \quad \text{in } \widetilde{H}^{-1/2}(\Gamma_{\text{out}})\}. \end{aligned}$$

The adjoint operator is[2]

$$A^*\begin{pmatrix} q \\ \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} -i\omega\boldsymbol{v} - \nabla q \\ -\operatorname{div}(\mathfrak{a}\boldsymbol{v}) - i\omega q \end{pmatrix} \tag{C.1}$$

and the domain is

$$\begin{aligned} D(A^*) = \{&(q,\boldsymbol{v}) \,|\, q \in H^1(\Omega), \mathfrak{a}\boldsymbol{v} \in \boldsymbol{H}(\operatorname{div},\Omega), \\ &q|_{\Gamma_{\text{in}}} = 0, \quad (\mathfrak{a}\boldsymbol{v})\cdot\boldsymbol{n} = 0 \quad \text{in } H^{-1/2}(\Gamma_{\text{lat}}), \quad -i\omega\mathfrak{a}\boldsymbol{v}\cdot\boldsymbol{n} + \operatorname{DtN}^* q = 0 \text{ in } \widetilde{H}^{-1/2}(\Gamma_{\text{out}})\}. \end{aligned}$$

We recall that the operator DtN is defined in Section 3.2 in terms of the eigenpair $(\varphi_n, \lambda_n)_{N\in\mathbb{N}}$ of (3.5) and that the values κ_n are defined in (3.7).

We also record that the computation shows that the adjoint of the operator DtN is given by

$$\langle p, \operatorname{DtN}^* q\rangle \stackrel{\text{def}}{=} \langle \operatorname{DtN} p, q\rangle = -\sum_n p_n \overline{\overline{\lambda_n} q_n}$$

so that

$$\operatorname{DtN}^* q = -\sum_n \overline{\lambda}_n q_n \tag{C.2}$$

C.2 Stability estimates for A^*

Introduce $H^1_{\Gamma_{\text{in}}} := \{v \in H^1(\Omega) \,|\, v|_{\Gamma_{\text{in}}} = 0\}$. In the following, we will need the following observation about the sequences $\{\lambda_n\}_n$ and $\{\kappa_n\}_n$: Noting that there are only finitely many propagating modes and that we assumed (3.8) we have

$$\max_{n\in I_{\text{prop}}} \left(|\kappa_n|^{-1} + |\kappa_n|\right) \le C, \tag{C.3a}$$

$$\max_{n\in I_{\text{prop}}} |\sqrt{\lambda_n}| \le C, \tag{C.3b}$$

$$\max_{n\in\mathbb{N}} \frac{|\sqrt{\lambda_n}|}{|\kappa_n|} \le C \tag{C.3c}$$

[2] we use that $\mathfrak{a}$ is real-valued

for a constant that depends on ω.

Lemma 14

The solution $q \in H^1_{\Gamma_{in}}$ *of*

$$(\nabla v, \mathfrak{a}\nabla q)_{L^2(\Omega)} - \omega^2 (v,q)_{L^2(\Omega)} - \langle v, \mathrm{DtN}^* q\rangle = (v,f)_{L^2(\Omega)} \quad \forall v \in H^1_{\Gamma_{in}} \tag{C.4}$$

satisfies

$$\|q\|_{H^1(\Omega)} \leq CL\|f\|_{L^2(\Omega)}.$$

∎

Proof: We make the ansatz

$$q(x,z) = \sum_n q_n(z)\varphi_n(x)$$

and set

$$f_n(z) := (\varphi_n, f(\cdot,z))_{L^2(D)}.$$

By orthogonality properties of the functions $\{\varphi_n\}_n$ we have

$$\begin{aligned}
\|q\|^2_{L^2(\Omega)} &= \sum_n \|q_n\|^2_{L^2(0,L)}, \\
\|\sqrt{a}\nabla_x q\|^2_{L^2(\Omega)} &= \sum_n \lambda_n^2 \|q_n\|^2_{L^2(0,L)}, \\
\|\partial_z q\|^2_{L^2(\Omega)} &= \sum_n \|q_n'\|^2_{L^2(0,L)}, \\
\|f\|^2_{L^2(\Omega)} &= \sum_n \|f_n\|^2_{L^2(0,L)}.
\end{aligned}$$

Testing (C.4) with $v(x,z) = v_n(z)\varphi_n(x)$ with arbitrary $v_n \in H^1_{(0}(0,L)$ gives due to the orthogonalities satisfied by the functions $\{\varphi_n\}_n$

$$(v_n', q_n')_{L^2(0,L)} + \overline{\kappa_n^2}(v_n,q_n)_{L^2(0,L)} - v_n(L)\overline{\kappa_n q}_n(L) = (v_n, f_n)_{L^2(0,L)}$$

From Lemma 4, we conclude

$$\|q_n\|_{1,|\kappa_n|} \leq C \begin{cases} L\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{prop}}, \\ \kappa_n^{-1}\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{eva}}. \end{cases}$$

We arrive at

$$\begin{aligned}
\|\sqrt{a}\nabla_x q\|^2_{L^2(\Omega)} &= \sum_n \lambda_n^2 \|q_n\|^2_{L^2(0,L)} \overset{(C.3)}{\lesssim} L^2 \sum_{n\in I_{\text{prop}}} \frac{\lambda_n^2}{|\kappa_n|^2}\|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} \frac{\lambda_n^2}{|\kappa_n|^4}\|f_n\|^2_{L^2(0,L)} \overset{(C.3)}{\lesssim} L^2\|f\|^2_{L^2(\Omega)}, \\
\|\partial_z q\|^2_{L^2(\Omega)} &= \sum_n \|q_n'\|^2_{L^2(0,L)} \lesssim L^2 \sum_{n\in I_{\text{prop}}} \|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} \lesssim L^2\|f\|^2_{L^2(\Omega)}, \\
\|q\|^2_{L^2(\Omega)} &= \sum_n \|q_n\|^2_{L^2(0,L)} \lesssim L^2 \sum_{n\in I_{\text{prop}}} |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-4}\|f_n\|^2_{L^2(0,L)} \lesssim L^2\|f\|^2_{L^2(\Omega)}.
\end{aligned}$$

∎

Lemma 15
The solution $q \in H^1_{\Gamma_{in}}$ *of*

$$(\nabla v, \mathfrak{a}\nabla q)_{L^2(\Omega)} - \omega^2 (v, q)_{L^2(\Omega)} - \langle v, \mathrm{DtN}^* q\rangle = (\nabla v, \boldsymbol{f})_{L^2(\Omega)} \quad \forall v \in H^1_{\Gamma_{in}} \tag{C.5}$$

satisfies

$$\|q\|_{H^1(\Omega)} \leq CL\|\boldsymbol{f}\|_{L^2(\Omega)}.$$

∎

Proof: We write the vector $\boldsymbol{f}$ as $\boldsymbol{f} = (\boldsymbol{f}_x, f_z)^\top$ with a vector-valued function $\boldsymbol{f}_x$ and a scalar function f_z. By linearity of the problem, we may consider the cases $(\boldsymbol{f}_x, 0)^\top$ and $(0, f_z)^\top$ as right-hand sides separately. For $\boldsymbol{f}_x = 0$, we proceed as in Lemma 14 by writing $f_z = \sum_n f_n(z)\varphi_n(x)$ and get with Lemma 4 for the corresponding functions q_n

$$\begin{aligned}\|q_n\|_{1,|\kappa_n|} &\leq C \begin{cases} L|\kappa_n|\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{prop}}, \\ \|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{eva}}. \end{cases} \\ &\leq CL\|f_n\|_{L^2(0,L)}\end{aligned}$$

since $\max_{n\in I_{\text{prop}}} |\kappa_n| \leq C$. We may repeat the calculations performed in Lemma 14 to establish

$$\begin{aligned}\|\sqrt{a}\nabla_x q\|^2_{L^2(\Omega)} &= \sum_n \lambda_n \|q_n\|^2_{L^2(0,L)} \leq CL^2 \sum_n \frac{\lambda_n}{|\kappa_n|^2}\|f_n\|^2_{L^2(0,L)} \leq C\|f_z\|^2_{L^2(\Omega)}, \\ \|\partial_z q\|^2_{L^2(\Omega)} &= \sum_n \|q_n'\|^2_{L^2(0,L)} \leq CL^2 \sum_n \|f_n\|^2_{L^2(0,L)} \leq CL^2\|f_z\|^2_{L^2(\Omega)}, \\ \|q\|^2_{L^2(\Omega)} &= \sum_n \|q_n\|^2_{L^2(0,L)} \leq CL^2 \sum_n |\kappa_n|^{-2}\|f_n\|^2_{L^2(0,L)} \leq CL^2\|f_z\|^2_{L^2(\Omega)}.\end{aligned}$$

For the case of the right-hand side $\boldsymbol{f} = (\boldsymbol{f}_x, 0)^\top$, we define $f_n(z) := (\nabla\varphi_n, \boldsymbol{f}_x(\cdot, z))_{L^2(D)} = (a\nabla\varphi_n, a^{-1}\boldsymbol{f}_x(\cdot, z))_{L^2(D)}$ and note by the fact that the functions $\{\|\sqrt{a}\nabla\varphi_n\|^{-1}_{L^2(D)}\nabla\varphi_n\}_n$ are an orthonormal (with respect to $(a\cdot,\cdot)_{L^2(D)}$) basis of its span that

$$\sum_n \lambda_n^{-1}\|f_n\|^2_{L^2(0,L)} = \int_0^L \sum_n \frac{1}{\|\sqrt{a}\nabla\varphi_n\|^2_{L^2(D)}} |(a\nabla\varphi_n, a^{-1}\boldsymbol{f}_x(\cdot, z))_{L^2(D)}|^2 \leq \|a^{-1}\boldsymbol{f}_x\|^2_{L^2(\Omega)}.$$

We expand the solution q as $q = \sum_n q_n(z)\varphi_n(x)$. Testing the equation with functions of the form $v_n(z)\varphi_n(x)$ yields again an equation for the coefficients q_n:

$$-\kappa_n^2 (v_n, q_n)_{L^2(0,L)} + (v_n', q_n')_{L^2(0,L)} - \overline{\kappa}_n v_n(L)\overline{q}_n(L) = (v_n, f_n)_{L^2(0,L)} \qquad \forall v_n \in H^1_{(0}(0, L)$$

By Lemma 4 we get

$$\|q_n\|_{1,|\kappa_n|} \leq C \begin{cases} L\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{prop}}\ , \\ |\kappa_n|^{-1}\|f_n\|_{L^2(0,L)} & \text{if } n \in I_{\text{eva}}\ . \end{cases}$$

Hence,

$$\|\sqrt{a}\nabla_x q\|^2_{L^2(\Omega)} = \sum_n \lambda_n \|q_n\|^2_{L^2(0,L)} \leq CL^2 \sum_{n\in I_{\text{prop}}} \lambda_n \|f_n\|^2_{L^2(0,L)} + C \sum_{n\in I_{\text{eva}}} \frac{\lambda_n}{|\kappa_n|^4} \|f_n\|^2_{L^2(0,L)} \leq CL\|\boldsymbol{f}_x\|^2_{L^2(\Omega)},$$

$$\|\partial_z q\|^2_{L^2(\Omega)} = \sum_n \|q_n'\|^2_{L^2(0,L)} \leq CL^2 \sum_{n\in I_{\text{prop}}} \|f_n\|^2_{L^2(0,L)} + C \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-2} \|f_n\|^2_{L^2(0,L)} \leq C\|\boldsymbol{f}_x\|^2_{L^2(\Omega)}$$

$$\|q\|^2_{L^2(\Omega)} \sum_n \|q_n\|^2_{L^2(0,L)} \leq CL^2 \sum_{n\in I_{\text{prop}}} |\kappa_n|^{-2} \|f_n\|^2_{L^2(0,L)} + \sum_{n\in I_{\text{eva}}} |\kappa_n|^{-4} \|f_n\|^2_{L^2(0,L)} \leq C\|\boldsymbol{f}_x\|^2_{L^2(\Omega)}.$$

Putting together the above results proves the claim. $\blacksquare$

THEOREM 3

There is a constant $C > 0$ (depending on $\mathfrak{a}$ and ω) such that for all $(\boldsymbol{v}, q) \in D(A^*)$

$$\|A^*(q, \boldsymbol{v})^\top\|_{L^2(\Omega)} \geq CL^{-1}\|(\boldsymbol{v}, q)\|_{L^2(\Omega)}.$$

$\blacksquare$

Proof: First, we note that $A^* : D(A^*) \to L^2(\Omega)$ is injective. Indeed, $A^*(q, \boldsymbol{v})^\top = 0$ implies $\boldsymbol{v} = i\omega\nabla q$ and therefore $q \in H^1(\Omega)$ satisfies a homogeneous second-order equation. Together with the boundary condition, one checks that $q = 0$ so that also $\boldsymbol{v} = 0$.

Abbreviate for the two components of $A^*(q, \boldsymbol{v})^\top$

$$\boldsymbol{f} := -i\omega\boldsymbol{v} - \nabla q \in L^2(\Omega), \qquad f := -\operatorname{div}(\mathfrak{a}\boldsymbol{v}) - i\omega q \in L^2(\Omega).$$

Hence, $(q, \boldsymbol{v})$ satisfy for smooth $(\widetilde{p}, \widetilde{\boldsymbol{u}})$

$$(\widetilde{\boldsymbol{u}}, -i\omega\boldsymbol{v})_{L^2(\Omega)} - (\widetilde{\boldsymbol{u}}, \nabla q)_{L^2(\Omega)} = (\widetilde{\boldsymbol{u}}, \boldsymbol{f})_{L^2(\Omega)},$$
$$(\widetilde{p}, -\operatorname{div}(\mathfrak{a}\boldsymbol{v}))_{L^2(\Omega)} - (\widetilde{p}, i\omega q)_{L^2(\Omega)} = (\widetilde{p}, f)_{L^2(\omega)}$$

Considering $\widetilde{p}$ with $\widetilde{p}|_{\Gamma_{\text{in}}} = 0$ and using the boundary conditions satisfied by $(q, \boldsymbol{v})$ (i.e., $(q, \boldsymbol{u}) \in D(A^*)$) yields after an integration by parts

$$(\widetilde{\boldsymbol{u}}, -i\omega\boldsymbol{v})_{L^2(\Omega)} - (\widetilde{\boldsymbol{u}}, \nabla q)_{L^2(\Omega)} = (\widetilde{\boldsymbol{u}}, \boldsymbol{f})_{L^2(\Omega)},$$
$$(\nabla\widetilde{p}, \mathfrak{a}\boldsymbol{v})_{L^2(\Omega)} - (\widetilde{p}, i\omega q)_{L^2(\Omega)} - \langle \widetilde{p}, \frac{1}{i\omega}\operatorname{DtN}^* q\rangle_{\Gamma_{\text{out}}} = (\widetilde{p}, f)_{L^2(\Omega)}$$

Selecting $\widetilde{\boldsymbol{u}} = \frac{1}{i\omega}\mathfrak{a}\nabla\widetilde{p}$ yields

$$(\nabla\widetilde{p}, \mathfrak{a}\boldsymbol{v})_{L^2(\Omega)} - \frac{1}{i\omega}(\mathfrak{a}\nabla\widetilde{p}, \nabla q)_{L^2(\Omega)} = \frac{1}{i\omega}(\nabla\widetilde{p}, \boldsymbol{f})_{L^2(\Omega)},$$
$$(\nabla\widetilde{p}, \mathfrak{a}\boldsymbol{v})_{L^2(\Omega)} - (\widetilde{p}, i\omega q) - \langle \widetilde{p}, \frac{1}{i\omega}\operatorname{DtN}^* q\rangle_{\Gamma_{\text{out}}} = (\widetilde{p}, f)_{L^2(\Omega)}$$

so that, by subtracting these two equations, we arrive at

$$-\frac{1}{i\omega}(\nabla\widetilde{p}, \nabla q)_{L^2(\Omega)} + (\widetilde{p}, i\omega q)_{L^2(\Omega)} + \langle \widetilde{p}, \frac{1}{i\omega}\operatorname{DtN}^* q\rangle_{\Gamma_{\text{out}}} = \frac{1}{i\omega}(\nabla\widetilde{p}, \boldsymbol{f})_{L^2(\Omega)} - (\widetilde{p}, f)_{L^2(\Omega)}$$

Rearranging terms yields

$$(\nabla\widetilde{p},\nabla q)_{L^2(\Omega)} - \omega^2(\widetilde{p},q)_{L^2(\Omega)} - \langle\widetilde{p},\mathrm{DtN}^*\,q\rangle_{\Gamma_{\mathrm{out}}} = -(\nabla\widetilde{p},\boldsymbol{f})_{L^2(\Omega)} + i\omega(\widetilde{p},f)_{L^2(\Omega)}$$

From Lemmas 14, 15 we infer

$$\|q\|_{H^1(\Omega)} \leq CL\left[\|\boldsymbol{f}\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right],$$

which in turn yields

$$\|(q,\boldsymbol{v})\|^2_{L^2(\Omega)} \leq \|q\|^2_{L^2(\Omega)} + \|\boldsymbol{v}\|^2_{L^2(\Omega)} \leq \|q\|^2_{L^2(\Omega)} + 2\omega^{-2}\|\nabla q\|^2_{L^2(\Omega)} + 2\omega^{-2}\|\boldsymbol{f}\|^2_{L^2(\Omega)}.$$

In total, we arrive at $\|(q,\boldsymbol{v})\|_{L^2(\Omega)} \leq CL\left[\|\boldsymbol{f}\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right]$, i.e., $\|(q,\boldsymbol{v})\|_{L^2(\Omega)} \leq CL\|A^*(q,\boldsymbol{v})^\top\|_{L^2(\Omega)}$, which is the claim. ∎