



# A quasilinear transmission problem with application to Maxwell equations with a divergence-free $\mathcal{D}$ -field



Tomáš Dohnal, Giulio Romani\*, Daniel P. Tietz

*Institut für Mathematik, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle (Saale), Germany*

## ARTICLE INFO

### Article history:

Received 20 September 2021  
Available online 5 February 2022  
Submitted by P. D’Ancona

### Keywords:

Quasilinear elliptic problem  
Transmission problem  
Regularity  
A-priori estimates  
Maxwell equations  
Asymptotic analysis

## ABSTRACT

Maxwell equations in the absence of free charges require initial data with a divergence-free displacement field  $\mathcal{D}$ . In materials in which the dependence  $\mathcal{D} = \mathcal{D}(\mathcal{E})$  is nonlinear the quasilinear problem  $\nabla \cdot \mathcal{D}(\mathcal{E}) = 0$  is hence to be solved. In many applications, e.g. in the modelling of wave packets, an approximative asymptotic ansatz of the electric field  $\mathcal{E}$  is used, which satisfies this divergence condition at  $t = 0$  only up to a small residual. We search then for a small correction of the ansatz to enforce  $\nabla \cdot \mathcal{D}(\mathcal{E}) = 0$  at  $t = 0$  and choose this correction in the form of a gradient field. In the usual case of a power type nonlinearity in  $\mathcal{D}(\mathcal{E})$  this leads to the sum of the Laplace and  $p$ -Laplace operators. We also allow for the medium to consist of two different materials so that a transmission problem across an interface is produced. We prove the existence of the correction term for a general class of nonlinearities and provide regularity estimates for its derivatives, independent of the  $L^2$ -norm of the original ansatz. In this way, when applied to the wave packet setting, the correction term is indeed asymptotically smaller than the original ansatz. We also provide numerical experiments to support our analysis.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

This paper is concerned with the study of the existence and regularity estimates for weak solutions of the second-order quasilinear problem defined on the whole space  $\mathbb{R}^n$ ,  $n \geq 2$ ,

$$\nabla \cdot \mathcal{D}(x, \nabla \phi + U_0(x)) = 0, \quad x \in \mathbb{R}^n, \tag{1.1}$$

where  $\nabla \cdot$  stands for the divergence operator,  $U_0 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a prescribed vector field and the function  $\mathcal{D} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is such that  $\mathcal{D}(x, \cdot)$  is growing both linearly and nonlinearly at infinity, namely

$$\mathcal{D}(x, v) := \epsilon_1(x)v + \epsilon_f(x)f(v) \quad \text{for } (x, v) \in \mathbb{R}^n \times \mathbb{R}^n, \tag{1.2}$$

\* Corresponding author.

E-mail addresses: [tomas.dohnal@mathematik.uni-halle.de](mailto:tomas.dohnal@mathematik.uni-halle.de) (T. Dohnal), [giulio.romani@mathematik.uni-halle.de](mailto:giulio.romani@mathematik.uni-halle.de) (G. Romani), [daniel.tietz@mathematik.uni-halle.de](mailto:daniel.tietz@mathematik.uni-halle.de) (D.P. Tietz).

with positive coefficients  $\epsilon_1, \epsilon_f : \mathbb{R}^n \rightarrow \mathbb{R}_+$  and where  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a superlinear vector field, the precise behaviour of which will be specified in Section 2. However, the model growth that one should have in mind is  $f(v) = |v|^{p-1}v$  for some  $p > 1$ . Our analysis is confined to the case of positive  $\epsilon_1, \epsilon_f$ . An essential feature of our problem is that the coefficients  $\epsilon_1, \epsilon_f$ , although smooth in each half-space  $\mathbb{R}^n_{\pm} := \{x \in \mathbb{R}^n \mid \pm x_1 > 0\}$ , are not continuous on  $\Gamma := \{x \in \mathbb{R}^n \mid x_1 = 0\}$ . This means that (1.1) has to be understood pointwise in  $\mathbb{R}^n_{\pm}$  and becomes a transmission problem with the straight interface  $\Gamma$ , which has to be coupled with the interface condition

$$[[\mathcal{D}(\cdot, \nabla\phi + U_0) \cdot e_1]](x) = 0, \quad x \in \Gamma. \tag{1.3}$$

Here, the symbol  $[[\varphi]]$  denotes the jump of  $\varphi$  on  $\Gamma$ . More precisely, for  $x \in \Gamma$  we define

$$[[\varphi]](x) := \lim_{h \rightarrow 0^+} (\varphi(x + he_1) - \varphi(x - he_1)),$$

where  $e_1 := (1, 0, \dots, 0)^T$  and similarly we define  $e_k$  for  $k \in \{2, \dots, n\}$ . We usually omit the variable  $x$  in the notation  $[[\varphi]](x)$ . In Sobolev spaces the interface condition has to be understood in a trace sense, see Definition 2.1.

**Motivation: a Maxwell problem** Problem (1.1)-(1.2)-(1.3) is motivated by the study of Maxwell equations when one investigates the configuration of two materials separated by a straight interface. Let us illustrate this background in more detail. Maxwell equations in nonlinear dielectric materials without free charges and currents are described by

$$\mu_0 \partial_t \mathcal{H} = -\nabla \times \mathcal{E}, \quad \epsilon_0 \partial_t \mathcal{D} = \nabla \times \mathcal{H}, \quad \nabla \cdot \mathcal{D} = \nabla \cdot \mathcal{H} = 0, \quad x \in \mathbb{R}^n, t > 0 \tag{1.4}$$

with<sup>1</sup>

$$n = 2 \quad \text{or} \quad n = 3,$$

where  $\epsilon_0$  and  $\mu_0$  are the permittivity and the permeability of the free space respectively,  $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3)^T$  and  $\mathcal{H} = (\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)^T$  are the electric and the magnetic field respectively, and the electric displacement field  $\mathcal{D} = (\mathcal{D}_1(\mathcal{E}), \mathcal{D}_2(\mathcal{E}), \mathcal{D}_3(\mathcal{E}))^T$  depends nonlinearly on the electric field. In the usual case of Kerr isotropic media, one has

$$\mathcal{D}(\cdot, \mathcal{E})(x, t) = \epsilon_1(x)\mathcal{E}(x, t) + \epsilon_3(x)(\mathcal{E}(x, t) \cdot \mathcal{E}(x, t))\mathcal{E}(x, t). \tag{1.5}$$

In the following we set  $\epsilon_0 = \mu_0 = 1$  without loss of generality. Here we consider two different materials in  $\mathbb{R}^n_-$  and  $\mathbb{R}^n_+$  respectively, divided by the interface  $\Gamma$ . We model this configuration by allowing a discontinuity of the linear and cubic susceptibility coefficients  $\epsilon_1$  and  $\epsilon_3$  across the interface, so that we have

$$\epsilon_j := \epsilon_j^- \chi_{\mathbb{R}^n_-} + \epsilon_j^+ \chi_{\mathbb{R}^n_+} \quad \text{for } j \in \{1, 3\},$$

where  $\chi_{\Omega}$  denotes the characteristic function of the set  $\Omega$ . We make the usual assumption that  $\epsilon_1, \epsilon_3$  are positive a.e. and bounded:  $0 < d \leq \epsilon_1^{\pm}, \epsilon_3^{\pm} \in L^{\infty}(\mathbb{R}^n_{\pm})$  for some constant  $d$ . Because of the inhomogeneity of the material, Maxwell equations (1.4) have to be coupled with suitable interface conditions. In particular, since  $\Gamma = \{x_1 = 0\}$ , one has to prescribe

<sup>1</sup> In the two-dimensional case  $n = 2$  we assume without loss of generality  $(\mathcal{E}, \mathcal{H}, \mathcal{D}) = (\mathcal{E}, \mathcal{H}, \mathcal{D})(x_1, x_2)$  and here  $\nabla \times = (\partial_1, \partial_2, 0)^T \times$  and  $\nabla \cdot = (\partial_1, \partial_2, 0)^T \cdot$ .

$$[[\mathcal{E}_2]] = [[\mathcal{E}_3]] = 0, \quad [[\mathcal{H}_1]] = [[\mathcal{H}_2]] = [[\mathcal{H}_3]] = 0, \quad [[\mathcal{D}_1]] = 0, \tag{1.6}$$

see e.g. [18, Section 33-3].

If the material is homogeneous or periodic in the  $x_2, x_3$  variables and one considers localised quasi-monochromatic signals, it is a standard practice in the physics literature to use an approximation by travelling pulses modulated by an envelope, see e.g. [29,35]. A formal ansatz for an asymptotic approximation of a solution  $\mathcal{U} := (\mathcal{E}, \mathcal{H})^\top$  of the Maxwell system (1.4) is then<sup>2</sup>

$$\mathcal{U}_{ans}(x, t) := \begin{pmatrix} \mathcal{E}_{ans}(x, t) \\ \mathcal{H}_{ans}(x, t) \end{pmatrix} := \varepsilon A(\varepsilon(x_{||} - \nu_g t), \varepsilon^2 t) m_{\mathcal{U}}(x_1) e^{i(k_0 \cdot x_{||} - \omega_0 t)} + c.c., \tag{1.7}$$

where  $x_{||} := (x_2, \dots, x_n)^\top$  denotes the components tangential to the interface,  $\nu_g \in \mathbb{R}^{n-1}$  is the group-velocity of the mode,  $k_0 \in \mathbb{R}^{n-1}$  is the wave vector,  $\omega_0 \in \mathbb{R}$  represents a frequency, and  $0 < \varepsilon \ll 1$  is a small asymptotic parameter. The vector function  $m_{\mathcal{U}} : \mathbb{R} \rightarrow \mathbb{C}^6$  is an exponentially decaying “interface eigenvector” of a linear problem associated with (1.4) (i.e.  $\epsilon_3 \equiv 0$ ), whereas the profile  $A : \mathbb{R}^{n-1} \times \mathbb{R} \rightarrow \mathbb{C}$  has to solve a suitable nonlinear amplitude equation of Schrödinger kind in the real slow variables  $X := \varepsilon(x_{||} - \nu_g t)$  and  $T := \varepsilon^2 t$ , see e.g. [21,16,35]. Therefore,  $A$  is smooth and exponentially decaying when sufficient conditions are assumed at the initial time  $T = 0$  of the nonlinear Schrödinger equation, see e.g. [46].

The above ansatz  $\mathcal{U}_{ans}$  has the property that the divergence equation for the  $\mathcal{H}$ -field, i.e.  $\nabla \cdot \mathcal{H} = 0$ , is automatically satisfied for all times thanks to the choice of  $m_{\mathcal{U}}$ , see Section 5. Moreover, the choice of the profile  $A$  is such that when  $\mathcal{U}_{ans}$  is inserted in the Maxwell system, it produces a sufficiently small residuum in  $\varepsilon$  for large time scales  $t \in [0, T_0(\varepsilon)]$  with  $T_0(\varepsilon) \in \mathcal{O}(\varepsilon^{-2})$  as  $\varepsilon \rightarrow 0$ . Finally, one can notice that

$$\partial_t(\nabla \cdot \mathcal{D}) = \nabla \cdot \partial_t \mathcal{D} = \nabla \cdot (\nabla \times \mathcal{H}) = 0,$$

which means that whenever the divergence of the  $\mathcal{D}$ -field vanishes at  $t = 0$ , it vanishes for all times. The problem is that  $\nabla \cdot \mathcal{D}(\cdot, \mathcal{E}_{ans}^{(0)}) \neq 0$  for

$$\mathcal{E}_{ans}^{(0)}(x) := \mathcal{E}_{ans}(x, 0) = \varepsilon A(\varepsilon x_{||}, 0) m_{\mathcal{E}}(x_1) e^{ik_0 \cdot x_{||}} + c.c., \tag{1.8}$$

where  $m_{\mathcal{E}}$  denotes the vector of the first three components of  $m_{\mathcal{U}}$ . Nonetheless one expects that this divergence condition should become satisfied by means of a suitable small correction of the ansatz above. This is precisely the motivation behind the present work. In other words, our aim is to find a vector field  $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$  which is asymptotically smaller in  $\varepsilon$  than  $\mathcal{E}_{ans}^{(0)}$  and such that

$$\mathcal{E}^{(0)} := \mathcal{E}_{ans}^{(0)} + V \tag{1.9}$$

solves

$$\nabla \cdot \mathcal{D}(\cdot, \mathcal{E}^{(0)}) = 0, \tag{1.10}$$

so that it can be used as an initial condition for the electric field. Of course,  $\mathcal{E}^{(0)}$  has also to fulfil conditions (1.6), i.e.

$$[[\mathcal{E}_2^{(0)}]] = [[\mathcal{E}_3^{(0)}]] = [[\mathcal{D}_1(\mathcal{E}^{(0)})]] = 0. \tag{1.11}$$

<sup>2</sup> Here c.c. denotes the complex conjugate of the previous term.

This correction of higher-order in  $\varepsilon$  will not derange the behaviour of  $\mathcal{E}_{ans}$  in the curl-equations. Indeed, if this step is achieved, one then prescribes  $\mathcal{U}_{ans}(x, 0) + (V, 0)^\top(x)$  as the initial condition for the electromagnetic field, which then satisfies the two divergence equations in (1.4), and which is asymptotically close to the initial value of the explicit ansatz  $\mathcal{U}_{ans}$ . Having that in hand, one can proceed following the strategy of [43] and get the existence of an exact solution of the Maxwell equations (1.4) which is well approximated by  $\mathcal{U}_{ans}$  for a large time scale. This part of the analysis is presented in detail in the forthcoming [47].

**General problem and aims** We rewrite our problem (1.1)-(1.2)-(1.3) as

$$\begin{cases} -\nabla \cdot a(x, \nabla \phi) = b(x), & x \in \mathbb{R}_\pm^n, & \text{(a)} \\ \llbracket (a(\cdot, \nabla \phi) + \varepsilon_1 U_0) \cdot e_1 \rrbracket (x) = 0, & x \in \Gamma, & \text{(b)} \end{cases} \quad (1.12)$$

with  $n \geq 2$ ,

$$a(\cdot, \nabla \phi) := \varepsilon_f f(U_0 + \nabla \phi) + \varepsilon_1 \nabla \phi \quad (= \mathcal{D}(U_0 + \nabla \phi) - \varepsilon_1 U_0)$$

and

$$b := \nabla \cdot (\varepsilon_1 U_0),$$

and where  $U_0 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a given vector field which satisfies the (linear) transmission condition  $\llbracket \varepsilon_1 U_0 \cdot e_1 \rrbracket = 0$ . Instead of (1.12b) we often use the shorter notation  $\llbracket \mathcal{D}_1(U_0 + \nabla \phi) \rrbracket = 0$  as in (1.3).

The Maxwell problem described above corresponds then to the case  $n = 2$  or  $n = 3$ ,  $U_0 = \mathcal{E}_{ans}^{(0)}$ ,  $f(v) = |v|^2 v$ , when we look for an irrotational vector field  $V$ , i.e.

$$V = \nabla \phi$$

for some  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ . Note that this formulation does not consider the first two transmission conditions  $\llbracket \mathcal{E}_2^{(0)} \rrbracket = \llbracket \mathcal{E}_3^{(0)} \rrbracket = 0$ , but they will be automatically satisfied in our application, see Section 5.

By (1.12a)-(1.12b) we aim to consider a more general setting of a second-order transmission problem in  $\mathbb{R}^n$  which involves a quasilinear operator, namely the divergence of both a linear term (which therefore produces an anisotropic Laplacian) and a nonlinear term, the growth of which is superlinear and behaves like a power. Hence, our difficulties in solving such a problem are threefold: we have to deal with a sort of anisotropic  $(p, 2)$ -Laplacian operator, with the unboundedness of the domain  $\mathbb{R}_\pm^n$  and furthermore with transmission conditions on the interface  $\Gamma$ . As a consequence, we first need to find a suitable functional setting in which to prove the existence of a solution  $\phi$ . Then, under additional regularity conditions on the coefficients on  $\mathbb{R}_\pm^n$ , we prove regularity estimates for  $\nabla \phi$  which *do not involve*  $\|U_0\|_2$ . This is of great importance for the above application:  $\|U_0\|_2 = \|\mathcal{E}_{ans}^{(0)}\|_2$  is asymptotically of order  $\mathcal{O}(\varepsilon^{1/2})$  and we are looking for  $V = \nabla \phi$  so that it is a correction of  $U_0$ , i.e.  $\|V\|_2 = \mathcal{O}(\varepsilon^{1/2})$  and for which (1.9)-(1.10) hold. This means that the estimates for  $\|\nabla \phi\|_2$  (and for its derivatives) should not depend on  $\|U_0\|_2$ . For details, see Section 5.

**Previous results** As briefly mentioned, problem (1.12a)-(1.12b) may be classified as a quasilinear transmission problem driven by an operator with a  $(p, 2)$ -growth.

A problem is referred to as *transmission* (or *diffraction*) when the domain in which the equation is defined is split in two or more subdomains in which the coefficients are regular, while at the interfaces they present jump discontinuities, and here the behaviour of the solutions is driven by some compatibility conditions. One usually imposes a condition on the jump of the solutions as well as on the normal derivative with respect to the interface. This class of problems is of great importance for physics and other applied sciences, since they can be derived from many models in which composite materials are treated, not only in the context of propagation of electromagnetic signals, but also e.g. in crack problems [2], thermodynamics

[38] and mathematical biology [10,14]. Moreover, a large amount of papers exist on the numerics of such problems, see e.g. [13,39,15,49] and references therein.

The theoretical treatment began in the '60s with the classical works of the Russian school by Oleĭnik, Ladyzhenskaya, Rivkind, Ural'tseva, and coauthors, see e.g. [40,27,28,26], where linear and quasilinear elliptic and parabolic transmission problems with smooth interfaces were first considered. In these works the main aim was to show that the solutions of the weak formulation are indeed classical, i.e. of class  $C^{2,\alpha}$  in the subdomains where the coefficients are regular, and locally of class  $C^{1,\alpha}$  near the interface, so that the transmission conditions are satisfied pointwise. Such results were later on improved and extended in [32,25,17,45]; here, however, denoting by  $z$  the unknown-variable and by  $\xi$  the gradient-variable, the growth of the function  $a(x, z, \xi)$  with respect to  $\xi$  is always supposed linear. In [31] the authors analyse the case of two adjacent materials which behave according to a model with nonlinear growth of power-type in the gradient-variable (which may also be different from material to material); such results have been later improved in [22,23]. More recently, quasilinear transmission problems even with a wild growth of the function  $a$  in the gradient-variable have been investigated in [7] by means of Orlicz spaces techniques. We also mention another direction of research for transmission problems focused on the analysis of non-smooth interfaces [11].

All the mentioned results (with the exception of [32]) deal with the case of a bounded domain. Equation (1.12a), instead, is set over the whole  $\mathbb{R}^n$ : this means, roughly speaking, that the growth in the function  $a$  with respect to  $\xi$  influences the choice of the function space. Having in mind our Maxwell setting, we assume that  $a(x, z, \xi) = a(x, \xi)$  and  $a(x, \cdot)$  consists of a linear term and power-like nonlinear terms. Therefore the operator we are dealing with can be thought of as a sum of a Laplacian and a  $p$ -Laplacian. Quasilinear problems of this kind, in the literature called  $(p, q)$ -Laplacian or *double phase* problems, have recently gained a lot of attention in the mathematical community regarding existence and regularity issues. Focusing here on those contributions in which the equation is posed on the whole  $\mathbb{R}^n$ , we mention [30,9,6], in which also a nonlinear right-hand side is considered, and [4,3,5], where in addition a more general double phase operator is considered. In these works the techniques are mainly variational, taking advantage of the mountain pass- or linking-type geometry which the functional associated with the equation enjoys. Orlicz spaces come into play in [4,3] due to the generality of the operator considered. In [8] double phase problems are shown to emerge in the context of static solutions for Lorentz invariant hyperbolic equations on  $\mathbb{R}^{3+1}$ . On the other hand, the interest in the study of *local regularity* for  $(p, q)$ -Laplacian equations initiated by Marcellini in [33] significantly grew in the last two decades, and we refer to [36,37,34] for an overview on the subject. The focus of these works is on finding sharp conditions on the growth of  $a(x, \cdot)$  with respect to the dimension  $n$ , such that the solutions of (1.12a) are locally of class  $C^{1,\alpha}$  in the case the coefficients are at least continuous.

We aim to prove standard Sobolev regularity for our solution (which is enough for our application) by means of the well-established method of the difference quotient. Nevertheless we also provide global estimates which are independent on the  $L^2$ -norm of the given vector field  $U_0$  as described above. This complicates the analysis.

The rest of the paper is organised as follows. In Section 2 we motivate and describe the suitable functional setting for problem (1.12a)-(1.12b) and we give the precise statements of the assumptions and of our main results. In Section 3 we prove existence and an estimate for  $\nabla\phi$  in terms of the right-hand side  $b$  of (1.12a) and of  $U_0$  but independent of  $\|U_0\|_2$ . Section 4 contains the proof of the estimates for the derivatives of  $\nabla\phi$  in which the discontinuity of the coefficients at the interface comes into play. We shortly illustrate then how analogous estimates are also achievable for higher derivatives of  $\nabla\phi$ . Finally in Section 5 we apply such results to the original divergence problem (1.9)-(1.10)-(1.11) arisen in the context of Maxwell equations described above and we provide a numerical verification of our results. The short Section 6 contains a discussion about questions which are left open by our analysis, and concludes the paper.

**Notation** As usual, for  $\Omega \subset \mathbb{R}^n$ ,  $C_0^\infty(\Omega)$  denotes the space of  $C^\infty$ -functions with compact support in  $\Omega$  and  $\mathcal{D}'(\Omega)$  its dual space.  $L^p(\mathbb{R}^n)^m$  is the Lebesgue space of  $p$ -integrable functions over  $\mathbb{R}^n$  which take values in  $\mathbb{R}^m$ . Of course, when  $m = 1$  we omit the second exponent. The norm in  $L^p$ -spaces for  $1 \leq p \leq \infty$  is always denoted by  $\|\cdot\|_p$  (i.e. for both scalar and vector case) and  $p' := \frac{p}{p-1}$  and  $p^* := \frac{np}{n-p}$  denote the conjugate exponent and the critical exponent for the Sobolev embedding  $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^{p^*}(\mathbb{R}^n)$ , respectively. We write  $f \in L^p(\mathbb{R}_\pm^n)$  if  $f = f_- \chi_{\mathbb{R}_-^n} + f_+ \chi_{\mathbb{R}_+^n}$  with  $f_+ \in L^p(\mathbb{R}_+^n)$  and  $f_- \in L^p(\mathbb{R}_-^n)$ . The spaces  $W^{1,p}(\mathbb{R}_\pm^n)$  are analogously defined. The ball with centre  $x_0 \in \mathbb{R}^n$  and radius  $r > 0$  is denoted by  $B_r(x_0)$ . Moreover, we refrain from writing the domain of integration when integration is meant on the whole  $\mathbb{R}^n$ . The differential  $dx$  will be often omitted.

## 2. Variational structure, assumptions and functional setting

**Variational formulation of the problem** Proceeding formally and postponing to the next paragraph the precise assumptions on  $U_0$  and  $f$ , we start by showing that if we suppose that  $f$  is irrotational, i.e.  $f = \nabla F$  for some  $F \in C^1(\mathbb{R}^n)$ , problem (1.12a) has a variational structure with the energy functional

$$\mathcal{J}(\phi) := \int \epsilon_f F(U_0 + \nabla \phi) + \int \epsilon_1 \frac{|\nabla \phi|^2}{2} + \int \epsilon_1 U_0 \cdot \nabla \phi.$$

Notice that, due to the particular form of the right-hand side  $b$ ,  $\mathcal{J}$  involves only the gradient of  $\phi$  and not the function  $\phi$  itself. Indeed, we claim that the Euler-Lagrange equation associated with  $\mathcal{J}$  is (1.12a) and that critical points of  $\mathcal{J}$  satisfy the interface condition (1.12b). For any  $\eta \in C_0^\infty(\mathbb{R}^n)$  we get

$$\mathcal{J}'[\phi](\eta) = \int \epsilon_f f(U_0 + \nabla \phi) \cdot \nabla \eta + \int \epsilon_1 (U_0 + \nabla \phi) \cdot \nabla \eta.$$

Splitting each integral over the two half-spaces  $\mathbb{R}_\pm^n$  and applying formally Gauß's theorem to each of them, we get

$$\begin{aligned} \int \epsilon_1 (U_0 + \nabla \phi) \cdot \nabla \eta &= \int_{\mathbb{R}_-^n} \epsilon_1^- (U_0 + \nabla \phi) \cdot \nabla \eta + \int_{\mathbb{R}_+^n} \epsilon_1^+ (U_0 + \nabla \phi) \cdot \nabla \eta \\ &= - \int \nabla \cdot (\epsilon_1 (U_0 + \nabla \phi)) \eta - \int_{\Gamma} ((\epsilon_1 (U_0 + \nabla \phi))_+ - (\epsilon_1 (U_0 + \nabla \phi))_-) \cdot e_1 \eta, \end{aligned}$$

where  $u_\pm(x) := \lim_{h \rightarrow 0^\pm} u(x + he_1)$ ,  $x \in \Gamma$ . Analogous computations hold for the first term

$$\begin{aligned} \int \epsilon_f f(U_0 + \nabla \phi) \cdot \nabla \eta &= - \int \nabla \cdot (\epsilon_f f(U_0 + \nabla \phi)) \eta \\ &\quad - \int_{\Gamma} ((\epsilon_f f(U_0 + \nabla \phi))_+ - (\epsilon_f f(U_0 + \nabla \phi))_-) \cdot e_1 \eta. \end{aligned}$$

All in all smooth critical points of  $\mathcal{J}$  satisfy

$$\int \nabla \cdot (\epsilon_f f(U_0 + \nabla \phi) + \epsilon_1 \nabla \phi) \eta + \int \nabla \cdot (\epsilon_1 U_0) \eta + \int_{\Gamma} (\mathcal{D}(\cdot, U_0 + \nabla \phi)_+ - \mathcal{D}(\cdot, U_0 + \nabla \phi)_-) \cdot e_1 \eta = 0.$$

Taking now  $\eta \in C_0^\infty(\mathbb{R}^n)$  with  $\text{supp } \eta \subset \mathbb{R}_\pm^n$ , the boundary integral vanishes and we obtain that the critical points of  $\mathcal{J}$  satisfy (1.12a) in  $\mathbb{R}_\pm^n$ . If, on the other hand,  $\text{supp } \eta \cap \Gamma \neq \emptyset$ , then the boundary integral tells us

that the quantity  $(\mathcal{D}(\cdot, U_0 + \nabla\phi)_+ - \mathcal{D}(\cdot, U_0 + \nabla\phi)_-) \cdot e_1 = \llbracket \mathcal{D}_1(\cdot, U_0 + \nabla\phi) \rrbracket$  vanishes a.e., which is condition (1.12b).

As we work in a Sobolev space, we need to interpret the interface condition via the *normal trace* of  $\mathcal{D}(\cdot, U_0 + \nabla\phi)$  on  $\mathbb{R}_+^n$  and  $\mathbb{R}_-^n$ . The standard definition of the normal trace for  $H_{\text{div}}$ -functions is not suitable in our setting because our solutions do not satisfy  $\mathcal{D}(\cdot, U_0 + \nabla\phi) \in L^2(\mathbb{R}^n)$ . Instead we define the trace in the sense of distributions.

**Definition 2.1.** For

$$S_{\pm} := \left\{ \mathcal{D} \in L^1_{loc}(\mathbb{R}_{\pm}^n)^n \mid \nabla \cdot \mathcal{D} \in L^1_{loc}(\mathbb{R}_{\pm}^n) \right\}$$

we define the traces

$$\mathcal{T}_+ : S_+ \rightarrow \mathfrak{D}'(\Gamma), \quad (\mathcal{T}_+ \mathcal{D})[\psi] := - \int_{\mathbb{R}_+^n} \nabla \cdot \mathcal{D} \hat{\psi} + \mathcal{D} \cdot \nabla \hat{\psi} \quad \text{for all } \psi \in C_0^\infty(\Gamma),$$

where  $\hat{\psi} \in C_0^\infty(\overline{\mathbb{R}_+^n})$  such that  $\hat{\psi}|_\Gamma = \psi$ , and analogously

$$\mathcal{T}_- : S_- \rightarrow \mathfrak{D}'(\Gamma), \quad (\mathcal{T}_- \mathcal{D})[\psi] := \int_{\mathbb{R}_-^n} \nabla \cdot \mathcal{D} \hat{\psi} + \mathcal{D} \cdot \nabla \hat{\psi} \quad \text{for all } \psi \in C_0^\infty(\Gamma),$$

where  $\hat{\psi} \in C_0^\infty(\overline{\mathbb{R}_-^n})$  such that  $\hat{\psi}|_\Gamma = \psi$ .

The existence of  $\hat{\psi}$  was proved, e.g. in [50]. The fact that the trace is independent of the choice of  $\hat{\psi}$  follows directly from the definition of the weak divergence. With this definition the interface condition (1.12b) in trace sense is

$$\mathcal{T}_+ \mathcal{D}(U_0 + \nabla\phi) = \mathcal{T}_- \mathcal{D}(U_0 + \nabla\phi) \quad \text{in } \mathfrak{D}'(\Gamma). \tag{2.1}$$

As the above calculation shows, if  $\nabla \cdot \mathcal{D}(\cdot, U_0 + \nabla\phi)$  exists in the classical sense, then (2.1) is equivalent to (1.12b) pointwise on  $\Gamma$ .

**Assumptions** Trying to retain the essential features of the Maxwell context described in the introduction, we assume the following conditions for  $f, \epsilon_1, \epsilon_f, U_0$  throughout the paper.

- (H0)  $\epsilon_1, \epsilon_f \in L^\infty(\mathbb{R}^n)$  and there exists a constant  $d$  such that  $\epsilon_1(x), \epsilon_f(x) \geq d > 0$  a.e. in  $\mathbb{R}^n$ ;
- (F0) there exist  $p > 1$  and  $F \in C^1(\mathbb{R}^n)$  convex so that  $F(0) = 0, f = \nabla F$ , and  $F(v) \geq \mu_p |v|^{p+1}$  for some  $\mu_p > 0$ ;
- (F1) there exist  $1 < \alpha \leq p$  and constants  $0 < \lambda_p \leq \Lambda_p$  and  $\Lambda_\alpha \geq 0$  such that for all  $v \in \mathbb{R}^n$ 
  - i)  $|f(v)| \leq \Lambda_p |v|^p + \Lambda_\alpha |v|^\alpha$ ;
  - ii)  $f(v) \cdot v > \lambda_p |v|^{p+1}$ ;
- (A0)  $U_0 \in L^2(\mathbb{R}^n)^n \cap L^{p+1}(\mathbb{R}^n)^n$ ;
- (A1)  $b := \nabla \cdot (\epsilon_1 U_0) \in L^2(\mathbb{R}^2) \cap L^1(\log, \mathbb{R}^2)$  if  $n = 2$  and  $b \in L^2(\mathbb{R}^n) \cap L^{\frac{2n}{n+2}}(\mathbb{R}^n)$  if  $n \geq 3$ , where  $L^1(\log, \mathbb{R}^2) := \{ \varphi \in L^1_{loc}(\mathbb{R}^2) \mid \|\varphi\|_{L^1(\log, \mathbb{R}^2)} := \int \log(2 + |x|) |\varphi(x)| dx < \infty \}$ . To unify the two cases, we write

$$\| \| b \| \| := \begin{cases} \| b \|_{L^1(\log, \mathbb{R}^2)} & \text{if } n = 2, \\ \| b \|_{L^{\frac{2n}{n+2}}} & \text{if } n \geq 3; \end{cases} \tag{2.2}$$

(A2)  $U_0$  does not weakly solve the nonlinear equation  $\nabla \cdot \mathcal{D}(\cdot, \mathcal{E}) = 0$ , where  $\mathcal{D}(\cdot, \mathcal{E})$  is defined in (1.2) (i.e.  $\phi = 0$  is not a solution in the sense of Definition 2.2), but satisfies the transmission condition  $[[\epsilon_1 U_0 \cdot e_1]] = 0$  (in the sense of Definition 2.1).

To obtain the existence and  $L^2$ -estimates for the derivatives of  $\nabla \phi$ , we need moreover

- (H1)  $\epsilon_1, \epsilon_f \in W^{1,\infty}(\mathbb{R}^n_{\pm})$ ;
- (F2)  $f \in C^1(\mathbb{R}^n)^n$  and there exist constants  $0 < \tilde{\lambda}_p \leq \tilde{\Lambda}_p$  and  $\tilde{\Lambda}_\alpha \geq 0$  such that the Jacobi matrix  $J_f$  of  $f$  satisfies for all  $v, \xi \in \mathbb{R}^n$ 
  - i)  $|J_f(v)| \leq \tilde{\Lambda}_p |v|^{p-1} + \tilde{\Lambda}_\alpha |v|^{\alpha-1}$ ;
  - ii)  $(J_f(\xi)v) \cdot v \geq \tilde{\lambda}_p |\xi|^{p-1} |v|^2$ ;
- (A3)  $\partial_k U_0 \in L^2(\mathbb{R}^n)^{n \times n} \cap L^{p+1}(\mathbb{R}^n)^{n \times n}$  for  $k \in \{2, \dots, n\}$ , and  $\partial_1 U_0 \in L^2(\mathbb{R}^n_{\pm})^{n \times n} \cap L^{p+1}(\mathbb{R}^n_{\pm})^{n \times n}$ ;
- (A4)  $\partial_k b \in L^2(\mathbb{R}^n)$ ,  $k \in \{2, \dots, n\}$ , and  $\partial_1 b \in L^2(\mathbb{R}^n_{\pm})$ .

**Remark 1.** Assumptions (H0)-(H1) on the coefficients  $\epsilon_1, \epsilon_f$  give us the structure of a transmission problem.

**Remark 2.** The simplest (and physically relevant) model nonlinearity is  $f(v) = |v|^{p-1}v$  with  $p > 1$ , but also nonlinearities of the kind  $f(v) = \sum_{i=1}^N |v|^{q_i}v$  with  $q_i > 0$  satisfy (F0)-(F2) with  $\alpha = 1 + \min_i q_i$  and  $p = 1 + \max_i q_i$  and are allowed.

**Remark 3.** Assumption (A2) reflects the situation that we encounter in our application to the Maxwell setting, where the vector field  $U_0$ , being just a solution of the linearised equation and not of the nonlinear equation  $\nabla \cdot \mathcal{D}(\cdot, \mathcal{E}) = 0$ , still satisfies the transmission condition  $[[\mathcal{D}_1(\cdot, U_0)]] = 0$ . Notice that such a transmission condition allows to apply Gauß’s theorem and transform the third term in the functional  $\mathcal{J}$  into  $\int b\phi$ , which exists thanks to assumption (A1), see Section 3.

**The main results** From assumptions (H0), (F1), and (A0), it is easy to see that the suitable functional space in which to look for critical points of the functional  $\mathcal{J}$  is

$$\mathfrak{D}_{2,p+1} := D_0^{1,2}(\mathbb{R}^n) \cap D_0^{1,p+1}(\mathbb{R}^n), \tag{2.3}$$

where for  $q \geq 1$

$$D_0^{1,q}(\mathbb{R}^n) := \overline{C_0^\infty(\mathbb{R}^n)}^{|\cdot|_{1,q}} \quad \text{with the norm} \quad |u|_{1,q} := \|\nabla u\|_q$$

is the *homogeneous Sobolev space*, sometimes in the literature referred to as the *Beppo Levi space*. The norm on  $\mathfrak{D}_{2,p+1}$  is defined as  $\|\cdot\|_{\mathfrak{D}} := |\cdot|_{1,2} + |\cdot|_{1,p+1}$ .

**Definition 2.2.** We say that  $\phi \in \mathfrak{D}_{2,p+1}$  is a *weak solution* of problem (1.12a)-(1.12b) if

$$\int \epsilon_f f(U_0 + \nabla \phi) \cdot \nabla \eta + \int \epsilon_1 (U_0 + \nabla \phi) \cdot \nabla \eta = 0 \quad \text{for all } \eta \in \mathfrak{D}_{2,p+1}. \tag{2.4}$$

Weak solutions indeed satisfy the interface condition (1.12b) in the trace sense, which follows from the definition and from the fact that  $\nabla \cdot \mathcal{D}(U_0 + \nabla \phi) = 0$  pointwise almost everywhere, see Theorem 2.2 and its proof.

Notice that (1.12a) depends just on the gradient of the unknown function, and not on the function itself. This is the first difference between our analysis and the quasilinear problems on the whole  $\mathbb{R}^n$  of the same kind which appear in the literature: indeed the  $p$ -Laplace-kind operator is often coupled with the term  $|u|^{p-2}u$ , so that one can work within the functional framework of classical Sobolev spaces, see e.g. [30,5,6].



Our main results are the following:

**Theorem 2.1.** *Let  $\epsilon_1, \epsilon_f, f, U_0$  satisfy assumptions (H0), (F0)-(F1), (A0)-(A2). Then there exists a non-trivial minimum  $\phi$  of the functional  $\mathcal{J}$  in  $\mathfrak{D}_{2,p+1}$  and there holds*

$$\int |\nabla\phi|^2 + \int |U_0 + \nabla\phi|^{p-1} |\nabla\phi|^2 \leq C \left( \|U_0\|_{p+1}^{p+1} + \|U_0\|_{\alpha+1}^{\alpha+1} + \|b\|_2^2 + \|b\|^2 \right), \tag{2.5}$$

where the constant  $C$  depends only on  $\Lambda_p, \lambda_p, \Lambda_\alpha, d, \|\epsilon_1\|_\infty$ , and  $\|\epsilon_f\|_\infty$ . Moreover,  $\phi$  is a weak solution of (1.12a)-(1.12b).

**Theorem 2.2.** *Let  $\epsilon_1, \epsilon_f, f, U_0$  satisfy assumptions (H0)-(H1), (F0)-(F2), (A0)-(A4) and let  $\phi$  be a minimiser of Theorem 2.1. Then for  $\nabla\phi$  the tangential (with respect to  $\Gamma$ ) derivatives  $\partial_k\nabla\phi, k \in \{2, \dots, n\}$ , satisfy  $\partial_k\nabla\phi \in L^2(\mathbb{R}^n)^n$  and*

$$\int |\partial_k\nabla\phi|^2 + \int |U_0 + \nabla\phi|^{p-1} |\partial_k\nabla\phi|^2 \leq C_k \left( \|U_0\|_{p+1}^{p+1} + \|\partial_k U_0\|_{p+1}^{p+1} + \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} + \|b\|_2^2 + \|b\|^2 + \|\partial_k b\|_2^2 \right). \tag{2.6}$$

The normal derivative  $\partial_{11}\phi$  exists in  $L^2(\mathbb{R}_\pm^n)$  and there holds

$$\int_{\mathbb{R}_\pm^n} |\partial_{11}\phi|^2 + \int_{\mathbb{R}_\pm^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \leq C_1 \left( \|U_0\|_{p+1}^{p+1} + \sum_{k=1}^n \|\partial_k U_0\|_{L^{p+1}(\mathbb{R}_\pm^n)}^{p+1} + \|U_0\|_{\alpha+1}^{\alpha+1} + \sum_{k=1}^n \|\partial_k U_0\|_{L^{\alpha+1}(\mathbb{R}_\pm^n)}^{\alpha+1} + \|b\|_2^2 + \|b\|^2 + \sum_{k=2}^n \|\partial_k b\|_{L^2(\mathbb{R}_\pm^n)}^2 \right). \tag{2.7}$$

The constants  $C_k, C_1$  depend only on  $\Lambda_p, \lambda_p, \Lambda_\alpha, \tilde{\Lambda}_p, \tilde{\lambda}_p, \tilde{\Lambda}_\alpha, d, \|\epsilon_1\|_{W^{1,\infty}(\mathbb{R}_\pm^n)}$ , and  $\|\epsilon_f\|_{W^{1,\infty}(\mathbb{R}_\pm^n)}$ .

Moreover, equation (1.12a) is satisfied pointwise almost everywhere and the interface condition (1.12b) holds in the trace sense of Definition 2.1.

**Remark 4.** In estimate (2.7) only the first component  $\partial_{11}\phi$  of the normal derivative  $\partial_1\nabla\phi$  appears. Indeed, for the remaining components there holds  $\partial_{1k}\phi = \partial_{k1}\phi$  a.e. in  $\mathbb{R}_\pm^n, k \in \{2, \dots, N\}$ , and thus estimate (2.6) can be directly applied. Hence, only  $\partial_{11}\phi$  has to be studied separately. For details, see Section 4.3.

**Remark 5.** An upper bound for  $\|\nabla\phi\|_2^2 + \|\nabla\phi\|_{p+1}^{p+1}$  analogous to (2.5) with the same terms on the right-hand side can be obtained by (3.9). However, we preferred to state Theorem 2.1 this way to have a better comparison with the left-hand side of the inequalities (2.6)-(2.7).

**Remark 6.** Notice that the quantity on the right-hand side of (2.5) does *not* involve  $\|U_0\|_2$ . This will be crucial in our application to the asymptotic Maxwell problem in Section 5 in order to show that  $\nabla\phi$  is actually an  $L^2$ -correction of  $U_0$ , as desired.

**Remark 7.** If one prescribes higher regularity on  $\epsilon_1, \epsilon_f, U_0, b$ , as well as a control on the higher derivatives of  $f$  analogously to assumptions (F1)-(F2), one may further infer similar estimates for the higher derivatives of  $\nabla\phi$  to the ones in (2.6)-(2.7). The proof is then analogous to that of Theorem 2.2.

**The functional setting: homogeneous Sobolev spaces** Before giving the proofs of Theorems 2.1-2.2 in the next sections, for the sake of a self-contained exposition and in order to state some properties of such spaces

that are needed in the sequel, we take now a small detour and define homogeneous Sobolev spaces following [19, Chapter II], see also [24, Sec 6]. For  $\Omega \subset \mathbb{R}^n$  arbitrary measurable domain,  $m \in \mathbb{N}$ , and  $q \geq 1$ , we start by defining

$$D^{m,q}(\Omega) := \{u \in L^1_{loc}(\Omega) \mid D^\ell u \in L^q(\Omega), |\ell| = m\}$$

with the norm  $|u|_{m,q} := \|\nabla^m u\|_q$ . If  $\Omega$  is unbounded, a control on the norm of the highest derivative does *not* imply a control on the norms of the function itself and of all other derivatives. In other words  $D^{m,q}(\Omega) \setminus W^{m,q}(\Omega) \neq \emptyset$ . Nevertheless, if  $q > 1$  any function  $u \in D^{m,q}(\Omega)$  belongs to  $W^{m,q}_{loc}(\Omega)$  and

$$\|u\|_{W^{m,q}(\Omega')} \leq C \left( \sum_{|\ell|=m} \|D^\ell u\|_{L^q(\Omega')} + \|u\|_{L^1(\Omega')} \right) \tag{2.8}$$

for any  $\Omega' \subset\subset \Omega$ , see [19, Lemma II.6.1].

An immediate disadvantage of these spaces is that, if  $P \in \mathcal{P}^m := \{\text{polynomials of degree } \leq m-1\}$ , then it is clear that  $|u|_{m,q} = |u + P|_{m,q}$ . One can avoid this by defining

$$\mathring{D}^{m,q}(\Omega) := \{[u]_m \mid u \in D^{m,q}(\Omega)\},$$

where  $[u]_m := \{w \in D^{m,q}(\Omega) \mid w = u + P, \text{ for some } P \in \mathcal{P}^m\}$  and one can show that  $\mathring{D}^{m,q}(\Omega)$  is a Banach space. Since  $C^\infty_0(\Omega) \subset \mathring{D}^{m,q}(\Omega)$  via the natural inclusion  $i : u \mapsto i(u) := [u]_m$ , the space  $D^{m,q}_0(\Omega) := \overline{C^\infty_0(\Omega)}^{|\cdot|_{m,q}}$  is isomorphic to a closed subset of  $\mathring{D}^{m,q}(\Omega)$ , so it is a Banach space too.

**Remark 8.** Notice that this means that if we consider  $u \in D^{1,q}(\mathbb{R}^n)$ , then it coincides with an element of  $D^{1,q}(\mathbb{R}^n)$ , so  $u \in L^1_{loc}(\mathbb{R}^n)$ . This in turns implies that  $u \in W^{1,q}_{loc}(\mathbb{R}^n)$  by (2.8).

The spaces  $\mathring{D}^{m,q}(\Omega)$  and  $D^{m,q}_0(\Omega)$  turn out to be separable for  $1 \leq q < +\infty$  and reflexive for  $1 < q < \infty$ , see [19, Exercise II.6.2] or [42, Theorem 2.2]. This implies that our space  $\mathfrak{D}_{2,p+1}$  defined in (2.3) is a reflexive Banach space. Indeed, the spaces  $D^{1,2}_0(\mathbb{R}^n)$  and  $D^{1,p+1}_0(\mathbb{R}^n)$  form a conjugate couple of Banach spaces (i.e. their intersection is dense in both spaces) and therefore, see [1, Theorem 8.III],

$$\begin{aligned} \mathfrak{D}''_{2,p+1} &= \left( D^{1,2}_0(\mathbb{R}^n) \cap D^{1,p+1}_0(\mathbb{R}^n) \right)'' = \left( D^{1,2}_0(\mathbb{R}^n)' + D^{1,p+1}_0(\mathbb{R}^n)' \right)' \\ &= D^{1,2}_0(\mathbb{R}^n)'' \cap D^{1,p+1}_0(\mathbb{R}^n)'' = D^{1,2}_0(\mathbb{R}^n) \cap D^{1,p+1}_0(\mathbb{R}^n) = \mathfrak{D}_{2,p+1}. \end{aligned}$$

Finally, we recall a result which allows us to split a function which belongs to  $\mathring{D}^{1,q}(\Omega)$  into a ‘‘Sobolev’’ term in  $W^{1,q}$  and a ‘‘regular’’ term which might not be in  $W^{1,q}$  but over which we may have a control in some Lebesgue space.

**Proposition 2.3** ([41], Theorem 2.2 (iii)). *For  $q \in [1, \infty)$  there exist linear maps  $T_0 : \mathring{D}^{1,q}(\mathbb{R}^n) \rightarrow W^{1,q}(\mathbb{R}^n)$  and  $T_\infty : \mathring{D}^{1,q}(\mathbb{R}^n) \rightarrow C^\infty(\mathbb{R}^n)$  such that  $[u] = [T_0(u)] + [T_\infty(u)]$  for all  $[u] \in \mathring{D}^{1,q}$  and*

$$\|T_0(u)\|_{W^{1,q}} \leq C|u|_{1,q}, \quad |T_\infty(u)|_{1,q} \leq |u|_{1,q}, \quad \|\nabla T_\infty(u)\|_\infty \leq |u|_{1,q},$$

where  $C = C(n) > 0$ . Moreover  $T_\infty$  may be chosen to satisfy

$$\begin{aligned} |T_\infty(u)|(x) &\leq C|u|_{1,q} \log(2 + |x|), & x \in \mathbb{R}^n & \quad \text{if } q = n, \\ \|T_\infty(u)\|_{q^*} &\leq C|u|_{1,q} & & \quad \text{if } q < n, \end{aligned}$$

where we recall  $q^* = \frac{nq}{n-q}$ .

**Remark 9.** This result motivates our definition of the norm  $\|\cdot\|$ .

**3. Existence and estimates for  $\nabla\phi$  (proof of Theorem 2.1)**

**Notation** In the sequel we often use the symbol  $\lesssim$  to indicate that an inequality holds up to a multiplicative constant depending only on the structural constants  $n, \Lambda_p, \lambda_p, \Lambda_\alpha, \tilde{\Lambda}_p, \tilde{\lambda}_p, \tilde{\Lambda}_\alpha, d, \|\epsilon_1\|_{W^{1,\infty}(\mathbb{R}^{\pm}_x)}$ , and  $\|\epsilon_f\|_{W^{1,\infty}(\mathbb{R}^{\pm}_x)}$  but not on  $U_0, b$ , and  $\phi$ . Moreover  $\delta$  is a small positive parameter and  $C_\delta$  denotes a positive constant which smoothly depends on  $\delta$  and the value of which may vary from line to line.

**Minimisation** Under assumptions (H0), (F0)-(F1), and (A0), the existence part of Theorem 2.1 essentially follows by the standard minimisation method. First, note that the functional  $\mathcal{J}$  is well-defined on  $\mathfrak{D}_{2,p+1}$ , since

$$\begin{aligned} \left| \int \epsilon_f F(U_0 + \nabla\phi) \right| &\leq \|\epsilon_f\|_\infty \int |F(U_0 + \nabla\phi)| \\ &\leq \|\epsilon_f\|_\infty \left( \Lambda_p \int |U_0 + \nabla\phi|^{p+1} + \Lambda_\alpha \int |U_0 + \nabla\phi|^{\alpha+1} \right) \\ &\lesssim \|U_0\|_2^2 + \|U_0\|_{p+1}^{p+1} + \|\nabla\phi\|_2^2 + \|\nabla\phi\|_{p+1}^{p+1} \end{aligned}$$

since  $F(0) = 0$  and  $\alpha \in (1, p]$ . Moreover,  $\mathcal{J}$  is coercive on  $\mathfrak{D}_{2,p+1}$ . Indeed, by (F0)

$$\begin{aligned} \mathcal{J}(\phi) &\geq d\mu_p \|U_0 + \nabla\phi\|_{p+1}^{p+1} + \frac{d}{2} \|\nabla\phi\|_2^2 - \|\epsilon_1\|_\infty \|U_0\|_2 \|\nabla\phi\|_2 \\ &\geq C \|\nabla\phi\|_{p+1}^{p+1} + \left(\frac{d}{2} - \delta\right) \|\nabla\phi\|_2^2 - C \|U_0\|_{p+1}^{p+1} - C_\delta \|U_0\|_2^2, \end{aligned}$$

and the coercivity of  $\mathcal{J}$  follows by choosing  $\delta$  sufficiently small. Finally, the weakly lower semicontinuity is ensured by the convexity of  $\mathcal{J}$ , see e.g. [20, Theorem 4.5]. Indeed,  $\mathcal{J}$  is a sum of a squared weighted  $L^2$ -norm and a nonlinear term driven by  $F$ , which is convex due to (F0). Since  $\mathfrak{D}_{2,p+1}$  is a reflexive Banach space as shown in Section 2, the direct method of the calculus of variations provides the existence of a global minimum of  $\mathcal{J}$ , which is in particular a weak solution of (1.12a)-(1.12b). Note that such a minimum is not trivial because otherwise  $\nabla \cdot \mathcal{D}(\cdot, U_0) = 0$  would weakly hold and thus violate assumption (A2).

**Estimate (2.5)** Testing (2.4) with  $\eta = \phi$ , we find

$$\int \epsilon_f f(U_0 + \nabla\phi) \cdot (U_0 + \nabla\phi) + \int \epsilon_1 |\nabla\phi|^2 = \int \epsilon_f f(U_0 + \nabla\phi) \cdot U_0 - \int \epsilon_1 U_0 \cdot \nabla\phi. \tag{3.1}$$

Since  $\epsilon_1$  and  $\epsilon_f$  are bounded from below by a positive constant, see (H0), the growth condition (F1.ii) implies that

$$\int \epsilon_f f(U_0 + \nabla\phi) \cdot (U_0 + \nabla\phi) + \int \epsilon_1 |\nabla\phi|^2 \geq d\lambda_p \int |U_0 + \nabla\phi|^{p+1} + d \int |\nabla\phi|^2. \tag{3.2}$$

On the other hand, by (F1.i), Hölder inequality with exponents  $\frac{p+1}{p}$  and  $p + 1$ , and a  $\delta$ -Young inequality we have

$$\begin{aligned} \left| \int \epsilon_f f(U_0 + \nabla\phi) \cdot U_0 \right| &\leq \|\epsilon_f\|_\infty \left( \Lambda_p \int |U_0 + \nabla\phi|^p |U_0| + \Lambda_\alpha \int |U_0 + \nabla\phi|^\alpha |U_0| \right) \\ &\leq \delta \int |U_0 + \nabla\phi|^{p+1} + C_\delta \int |U_0|^{p+1} + \delta \int |U_0 + \nabla\phi|^{\alpha+1} + C_\delta \int |U_0|^{\alpha+1} \end{aligned}$$

and, using  $x^{\alpha+1} \lesssim x^{p+1} + x^2$  for  $x \geq 0$ , since  $\alpha \in (1, p]$ ,

$$\begin{aligned} \int |U_0 + \nabla\phi|^{\alpha+1} &\lesssim \int |\nabla\phi|^{p+1} + \int |\nabla\phi|^2 + \int |U_0|^{\alpha+1} \\ &\lesssim \int |U_0 + \nabla\phi|^{p+1} + \int |\nabla\phi|^2 + \|U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1}. \end{aligned}$$

Therefore

$$\left| \int \epsilon_f f(U_0 + \nabla\phi) \cdot U_0 \right| \lesssim \delta \int |U_0 + \nabla\phi|^{p+1} + \delta \int |\nabla\phi|^2 + C_\delta \|U_0\|_{\alpha+1}^{\alpha+1} + C_\delta \|U_0\|_{p+1}^{p+1}, \quad (3.3)$$

where the first two terms on the right-hand side can be absorbed by the quantities on the right of (3.2). It remains to estimate the last term in (3.1). Since we aim to obtain an estimate of the  $L^2$ - and  $L^{p+1}$ -norms of  $\nabla\phi$  independently of the  $L^2$ -norm of  $U_0$ , we need to rely on the term  $b = \nabla \cdot (\epsilon_1 U_0)$ . By (A1)-(A2) we deduce  $b \in L^2(\mathbb{R}^n)$  as well as the transmission condition  $[[\epsilon_1 U_0 \cdot e_1]] = 0$ . Hence

$$- \int \epsilon_1 U_0 \cdot \nabla\phi = \int b\phi \quad (3.4)$$

and we aim to show that the right-hand side is well-defined in  $L^1$ , which is not trivial since we are working in homogeneous Sobolev spaces. To this purpose, we decompose  $\phi = u + v$  according to Proposition 2.3 with  $u := T_\infty(\phi)$  and  $v := T_0(\phi)$ . Hence we have  $v \in H^1(\mathbb{R}^n)$  with  $\|v\|_{H^1} \leq C|\phi|_{1,2}$  and either  $|u(x)| \leq C \log(2 + |x|)|\phi|_{1,2}$  if  $n = 2$  or  $\|u\|_{2^*} \leq C|\phi|_{1,2}$  if  $n \geq 3$ . Then

$$\begin{aligned} \left| \int b\phi \right| &\lesssim \|b\|_2 \|v\|_2 + \| \|b\| \| \phi \|_{1,2} \lesssim (\|b\|_2 + \| \|b\| \|) |\phi|_{1,2} \\ &\lesssim \delta \int |\nabla\phi|^2 + C_\delta (\|b\|_2^2 + \| \|b\| \|^2). \end{aligned} \quad (3.5)$$

Hence, (3.1)-(3.5) yield

$$(1 - C_1\delta) \int |U_0 + \nabla\phi|^{p+1} + (1 - C_2\delta) \int |\nabla\phi|^2 \lesssim C_\delta (\|U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \|b\|_2^2 + \| \|b\| \|^2),$$

for suitable constants  $C_1, C_2 > 0$ . To infer the desired estimate (2.5) it is now sufficient to choose  $\delta$  sufficiently small after having observed that

$$\begin{aligned} \int |U_0 + \nabla\phi|^{p-1} |\nabla\phi|^2 &\lesssim \int |U_0 + \nabla\phi|^{p+1} + \int |U_0 + \nabla\phi|^{p-1} |U_0|^2 \\ &\lesssim \int |U_0 + \nabla\phi|^{p+1} + \|U_0\|_{p+1}^{p+1}, \end{aligned} \quad (3.6)$$

where again Hölder inequality with conjugate exponents  $\frac{p+1}{p-1}$  and  $\frac{p+1}{2}$  and a  $\delta$ -Young inequality have been used. This concludes the proof of Theorem 2.1.

For later use, we note here that the ‘‘almost converse’’ inequality

$$\int |U_0 + \nabla\phi|^{p+1} \lesssim \int |U_0 + \nabla\phi|^{p-1} |\nabla\phi|^2 + \|U_0\|_{p+1}^{p+1} \quad (3.7)$$

holds. Indeed, by similar arguments as for (3.6) we have

$$\begin{aligned}
 \int |U_0 + \nabla\phi|^{p+1} &\leq \int |U_0 + \nabla\phi|^{p-1}|U_0|^2 + \int |U_0 + \nabla\phi|^{p-1}|\nabla\phi|^2 \\
 &\lesssim \|U_0\|_{p+1}^{p+1} + \int |\nabla\phi|^{p-1}|U_0|^2 + \int |U_0 + \nabla\phi|^{p-1}|\nabla\phi|^2 \\
 &\leq 2\|U_0\|_{p+1}^{p+1} + \|\nabla\phi\|_{p+1}^{p+1} + \int |U_0 + \nabla\phi|^{p-1}|\nabla\phi|^2.
 \end{aligned}
 \tag{3.8}$$

Now we further estimate

$$\begin{aligned}
 \|\nabla\phi\|_{p+1}^{p+1} &\lesssim \int |U_0 + \nabla\phi|^{p-1}|\nabla\phi|^2 + \int |U_0|^{p-1}|\nabla\phi|^2 \\
 &\leq \int |U_0 + \nabla\phi|^{p-1}|\nabla\phi|^2 + \delta\|\nabla\phi\|_{p+1}^{p+1} + C_\delta\|U_0\|_{p+1}^{p+1}
 \end{aligned}$$

by the  $\delta$ -Young inequality, so choosing  $\delta$  small enough one gets

$$\|\nabla\phi\|_{p+1}^{p+1} \lesssim \int |U_0 + \nabla\phi|^{p-1}|\nabla\phi|^2 + \|U_0\|_{p+1}^{p+1}.
 \tag{3.9}$$

Estimate (3.7) follows by combining (3.8) with (3.9).

#### 4. Estimates for the derivatives of $\nabla\phi$ (proof of Theorem 2.2)

In the previous section we proved the existence of  $\phi$  in  $\mathfrak{D}_{2,p+1}$ , together with estimate (2.5), under the sole assumption that the coefficients  $\epsilon_1, \epsilon_f$  are bounded from above and below by positive constants. Their regularity was not involved. However, for the existence of the derivatives of  $\nabla\phi$  as well as analogous estimates on them, the regularity of  $\epsilon_1, \epsilon_f$  comes into play. Assumption (H1) guarantees that their weak derivatives exist in the half-spaces  $\mathbb{R}_\pm^n$  but they might have a jump at the interface, which is the typical situation in applications. We show that the derivatives of  $\nabla\phi$  which are tangential to the interface exist on the whole  $\mathbb{R}^n$ , while the normal derivative  $\partial_{11}\phi$  is well-defined just in the two half-spaces. In both cases we obtain estimates analogous to the one found in the previous section.

We make use of the well-known method of difference quotients. In the context of transmission problems it was employed in [28] to obtain local regularity for linear equations and then extended in [45] to quasilinear problems of the kind (1.12a) with a function  $a(x, \cdot)$  with at most linear growth. On the other hand, this method was also applied to double phase quasilinear problems in [33]. Here we want to merge these two features and obtain analogous estimates for our double phase quasilinear transmission problem (1.12a)-(1.12b).

We start by defining the difference quotient of a function  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  in the direction  $e_k, k \in \{1, \dots, n\}$  as

$$\partial_k^h u(x) := \frac{u(x + he_k) - u(x)}{h}, \quad 0 < |h| \ll 1.$$

We recall that the action of the difference quotient on the product and the integral is similar to the derivative. Namely for  $u, v : \mathbb{R}^n \rightarrow \mathbb{R}$  there holds

$$\partial_k^h (uv)(x) = u(x + he_k)\partial_k^h v(x) + \partial_k^h u(x)v(x)$$

and

$$\int u(x)\partial_k^h v(x) dx = - \int \partial_k^{-h} u(x)v(x) dx.
 \tag{4.1}$$

Moreover, we will make use of the following well-known properties of the difference quotient, see e.g. [28, Chp. 2 Lemma 4.6] or [33, Lemma 2.7]:

**Lemma 4.1.** *Let  $\Omega'$  be an open set compactly contained in  $\Omega \subset \mathbb{R}^n$  with  $h_0 := \text{dist}(\Omega', \mathbb{R}^n \setminus \Omega) > 0$ .*

1. *If  $v \in W^{1,q}(\Omega)$  for some  $q \geq 1$ , then for every  $|h| \leq h_0$  there holds  $\int_{\Omega'} |\partial_k^h v|^q \leq \int_{\Omega} |\partial_k v|^q$ ;*
2. *If  $v \in L^q(\Omega)$  for some  $q > 1$  and if there exists a constant  $C > 0$  such that  $\|\partial_k^h v\|_{L^q(\Omega')} \leq C$  for every  $|h| \leq h_0$ , then  $\partial_k v \in L^q(\Omega')$  and  $\|\partial_k v\|_{L^q(\Omega')} \leq C$ ;*
3. *If  $v \in W^{1,q}(\Omega)$  for some  $q > 1$ , then  $\partial_k^h v \rightarrow \partial_k v$  strongly in  $L^q(\Omega')$  for every  $k \in \{1, \dots, n\}$ .*

When  $\Omega = \mathbb{R}^n$ , points 1.-3. hold also for  $\Omega' = \mathbb{R}^n$  and any  $h_0 > 0$ .

#### 4.1. Local estimates for the derivatives of $\nabla\phi$

Let  $k \in \{1, \dots, n\}$  and

$$\zeta \in \begin{cases} C_0^\infty(\mathbb{R}_\pm^n) & \text{if } k = 1, \\ C_0^\infty(\mathbb{R}^n) & \text{if } k \geq 2, \end{cases} \quad (4.2)$$

with  $0 < \zeta \leq 1$  in the interior of its support. In this section, we investigate at once *local* estimates for the normal derivative of  $\nabla\phi$ , which corresponds to the case  $k = 1$ , and for the tangential derivatives of  $\nabla\phi$ , i.e. the case  $k \in \{2, \dots, n\}$ . In both cases we aim to obtain the estimate

$$\begin{aligned} \int |\partial_k \nabla\phi|^2 \zeta^2 + \int |U_0 + \nabla\phi|^{p-1} |\partial_k \nabla\phi|^2 \zeta^2 &\leq C \left( \|U_0\|_{p+1}^{p+1} + \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k U_0\|_{p+1}^{p+1} \right. \\ &\quad \left. + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} + \|b\|_2^2 + \|b\|^2 + \|\partial_k b\|_2^2 \right), \end{aligned} \quad (4.3)$$

where in the case  $k = 1$  the Lebesgue norms of  $\partial_k U_0$  and  $\partial_k b$  have to be understood on  $\mathbb{R}_\pm^n$ . To this purpose it is important to choose the test function  $\zeta$  as in (4.2), and hence, when the normal derivative is considered,  $\text{supp } \zeta$  is strictly contained in one of the two half-spaces. We stress the fact that the constant  $C$  in (4.3) will depend on  $\|\nabla\zeta\|_\infty$  in addition to the structural constants of the problem as detailed in Theorem 2.2.

We test (2.4) with  $\eta = \partial_k^{-h} (\partial_k^h \phi \zeta^2)$  for  $h > 0$  when  $k \in \{2, \dots, n\}$  and  $0 < h < \text{dist}(\text{supp } \zeta, \Gamma)$  when  $k = 1$ . Analogously to (3.4) we apply Gauß's theorem, obtaining

$$\int \epsilon_f f(U_0 + \nabla\phi) \cdot \partial_k^{-h} \nabla (\partial_k^h \phi \zeta^2) + \int \epsilon_1 \nabla\phi \cdot \partial_k^{-h} \nabla (\partial_k^h \phi \zeta^2) = \int b \partial_k^{-h} (\partial_k^h \phi \zeta^2). \quad (4.4)$$

We begin by considering the linear terms. First by (4.1),

$$\begin{aligned} - \int \epsilon_1 \nabla\phi \cdot \partial_k^{-h} \nabla (\partial_k^h \phi \zeta^2) &= \int \partial_k^h (\epsilon_1 \nabla\phi) \cdot \nabla (\partial_k^h \phi \zeta^2) \\ &= \int \epsilon_1 (\cdot + h e_k) |\nabla \partial_k^h \phi|^2 \zeta^2 + 2 \int \epsilon_1 (\cdot + h e_k) \nabla \partial_k^h \phi \cdot \partial_k^h \phi \zeta \nabla \zeta \\ &\quad + \int \partial_k^h \epsilon_1 \nabla\phi \cdot \nabla \partial_k^h \phi \zeta^2 + 2 \int \partial_k^h \epsilon_1 \nabla\phi \cdot \partial_k^h \phi \zeta \nabla \zeta. \end{aligned} \quad (4.5)$$

We estimate the first term on the right in (4.5) from below by

$$\int \epsilon_1 (\cdot + h e_k) |\nabla \partial_k^h \phi|^2 \zeta^2 \geq d \int |\nabla \partial_k^h \phi|^2 \zeta^2, \quad (4.6)$$

and we estimate from above all other terms. Here we use  $\delta$ -Young inequalities, Lemma 4.1, and the estimate  $|\partial_k^h \epsilon_1| \leq \|\epsilon_1\|_{W^{1,\infty}(\mathbb{R}_{\pm}^n)}$ , which holds due to our choice of  $h$ :

$$\begin{aligned} \left| \int \epsilon_1(\cdot + he_k) \nabla \partial_k^h \phi \cdot \partial_k^h \phi \zeta \nabla \zeta \right| &\leq \|\epsilon_1\|_{\infty} \|\nabla \zeta\|_{\infty} \left( \delta \int |\nabla \partial_k^h \phi|^2 \zeta^2 + C_{\delta} \|\partial_k^h \phi \zeta\|_2^2 \right) \\ &\lesssim \delta \int |\nabla \partial_k^h \phi|^2 \zeta^2 + C_{\delta} \|\nabla \phi\|_2^2, \end{aligned} \tag{4.7}$$

$$\left| \int \partial_k^h \epsilon_1 \nabla \phi \cdot \nabla \partial_k^h \phi \zeta^2 \right| \leq \|\epsilon_1\|_{W^{1,\infty}(\mathbb{R}_{\pm}^n)} \left( \delta \int |\nabla \partial_k^h \phi|^2 \zeta^2 + C_{\delta} \|\nabla \phi\|_2^2 \right), \tag{4.8}$$

$$\left| \int \partial_k^h \epsilon_1 \nabla \phi \cdot \partial_k^h \phi \zeta \nabla \zeta \right| \leq \|\epsilon_1\|_{W^{1,\infty}(\mathbb{R}_{\pm}^n)} \|\nabla \zeta\|_{\infty} (\|\nabla \phi\|_2^2 + \|\partial_k^h \phi \zeta\|_2^2) \lesssim \|\nabla \phi\|_2^2. \tag{4.9}$$

Notice that all terms on the right-hand sides of (4.7)-(4.9) either may be absorbed by (4.6) or depend just on  $\|\nabla \phi\|_2$ , which can be estimated by known quantities using Theorem 2.1. We underline the fact that here, as well as in the following estimates, the constants hidden in the symbol  $\lesssim$  may also depend on  $\|\nabla \zeta\|_{\infty}$ .

The last term in (4.4), once integrated by parts as in (4.1), can be easily estimated:

$$\left| \int b \partial_k^{-h} (\partial_k^h \phi \zeta^2) \right| \leq \|\partial_k^h b \zeta\|_2 \|\partial_k^h \phi \zeta\|_2 \leq \|\partial_k b\|_2^2 + \|\nabla \phi\|_2^2. \tag{4.10}$$

The next step is to estimate the first term in (4.4). Following [33], we define

$$g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad g(x, \xi) := \epsilon_f(x) f(\xi)$$

and so for  $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$  one has

$$\begin{aligned} \partial_k^h g(x, u(x)) &= \frac{1}{h} \int_0^1 \frac{d}{dt} g(x + the_k, u(x) + th \partial_k^h u(x)) dt \\ &= \int_0^1 \left( \frac{\partial g}{\partial x_k}(x + the_k, u(x) + th \partial_k^h u(x)) + \sum_{j=1}^n \frac{\partial g}{\partial \xi_j}(x + the_k, u(x) + th \partial_k^h u(x)) \partial_k^h u_j(x) \right) dt \\ &= \int_0^1 \partial_k \epsilon_f(x + the_k) f(u(x) + th \partial_k^h u(x)) dt + \int_0^1 \epsilon_f(x + the_k) J_f(u(x) + th \partial_k^h u(x)) \partial_k^h u(x) dt. \end{aligned}$$

Hence, defining

$$Z_{th} := (U_0 + \nabla \phi) + th \partial_k^h (U_0 + \nabla \phi) = (1 - t)(U_0 + \nabla \phi) + t(U_0 + \nabla \phi)(\cdot + he_k),$$

we get for  $u = U_0 + \nabla \phi$

$$\begin{aligned} - \int \epsilon_f f(U_0 + \nabla \phi) \cdot \partial_k^{-h} \nabla (\partial_k^h \phi \zeta^2) &= \int \left( \int_0^1 \partial_k \epsilon_f(\cdot + the_k) f(Z_{th}) dt \right) \cdot (\nabla \partial_k^h \phi \zeta^2 + 2 \partial_k^h \phi \zeta \nabla \zeta) \\ &\quad + \int \left( \int_0^1 \epsilon_f(\cdot + the_k) J_f(Z_{th}) \partial_k^h (U_0 + \nabla \phi) dt \right) \cdot (\nabla \partial_k^h \phi \zeta^2 + 2 \partial_k^h \phi \zeta \nabla \zeta) \\ &=: T_1 + T_2 + S_{11} + S_{12} + S_{21} + S_{22}, \end{aligned} \tag{4.11}$$

using the discrete partial integration (4.1). The  $T$ -terms (resp.  $S$ -terms) originate from the two (resp. four) products contained in the first (resp. second) term. Let us analyse all these terms separately, starting with  $S_{21}$ , since it contains the quantity which is going to enter on the left-hand side of the desired inequality (2.6). Indeed, by means of (H0) and (F2.ii), one has

$$S_{21} := \int \left( \int_0^1 \epsilon_f(\cdot + the_k) J_f(Z_{th}) \nabla \partial_k^h \phi \, dt \right) \cdot \nabla \partial_k^h \phi \zeta^2 \geq d\tilde{\lambda}_p \int \left( \int_0^1 |Z_{th}|^{p-1} \, dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2. \quad (4.12)$$

All the remaining terms will be estimated from above by means of (F2.i), either by the terms on the left-hand sides of (4.12) and (4.6) multiplied by  $\delta$ , or by known quantities. For instance

$$\begin{aligned} |S_{11}| &\leq \|\epsilon_f\|_\infty \int \left( \tilde{\Lambda}_p \int_0^1 |Z_{th}|^{p-1} \, dt + \tilde{\Lambda}_\alpha \int_0^1 |Z_{th}|^{\alpha-1} \, dt \right) |\partial_k^h U_0| |\partial_k^h \nabla \phi| \zeta^2 \\ &\lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} \, dt + \int_0^1 |Z_{th}|^{\alpha-1} \, dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 \\ &\quad + C_\delta \int \left( \int_0^1 |Z_{th}|^{p-1} \, dt + \int_0^1 |Z_{th}|^{\alpha-1} \, dt \right) |\partial_k^h U_0|^2 \zeta^2, \\ &\lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} \, dt + 1 \right) |\nabla \partial_k^h \phi|^2 \zeta^2 \\ &\quad + C_\delta \int \left( \int_0^1 |Z_{th}|^{p-1} \, dt + \int_0^1 |Z_{th}|^{\alpha-1} \, dt \right) |\partial_k^h U_0|^2 \zeta^2, \end{aligned}$$

having used the  $\delta$ -Young inequality in the second step and the estimates  $x^{\alpha-1} \lesssim x^{p-1} + 1$  for  $x \geq 0$ , since  $\alpha \in (1, p]$ , in the last step. In the first term we recognise the “good” quantities. Noticing that by the definition of  $Z_{th}$  one has

$$\int_0^1 |Z_{th}|^{p-1} \, dt \lesssim |U_0 + \nabla \phi|^{p-1} + |U_0 + \nabla \phi|^{p-1}(\cdot + he_k), \quad (4.13)$$

and similarly for  $\alpha$  replacing  $p$ , we get via Hölder and Young inequalities

$$\begin{aligned} \int \left( \int_0^1 |Z_{th}|^{p-1} \, dt \right) |\partial_k^h U_0|^2 \zeta^2 &\lesssim \|U_0 + \nabla \phi\|_{p+1}^{p+1} + \|(U_0 + \nabla \phi)(\cdot + he_k)\|_{p+1}^{p+1} + 2\|\partial_k^h U_0 \zeta\|_{p+1}^{p+1} \\ &\lesssim 2\|U_0 + \nabla \phi\|_{p+1}^{p+1} + 2\|\partial_k U_0\|_{p+1}^{p+1} \\ &\lesssim \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 + \|U_0\|_{p+1}^{p+1} + \|\partial_k U_0\|_{p+1}^{p+1}, \end{aligned}$$

having used (3.7) in the last step. Similarly

$$\int \left( \int_0^1 |Z_{th}|^{\alpha-1} \, dt \right) |\partial_k^h U_0|^2 \zeta^2 \lesssim \int |U_0 + \nabla \phi|^{\alpha-1} |\partial_k^h U_0|^2 \zeta^2 + \int |U_0 + \nabla \phi|^{\alpha-1}(\cdot + he_k) |\partial_k^h U_0|^2 \zeta^2$$



$$\begin{aligned} &\lesssim \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k^h U_0 \zeta\|_{\alpha+1}^{\alpha+1} + \|\nabla \phi\|_{\alpha+1}^{\alpha+1} \\ &\lesssim \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} + \|\nabla \phi\|_{p+1}^{p+1} + \|\nabla \phi\|_2^2 \\ &\lesssim \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \|\nabla \phi\|_2^2 + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2. \end{aligned}$$

Here we used  $x^{\alpha+1} \lesssim x^{p+1} + x^2$  for  $x \geq 0$  in the second inequality since  $\alpha \in (1, p]$ , and (3.9) in the last inequality to estimate the third term. All in all we get

$$\begin{aligned} |S_{11}| &\lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} dt + 1 \right) |\nabla \partial_k^h \phi|^2 \zeta^2 + C_\delta \left( \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} \right. \\ &\quad \left. + \|U_0\|_{p+1}^{p+1} + \|\partial_k U_0\|_{p+1}^{p+1} + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 + \|\nabla \phi\|_2^2 \right). \end{aligned}$$

The terms  $S_{12}$  and  $S_{22}$  are estimated in the same way, obtaining

$$|S_{12}| \lesssim \|U_0\|_{\alpha+1}^{\alpha+1} + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 + \|\nabla \phi\|_2^2$$

and

$$|S_{22}| \lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} dt + 1 \right) |\nabla \partial_k^h \phi|^2 \zeta^2 + C_\delta \left( \|U_0\|_{p+1}^{p+1} + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 + \|\nabla \phi\|_2^2 \right).$$

It remains to estimate the  $T$ -terms in (4.11). Using (F1.i), we have

$$|T_1| \leq \|\epsilon_f\|_{W^{1,\infty}(\mathbb{R}_\pm)} \int \left( \Lambda_p \int_0^1 |Z_{th}|^p dt + \Lambda_\alpha \int_0^1 |Z_{th}|^\alpha dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2. \tag{4.14}$$

Concerning the first term, we exchange the integrals by Fubini’s theorem in order to use Hölder and  $\delta$ -Young inequalities in the  $x$ -integral, restoring the original integration order by a second application of Fubini’s theorem:

$$\begin{aligned} \int \left( \int_0^1 |Z_{th}|^p dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 &= \int_0^1 \left( \int |Z_{th}|^{\frac{p-1}{2}} |\nabla \partial_k^h \phi| \zeta |Z_{th}|^{\frac{p+1}{2}} \zeta \right) dt \\ &\leq \delta \int \left( \int_0^1 |Z_{th}|^{p-1} dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 + C_\delta \int \left( \int_0^1 |Z_{th}|^{p+1} dt \right) \\ &\lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 + C_\delta \left( \|U_0\|_{p+1}^{p+1} + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 \right). \end{aligned}$$

Note that in the last step we argued similarly to (4.13) with  $p + 1$  instead of  $p - 1$ , and then applied (3.7) to the resulting terms.

Analogously one may proceed with the second term in (4.14):

$$\int \left( \int_0^1 |Z_{th}|^\alpha dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 \leq \delta \int \left( \int_0^1 |Z_{th}|^{\alpha-1} dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 + C_\delta \int \left( \int_0^1 |Z_{th}|^{\alpha+1} dt \right)$$

$$\begin{aligned} &\lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} dt + 1 \right) |\nabla \partial_k^h \phi|^2 \zeta^2 \\ &\quad + C_\delta \left( \|U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \|\nabla \phi\|_2^2 + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 \right). \end{aligned}$$

Hence, from (4.14) we find

$$\begin{aligned} |T_1| &\lesssim \delta \int \left( \int_0^1 |Z_{th}|^{p-1} dt + 1 \right) |\nabla \partial_k^h \phi|^2 \zeta^2 \\ &\quad + C_\delta \left( \|U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \|\nabla \phi\|_2^2 + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 \right). \end{aligned}$$

Finally, a similar bound independent of  $\delta$  can be established for  $|T_2|$  since it does not involve  $\nabla \partial_k^h \phi$ :

$$|T_2| \lesssim \|U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \|\nabla \phi\|_2^2 + \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2.$$

Therefore, gathering the estimates we obtained for the terms in (4.5), (4.10), and (4.11), we infer

$$\begin{aligned} &(1 - C_1 \delta) \int |\nabla \partial_k^h \phi|^2 \zeta^2 + (1 - C_2 \delta) \int \left( \int_0^1 |Z_{th}|^{p-1} dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 \\ &\lesssim C_\delta \left( \int |U_0 + \nabla \phi|^{p-1} |\nabla \phi|^2 + \|\nabla \phi\|_2^2 + \|U_0\|_{\alpha+1}^{\alpha+1} \right. \\ &\quad \left. + \|\partial_k U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \|\partial_k U_0\|_{p+1}^{p+1} + \|b\|_2^2 + \|b\|_2^2 + \|\partial_k b\|_2^2 \right) \end{aligned} \quad (4.15)$$

for suitable constants  $C_1, C_2 > 0$ . Choosing a sufficiently small  $\delta$ , and since the second term on the left in (4.15) is positive, one deduces that  $\int |\nabla \partial_k^h \phi|^2 \zeta^2$  is uniformly bounded with respect to  $h$  and that the constant depends only on the  $W^{1,\infty}$ -norm of  $\zeta$ . Hence, by Lemma 4.1 one has  $\nabla \partial_k \phi \in L^2(\text{supp } \zeta)$  (with the  $L^2$ -norm bounded by the same right-hand side as in (4.15)) and  $\nabla \partial_k^h \phi \rightarrow \nabla \partial_k \phi$  in  $L^2(\text{supp } \zeta)$  as  $h \rightarrow 0$ , so pointwise a.e. in  $\text{supp } \zeta$ . Moreover, thanks to the continuity with respect to translations in the  $L^{p+1}$ -norm, one has

$$Z_{th} = (1-t)(U_0 + \nabla \phi) + t(U_0 + \nabla \phi)(\cdot + he_k) \rightarrow U_0 + \nabla \phi \quad \text{a.e. in } \mathbb{R}^n \quad \text{as } h \rightarrow 0.$$

Hence, applying Fubini's theorem and Fatou's lemma, one gets

$$\begin{aligned} \lim_{h \rightarrow 0} \int \left( \int_0^1 |Z_{th}|^{p-1} dt \right) |\nabla \partial_k^h \phi|^2 \zeta^2 &\geq \int_0^1 \left( \int \lim_{h \rightarrow 0} |Z_{th}|^{p-1} |\nabla \partial_k^h \phi|^2 \zeta^2 \right) dt \\ &= \int |U_0 + \nabla \phi|^{p-1} |\nabla \partial_k \phi|^2 \zeta^2, \end{aligned}$$

and therefore from (4.15), once estimate (2.5) found in Theorem 2.1 is applied, one infers (4.3).

#### 4.2. Tangential derivatives $\partial_k \nabla \phi, k \geq 2$ : estimate (2.6)

Let  $\zeta_0 \in C_0^\infty(\mathbb{R}^n)$  so that  $\zeta_0 \equiv 1$  on  $B_1(0)$  and  $\text{supp } \zeta_0 \subset B_2(0)$ , and let  $(\zeta_j)_j$  be the sequence of test functions defined as  $\zeta_j := \zeta_0(\frac{\cdot}{j})$ . Then the supports of  $(\zeta_j)_j$  grow to cover the whole  $\mathbb{R}^n$ ,  $\zeta_j(x) \rightarrow 1$  for all

$x \in \mathbb{R}^n$  as  $j \rightarrow +\infty$ , and  $\|\nabla\zeta_j\|_{W^{1,\infty}}$  are uniformly bounded. Hence an application of Fatou’s lemma to (4.3) with  $\zeta = \zeta_j$  yields (2.6).

Note that for the estimate of tangential derivatives of  $\nabla\phi$  introducing the factor  $\zeta$  in the test function  $\eta$  is actually not necessary. Indeed, testing with  $\eta = \partial_k^{-h}(\partial_k^h\phi)$  is allowed since Lemma 4.1 holds also for  $\Omega = \Omega' = \mathbb{R}^n$ . However, for the sake of a concise presentation in Section 4.3 we use a test function which can be used for all derivatives.

4.3. Normal derivative  $\partial_{11}\phi$ : estimate (2.7)

As mentioned in Remark 4, we are only left with the estimate for the first component of the normal derivative  $\partial_1\nabla\phi$ . Indeed, by the argument detailed in Section 4.1, for a  $\zeta \in C_0^\infty(\mathbb{R}_\pm^n)$  one infers the existence of  $\partial_1\nabla\phi$  in  $L^2(\text{supp}\zeta)$  together with the estimate

$$\int |\partial_1\nabla\phi|^2\zeta^2 + \int |U_0 + \nabla\phi|^{p-1}|\partial_1\nabla\phi|^2\zeta^2 \leq K_1, \tag{4.16}$$

where the constant  $K_1$  depends on the  $L^{p+1}$ - and  $L^{\alpha+1}$ -norms of  $U_0$ ,  $\partial_1U_0$ ,  $\partial_1b$  and on  $\|b\|$  in the half-space which contains  $\text{supp}\zeta$ . By the arbitrariness of  $\zeta \in C_0^\infty(\mathbb{R}_\pm^n)$ , one deduces that  $\partial_1\nabla\phi \in L_{loc}^2(\mathbb{R}_\pm^n)^n$ . Since  $\nabla\phi \in H_{loc}^1(\mathbb{R}_\pm^n)^n$ , it is easy to verify that  $\partial_{1k}\phi = \partial_{k1}\phi$  for all  $k \in \{1, \dots, n\}$  as functions in  $L^2(\mathbb{R}_\pm^n)$  and so also a.e. in  $\mathbb{R}_\pm^n$ . Hence, decomposing the domain of integration as  $\mathbb{R}^n \setminus \Gamma = \mathbb{R}_+^n \cup \mathbb{R}_-^n$ , we get for  $k \in \{2, \dots, n\}$

$$\int |\partial_{1k}\phi|^2 + \int |U_0 + \nabla\phi|^{p-1}|\partial_{1k}\phi|^2 = \int |\partial_{k1}\phi|^2 + \int |U_0 + \nabla\phi|^{p-1}|\partial_{k1}\phi|^2 \leq \tilde{C}_k, \tag{4.17}$$

where  $\tilde{C}_k$  is the right-hand side of (2.6). Hence we conclude that  $\partial_{1k}\phi = \partial_{k1}\phi$  in  $L^2(\mathbb{R}^n)$ , which means that  $\partial_{1k}\phi$ ,  $k \in \{2, \dots, n\}$  are estimated in Section 4.2.

The trick in (4.17) is of course not applicable for  $\partial_{11}\phi$ . Moreover, note that in the estimate we get from (4.16), i.e.

$$\int |\partial_{11}\phi|^2\zeta^2 + \int |U_0 + \nabla\phi|^{p-1}|\partial_{11}\phi|^2\zeta^2 \leq K_1, \tag{4.18}$$

the constant  $K_1$  does depend on  $\|\nabla\zeta\|_\infty$ . Hence, when one defines a sequence of uniformly bounded test functions  $(\zeta_j)_j$  similarly to Section 4.2, namely such as their support grows to cover e.g. the subspace  $\mathbb{R}_+^n$ , and  $\zeta_j \rightarrow 1$  pointwise on  $\mathbb{R}_+^n$ , then  $\|\nabla\zeta_j\|_\infty \rightarrow \infty$  as  $j \rightarrow \infty$  because  $\text{dist}(\text{supp}\zeta_j, \Gamma) \rightarrow 0$ . This implies that simply applying Fatou’s lemma to (4.18), as we did for the tangential derivatives, is insufficient. The aim of this section is thus to obtain an estimate for  $\partial_{11}\phi$  similar to (4.18) but independent of  $\|\nabla\zeta\|_\infty$ .

First, we claim that equation (1.12a) is actually satisfied *pointwise a.e.* in  $\mathbb{R}_\pm^n$ . Indeed, from the weak formulation (2.4) with  $\eta \in C_0^\infty(\mathbb{R}_\pm^n)$  and from (3.4) with such  $\eta$  instead of  $\phi$  we get by partial integration

$$-\int_{\mathbb{R}_\pm^n} \nabla \cdot (\epsilon_f f(U_0 + \nabla\phi) + \epsilon_1 \nabla\phi) \eta = \int_{\mathbb{R}_\pm^n} b\eta, \tag{4.19}$$

where the term on the left-hand side is well-defined. Indeed, concerning the second summand, the functions  $\partial_k(\epsilon_1\partial_k\phi) \in L_{loc}^1(\mathbb{R}_\pm^n)$  because of (4.18) for  $k = 1$  and (2.6) for  $k \in \{2, \dots, n\}$ . For the first summand we have,<sup>3</sup>

<sup>3</sup> We denote by  $J_f^{(k,j)}$  the  $(k, j)$ -element of the matrix  $J_f$  and by  $J_f^{(k,\cdot)}$  its  $k$ -th row.

$$\partial_k (\epsilon_f f_k(U_0 + \nabla\phi)) = \partial_k \epsilon_f f_k(U_0 + \nabla\phi) + \epsilon_f J_f^{(k,\cdot)}(U_0 + \nabla\phi) \cdot \partial_k(U_0 + \nabla\phi),$$

so by (F1.i)

$$\begin{aligned} \left| \int_{\text{supp } \eta} \partial_k \epsilon_f f_k(U_0 + \nabla\phi) \right| &\lesssim \|\epsilon_f\|_{W^{1,\infty}(\mathbb{R}_\pm^n)} \left( \int_{\text{supp } \eta} |U_0 + \nabla\phi|^p + \int_{\text{supp } \eta} |U_0 + \nabla\phi|^\alpha \right) \\ &\lesssim \int_{\text{supp } \eta} |U_0 + \nabla\phi|^p + |\text{supp } \eta| \lesssim \|U_0 + \nabla\phi\|_{p+1}^{p+1} + |\text{supp } \eta|, \end{aligned}$$

while by (F2.i)

$$\begin{aligned} \left| \int_{\text{supp } \eta} \epsilon_f J_f^{(k,\cdot)}(U_0 + \nabla\phi) \cdot \partial_k(U_0 + \nabla\phi) \right| &\lesssim \|\epsilon_f\|_\infty \int_{\text{supp } \eta} (|U_0 + \nabla\phi|^{p-1} + 1) (|\partial_k U_0| + |\partial_k \nabla\phi|) \\ &\lesssim \int_{\text{supp } \eta} |U_0 + \nabla\phi|^{p+1} + \int_{\text{supp } \eta} |U_0 + \nabla\phi|^{p-1} |\partial_k \nabla\phi|^2 + \int_{\text{supp } \eta} |\partial_k \nabla\phi|^2 \\ &\quad + \int_{\text{supp } \eta} |\partial_k U_0|^{p+1} + \int_{\text{supp } \eta} |\partial_k U_0|^2 + |\text{supp } \eta|. \end{aligned}$$

Here we used the trivial inequality  $t^{p-1} \lesssim t^p + 1$  for  $t \geq 0$ , then wrote  $p = \frac{p+1}{2} + \frac{p-1}{2}$  and applied the Cauchy-Schwarz inequality.

Then, by the arbitrariness of  $\eta \in C_0^\infty(\mathbb{R}_\pm^n)$  and invoking the fundamental lemma of calculus of variations, the fact that the solution  $\nabla\phi$  satisfies equation (1.12a) a.e. follows from (4.19).

Next, we prove the interface condition (1.12b) in the trace sense of Definition 2.1. As  $\nabla \cdot \mathcal{D}(U_0 + \nabla\phi) = 0$  a.e. in  $\mathbb{R}^n$ , we get

$$(T_\pm \mathcal{D})[\psi] = \mp \int_{\mathbb{R}_\pm^n} \mathcal{D} \cdot \nabla \hat{\psi},$$

where  $\hat{\psi} \in C_0^\infty(\mathbb{R}^n)$  is such that  $\hat{\psi}|_\Gamma = \psi$ . Hence we have

$$(T_+ \mathcal{D} - T_- \mathcal{D})[\psi] = - \int_{\mathbb{R}^n} \mathcal{D} \cdot \nabla \hat{\psi} = 0$$

using the definition of the weak solution and the fact that  $C_0^\infty(\mathbb{R}^n) \subset \mathfrak{D}_{2,p+1}$ .

In the final part of the proof we estimate the normal derivative  $\partial_{11}\phi$ . Let us now choose a half-space, say  $\mathbb{R}_+^n$ , and multiply equation (1.12a) by  $\partial_{11}\phi\zeta$ , with  $\zeta \in C_0^\infty(\mathbb{R}_+^n)$ . Integrating over  $\mathbb{R}_+^n$ , we get

$$\int_{\mathbb{R}_+^n} \nabla \cdot (\epsilon_f f(U_0 + \nabla\phi)) \partial_{11}\phi\zeta + \int_{\mathbb{R}_+^n} \nabla \cdot (\epsilon_1 \nabla\phi) \partial_{11}\phi\zeta = - \int_{\mathbb{R}_+^n} b \partial_{11}\phi\zeta. \quad (4.20)$$

We have

$$\int_{\mathbb{R}_+^n} \nabla \cdot (\epsilon_1 \nabla\phi) \partial_{11}\phi\zeta = \int_{\mathbb{R}_+^n} \epsilon_1 |\partial_{11}\phi|^2 \zeta + \sum_{k=2}^n \int_{\mathbb{R}_+^n} \epsilon_1 \partial_{kk}\phi \partial_{11}\phi\zeta + \sum_{k=1}^n \int_{\mathbb{R}_+^n} \partial_k \epsilon_1 \partial_k \phi \partial_{11}\phi\zeta \quad (4.21)$$

and for  $k \in \{2, \dots, n\}$

$$\left| \int_{\mathbb{R}_+^n} \epsilon_1 \partial_{kk} \phi \partial_{11} \phi \zeta \right| \lesssim \delta \int_{\mathbb{R}_+^n} |\partial_{11} \phi|^2 \zeta + C_\delta \int_{\mathbb{R}_+^n} |\partial_{kk} \phi|^2,$$

while for  $k \in \{1, \dots, n\}$

$$\left| \int_{\mathbb{R}_+^n} \partial_k \epsilon_1 \partial_k \phi \partial_{11} \phi \zeta \right| \lesssim \delta \int_{\mathbb{R}_+^n} |\partial_{11} \phi|^2 \zeta + C_\delta \int_{\mathbb{R}_+^n} |\nabla \phi|^2.$$

We can decompose the first term in (4.20) as

$$\begin{aligned} \int_{\mathbb{R}_+^n} \nabla \cdot (\epsilon_f f(U_0 + \nabla \phi)) \partial_{11} \phi \zeta &= \sum_{k=1}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k, \cdot)}(U_0 + \nabla \phi) \partial_k \nabla \phi \partial_{11} \phi \zeta \\ &+ \sum_{k=1}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k, \cdot)}(U_0 + \nabla \phi) \partial_k U_0 \partial_{11} \phi \zeta + \sum_{k=1}^n \int_{\mathbb{R}_+^n} \partial_k \epsilon_f f_k(U_0 + \nabla \phi) \partial_{11} \phi \zeta. \end{aligned} \tag{4.22}$$

The last two terms in (4.22) are easy to handle. For  $k \in \{1, \dots, n\}$  by (F2.i)

$$\begin{aligned} \left| \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k, \cdot)}(U_0 + \nabla \phi) \partial_k U_0 \partial_{11} \phi \zeta \right| &\lesssim \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{p-1} |\partial_k U_0| |\partial_{11} \phi| \zeta \\ &+ \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{\alpha-1} |\partial_k U_0| |\partial_{11} \phi| \zeta \\ &:= T_p + T_\alpha \end{aligned}$$

where

$$\begin{aligned} T_p &\leq C_\delta \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{p-1} |\partial_k U_0|^2 + \delta \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{p-1} |\partial_{11} \phi|^2 \zeta^2 \\ &\lesssim C_\delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{p+1} + \int_{\mathbb{R}_+^n} |\partial_k U_0|^{p+1} \right) + \delta \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{p-1} |\partial_{11} \phi|^2 \zeta^2, \end{aligned}$$

and

$$\begin{aligned} T_\alpha &\leq C_\delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{\alpha+1} + \int_{\mathbb{R}_+^n} |\partial_k U_0|^{\alpha+1} \right) + \delta \int_{\mathbb{R}_+^n} |U_0 + \nabla \phi|^{\alpha-1} |\partial_{11} \phi|^2 \zeta^2 \\ &\lesssim C_\delta \left( \|U_0\|_{\alpha+1}^{\alpha+1} + \|\nabla \phi\|_{p+1}^{p+1} + \|\nabla \phi\|_2^2 + \int_{\mathbb{R}_+^n} |\partial_k U_0|^{\alpha+1} \right) \end{aligned}$$

$$\begin{aligned}
& + \delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta^2 + \int_{\mathbb{R}_+^n} |\partial_{11}\phi|^2 \zeta^2 \right) \\
& \lesssim C_\delta \left( \|U_0\|_{\alpha+1}^{\alpha+1} + \|U_0\|_{p+1}^{p+1} + \int |U_0 + \nabla\phi|^{p-1} |\nabla\phi|^2 + \|\nabla\phi\|_2^2 + \int_{\mathbb{R}_+^n} |\partial_k U_0|^{\alpha+1} \right) \\
& + \delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta^2 + \int_{\mathbb{R}_+^n} |\partial_{11}\phi|^2 \zeta^2 \right),
\end{aligned}$$

where in the last step we applied (3.9). Hence, using (3.7) we can estimate

$$\begin{aligned}
\left| \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,\cdot)}(U_0 + \nabla\phi) \cdot \partial_k U_0 \partial_{11}\phi \zeta \right| & \lesssim C_\delta \left( \|U_0\|_{p+1}^{p+1} + \|U_0\|_{\alpha+1}^{\alpha+1} + \int |U_0 + \nabla\phi|^{p-1} |\nabla\phi|^2 + \|\nabla\phi\|_2^2 \right. \\
& \left. + \int_{\mathbb{R}_+^n} |\partial_k U_0|^{p+1} + \int_{\mathbb{R}_+^n} |\partial_k U_0|^{\alpha+1} \right) + \delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta^2 + \int_{\mathbb{R}_+^n} |\partial_{11}\phi|^2 \zeta^2 \right).
\end{aligned}$$

Similarly, by (F1.i) one has

$$\begin{aligned}
\left| \int_{\mathbb{R}_+^n} \partial_k \epsilon_f f_k(U_0 + \nabla\phi) \partial_{11}\phi \zeta \right| & \lesssim \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^p |\partial_{11}\phi| \zeta + \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^\alpha |\partial_{11}\phi| \zeta \\
& \leq C_\delta \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p+1} + \delta \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta \\
& + C_\delta \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{\alpha+1} + \delta \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{\alpha-1} |\partial_{11}\phi|^2 \zeta \\
& \lesssim \delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta + \int_{\mathbb{R}_+^n} |\partial_{11}\phi|^2 \zeta \right) \\
& + C_\delta \left( \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p+1} + \int |\nabla\phi|^2 + \|U_0\|_{p+1}^{p+1} + \|U_0\|_{\alpha+1}^{\alpha+1} \right).
\end{aligned}$$

The first term on the right hand side in (4.22) can be rewritten as follows:

$$\begin{aligned}
 & \sum_{k=1}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,\cdot)}(U_0 + \nabla\phi) \partial_k \nabla\phi \partial_{11}\phi \zeta \\
 &= \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(1,1)}(U_0 + \nabla\phi) |\partial_{11}\phi|^2 \zeta + \sum_{\substack{k,j=1 \\ (k,j) \neq (1,1)}}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,j)}(U_0 + \nabla\phi) \partial_{kj}\phi \partial_{11}\phi \zeta \\
 &= \sum_{k,j=1}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,j)}(U_0 + \nabla\phi) \partial_{1k}\phi \partial_{1j}\phi \zeta \\
 &+ \sum_{\substack{k,j=1 \\ (k,j) \neq (1,1)}}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,j)}(U_0 + \nabla\phi) (\partial_{kj}\phi \partial_{11}\phi - \partial_{1k}\phi \partial_{1j}\phi) \zeta.
 \end{aligned} \tag{4.23}$$

Since

$$\begin{aligned}
 \sum_{k,j=1}^n \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,j)}(U_0 + \nabla\phi) \partial_{1k}\phi \partial_{1j}\phi \zeta &= \int_{\mathbb{R}_+^n} \epsilon_f (J_f(U_0 + \nabla\phi) \partial_1 \nabla\phi) \cdot \partial_1 \nabla\phi \zeta \\
 &\geq d\tilde{\lambda}_p \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_1 \nabla\phi|^2 \zeta \\
 &\geq d\tilde{\lambda}_p \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta,
 \end{aligned}$$

by (F2.ii), we may recognise the term we want to keep in the final estimate. The second term on the right in (4.23) is estimated from above as usual by (F2.i) and the  $\delta$ -Young inequality:

$$\begin{aligned}
 \left| \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,j)}(U_0 + \nabla\phi) \partial_{kj}\phi \partial_{11}\phi \zeta \right| &\lesssim \delta \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta + \delta \int_{\mathbb{R}_+^n} |\partial_{11}\phi|^2 \zeta \\
 &+ C_\delta \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{kj}\phi|^2 + C_\delta \int_{\mathbb{R}_+^n} |\partial_{kj}\phi|^2 \zeta, \\
 \left| \int_{\mathbb{R}_+^n} \epsilon_f J_f^{(k,j)}(U_0 + \nabla\phi) \partial_{1k}\phi \partial_{1j}\phi \zeta \right| &\lesssim \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{1k}\phi|^2 + \int_{\mathbb{R}_+^n} |\partial_{1k}\phi|^2.
 \end{aligned}$$

This concludes the estimates for all terms in (4.22). Finally, the last term in (4.20) is straightforward:

$$\left| - \int_{\mathbb{R}_+^n} b \partial_{11}\phi \zeta \right| \leq \delta \|\partial_{11}\phi\|_2^2 + C_\delta \|b\|_2^2.$$

Recalling inequality (2.6), from equation (4.20) we may thus conclude (by choosing a sufficiently small  $\delta$ )

$$\int_{\mathbb{R}_+^n} |\partial_{11}\phi|^2 \zeta + \int_{\mathbb{R}_+^n} |U_0 + \nabla\phi|^{p-1} |\partial_{11}\phi|^2 \zeta \leq C_1^+, \tag{4.24}$$

where the constant  $C_1^+$  has the same dependencies as in Theorem 2.2. Notice that now the constant  $C_1^+$  does not depend on  $\|\zeta\|_{W^{1,\infty}(\mathbb{R}_+^n)}$  anymore, but only on  $\|\zeta\|_\infty$ . Hence, similarly to Section 4.2, one may consider a sequence of test functions  $(\zeta_j)_j \subset C_0^\infty(\mathbb{R}_+^n)$  with a uniformly bounded  $L^\infty$ -norm, which pointwise converge to 1 and the supports of which grow to cover the subspace  $\mathbb{R}_+^n$ . Then, applying Fatou’s lemma to (4.24) with  $\zeta = \zeta_j$ , we finally get the desired inequality (2.7). The argument for the half-space  $\mathbb{R}_-^n$  is completely analogous.

**5. Application: enforcing  $\nabla \cdot \mathcal{D}(\cdot, \mathcal{E}) = 0$  at  $t = 0$  in the Maxwell problem for  $n = 2$**

Let us return back to the Maxwell setting described in the introduction and apply the general results obtained in the previous sections to problem (1.9)-(1.10)-(1.11). We recall that we are considering two different materials of dielectric kind which are separated by the interface  $\Gamma = \{x_1 = 0\}$ . Within the whole section we restrict ourselves to the case  $n = 2$  and assume that the materials are independent of  $x_2, x_3$ . In addition, we choose the special case of the Kerr nonlinear dependence  $\mathcal{D} = \mathcal{D}(\mathcal{E})$  as in (1.5). With the notation of (1.2), one has

$$f(v) = |v|^2 v, \quad \epsilon_1(x) = \epsilon_1(x_1), \quad \epsilon_f(x) = \epsilon_3(x_1). \tag{5.1}$$

Of course, we require that the material coefficients  $\epsilon_1, \epsilon_3$  fulfil assumptions (H0)-(H1).

*5.1. Linear problem and numerical computation of the eigenvalues*

We start by analysing the *linear* Maxwell problem, i.e. when  $\mathcal{D} = \mathcal{D}_{lin} := \epsilon_1(x)\mathcal{E}$  or, in other words, the case  $\epsilon_3 \equiv 0$ . In this setting one may look for solutions of the form

$$\mathcal{E}(x, t) = \begin{pmatrix} \varphi_1(x_1) \\ \varphi_2(x_1) \\ \varphi_3(x_1) \end{pmatrix} e^{i(kx_2 - \omega t)} + \text{c.c.}, \quad \mathcal{H}(x, t) = \begin{pmatrix} \psi_1(x_1) \\ \psi_2(x_1) \\ \psi_3(x_1) \end{pmatrix} e^{i(kx_2 - \omega t)} + \text{c.c.} \tag{5.2}$$

with  $k \in \mathbb{R}, \omega \in \mathbb{R} \setminus \{0\}$  and  $\varphi, \psi : \mathbb{R} \rightarrow \mathbb{C}^3$ . We study *transverse magnetic* (TM) modes, that is  $\varphi = (\varphi_1, \varphi_2, 0)^\top$  and  $\psi = (0, 0, \psi_3)^\top$ . Note that in the special case of piecewise constant  $\epsilon_1$ , i.e.  $\epsilon_1 = \epsilon_1^+ \chi_{\mathbb{R}_+} + \epsilon_1^- \chi_{\mathbb{R}_-}$  with  $\epsilon_1^\pm \in \mathbb{R}$ , one can show that all solutions of the form (5.2) with  $\varphi, \psi \in L^2(\mathbb{R})^3$  are TM modes, see [12].

Arranging the nonvanishing components in the vector  $w(x_1) := (\varphi_1(x_1), \varphi_2(x_1), \psi_3(x_1))^\top$ , one sees that, in order for  $(\mathcal{E}, \mathcal{H})$  to be a solution of the linear Maxwell problem,  $w$  must satisfy the eigenvalue problem

$$\begin{cases} L(k, \omega)w = 0 & \text{in } \mathbb{R} \setminus \{0\}, \\ \llbracket \epsilon_1 w_1 \rrbracket = \llbracket w_2 \rrbracket = \llbracket w_3 \rrbracket = 0, \end{cases} \tag{5.3}$$

where the operator  $L(k, \omega) : D(L) \rightarrow L^2(\mathbb{R})^3$  is given by

$$L(k, \omega)w := \begin{pmatrix} \epsilon_1 \omega w_1 + k w_3 \\ \epsilon_1 \omega w_2 + i w_3' \\ k w_1 + i w_2' + \omega w_3 \end{pmatrix} \tag{5.4}$$

and its domain

$$D(L) := \{w \in L^2(\mathbb{R})^3 \mid w_2, w_3 \in H^1(\mathbb{R})\}$$



is dense in  $L^2(\mathbb{R})^3$ . Note that, because  $\omega \neq 0$ , an element  $w \in D(L)$  for which  $L(k, \omega)w = 0$  also satisfies the interface conditions, since  $w_2, w_3 \in H^1(\mathbb{R}) \hookrightarrow C(\mathbb{R})$  and  $\epsilon_1 \omega w_1 + k w_3 = 0$  implies that  $\epsilon_1 w_1$  is also continuous. Moreover, the first two equations of  $L(k, \omega)w = 0$  imply that

$$(\epsilon_1 w_1)' + \epsilon_1 i k w_2 = 0 \tag{5.5}$$

in  $\mathbb{R} \setminus \{0\}$ , which means that  $\nabla \cdot \mathcal{D}_{lin} = 0$  holds.

Wave numbers  $k$  for which  $w \in D(L)$  exists such that (5.3) holds are called *eigenvalues*. The dependence  $k = k(\omega)$  is called the *dispersion relation*. In general there can be more than one eigenvalue for a given  $\omega$ . Note that for the piecewise constant case the eigenvalue is unique (up to the sign), see the following remark.

**Remark 10.** One can show that it is always possible to find a solution  $w$  such that  $w_1$  and  $w_3$  are real and  $w_2$  is imaginary. Indeed, if one substitutes the ansatz  $(\tilde{w}_1, i\tilde{w}_2, \tilde{w}_3)$  in (5.3), one obtains a real problem for  $(\tilde{w}_1, \tilde{w}_2, \tilde{w}_3)$ .

One may also rewrite problem (5.3)-(5.4) as a second-order ODE on  $\mathbb{R}_+$  and  $\mathbb{R}_-$ . Indeed,

$$\begin{aligned} w_3'' &= i\epsilon_1' \omega w_2 + i\epsilon_1 \omega w_2' \\ &= \frac{\epsilon_1'}{\epsilon_1} w_3' - \epsilon_1 \omega (\omega w_3 + k w_1) \\ &= \frac{\epsilon_1'}{\epsilon_1} w_3' - \epsilon_1 \omega^2 w_3 + k^2 w_3. \end{aligned}$$

From  $[[w_2]] = 0$  we deduce  $[[\frac{w_3'}{\epsilon_1}]] = 0$ . Therefore, one needs to solve the eigenvalue problem

$$\begin{cases} -w_3'' + \frac{\epsilon_1'}{\epsilon_1} w_3' - \epsilon_1 \omega^2 w_3 = -k^2 w_3 & \text{in } \mathbb{R} \setminus \{0\}, \\ [[w_3]] = [[\frac{w_3'}{\epsilon_1}]] = 0 \end{cases} \tag{5.6}$$

and, once  $w_3$  is obtained, one may get the remaining components of  $w$  by the relations  $w_1 = -\frac{k}{\omega \epsilon_1} w_3$  and  $w_2 = -\frac{i}{\omega \epsilon_1} w_3'$ . We also see that the interface conditions  $[[\epsilon_1 w_1]] = [[w_2]] = 0$  at  $x_1 = 0$  are satisfied if  $w_3$  solves (5.6).

When  $\epsilon_1 = \epsilon_1^- \chi_{\mathbb{R}_-} + \epsilon_1^+ \chi_{\mathbb{R}_+}$  with constants  $\epsilon_1^\pm \in \mathbb{R}$ , equation (5.6) reduces to a second-order ODE with constant coefficients and one handily infers the existence of exponentially decaying solutions of the differential equation and the first interface condition in (5.6) of the form

$$w_3(x_1) = \begin{cases} C e^{-\sqrt{\lambda_{\omega,k}^-} x_1} & \text{if } x_1 \leq 0, \\ C e^{\sqrt{\lambda_{\omega,k}^+} x_1} & \text{if } x_1 > 0, \end{cases}$$

where  $C \in \mathbb{R} \setminus \{0\}$  if  $\lambda_{\omega,k}^\pm := k^2 - \epsilon_1^\pm \omega^2 > 0$ . In order for the second interface condition to be satisfied too, one needs

$$-\frac{\epsilon_1^+}{\epsilon_1^-} = \sqrt{\frac{\lambda_{\omega,k}^+}{\lambda_{\omega,k}^-}}, \tag{5.7}$$

which in turn yields the dispersion relation

$$k^2 = \omega^2 \frac{\epsilon_1^+ \epsilon_1^-}{\epsilon_1^+ + \epsilon_1^-}.$$

However, (5.7) implies that  $\epsilon_1^+$  and  $\epsilon_1^-$  should have opposite signs, which is not allowed by assumption (H0).

Therefore, next we consider the case of a non-constant  $\epsilon_1|_{\mathbb{R}_+}$  or  $\epsilon_1|_{\mathbb{R}_-}$ . Numerically we find examples for which (5.6) possesses a localised solution. We treat (5.6) as an eigenvalue problem with the eigenvalue  $k^2$ . By varying the parameter  $\omega$  we obtain a (non-explicit) dispersion relation  $k = k(\omega)$ .

**Numerical implementation** To simplify the numerics we write  $w_3$  as the sum of a  $C^1$ -component  $w_{3,r}$  and a component  $w_{3,s}$  of a simple form and with a discontinuous derivative.  $w_{3,s}$  can be expressed in term of  $w_{3,r}$  and hence we only need to solve for  $w_{3,r}$ . In detail, let  $w_3 = w_{3,r} + w_{3,s}$  with

$$w_{3,s}(x_1) = \begin{cases} w_{3,s}^- \in \mathbb{R} & \text{if } x_1 < 0, \\ w_{3,s}^+(x_1) & \text{if } x_1 \geq 0, \end{cases}$$

and choose the constant  $w_{3,s}^-$  such that  $w_{3,s}$  is continuous, i.e.  $w_{3,s}^- = w_{3,s}^+(0)$ . Note that with this choice  $w_3$  is continuous. Moreover,

$$\begin{aligned} \left[ \frac{w_3'}{\epsilon_1} \right] = 0 &\Leftrightarrow \epsilon_1^-(0) (w_{3,r}'(0) + (w_{3,s}^+)'(0)) = \epsilon_1^+(0) w_{3,r}'(0) \\ &\Leftrightarrow (w_{3,s}^+)'(0) = \frac{\epsilon_1^+(0) - \epsilon_1^-(0)}{\epsilon_1^-(0)} w_{3,r}'(0) =: \nu w_{3,r}'(0). \end{aligned} \tag{5.8}$$

If we set

$$w_{3,s}(x_1) = (\mathcal{L}w_{3,r})(x_1) := \begin{cases} -\text{sign } \nu w_{3,r}'(0) & \text{if } x_1 < 0, \\ -\text{sign } \nu w_{3,r}'(0) e^{-|\nu|x_1} & \text{if } x_1 \geq 0, \end{cases}$$

then (5.8) holds and the second interface condition in (5.6) is satisfied. Note that  $\mathcal{L} : C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$  is a linear operator and the equation for  $w_{3,r}$  is the differential equation

$$\begin{cases} \left( -\partial_{x_1}^2 + \frac{\epsilon_1'}{\epsilon_1} \partial_{x_1} - \epsilon_1 \omega^2 \right) ((I + \mathcal{L})w_{3,r}) = -k^2(I + \mathcal{L})w_{3,r} & \text{in } \mathbb{R} \setminus \{0\}, \\ \llbracket w_{3,r} \rrbracket = \llbracket w_{3,r}' \rrbracket = 0. \end{cases} \tag{5.9}$$

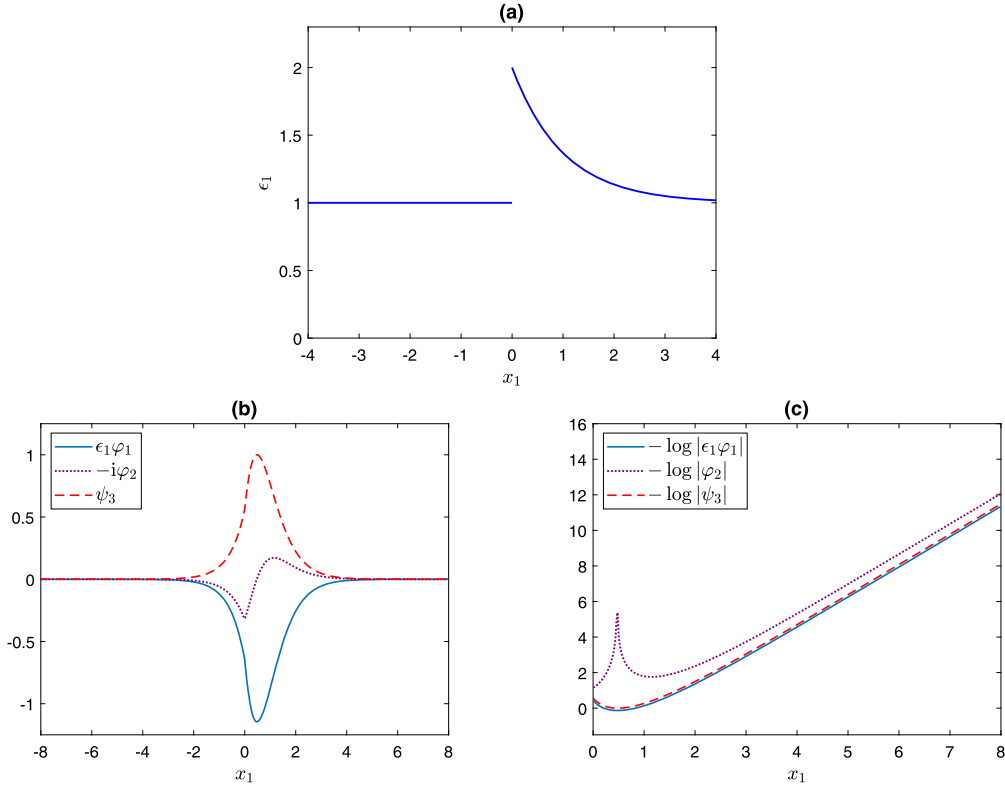
We look for solutions  $w_3$  which are eigenfunctions, i.e. with  $w_3 \in H^1(\mathbb{R})$ . In the special case when  $w_3(x_1) \rightarrow 0$  for  $|x_1| \rightarrow +\infty$ , the corresponding conditions for  $w_{3,r}$  are

$$\lim_{x_1 \rightarrow -\infty} w_{3,r}(x_1) = \text{sign } \nu w_{3,r}'(0), \quad \lim_{x_1 \rightarrow +\infty} w_{3,r}(x_1) = 0.$$

Note that we can freely choose  $w_{3,r}'(0)$ .

To solve (5.9) numerically for a fixed  $\omega \in \mathbb{R}$  we discretise the problem and apply a solver for a generalised eigenvalue problem, e.g. a solver based on a Krylov-Schur algorithm.

In Fig. 1(a) we plot a possible choice of a dielectric function  $\epsilon_1$  satisfying assumptions (H0)-(H1) and with a jump at the interface. We chose  $\epsilon_1$  constant for  $x_1 < 0$  and of the form  $1 + e^{-x_1}$  for  $x_1 \geq 0$ . The respective solution of (5.3) with an arbitrary normalisation is plotted in Fig. 1(b) and is exponentially decaying at  $\pm\infty$ , as shown in (c) for  $x_1 > 0$ . Note that it is a simple exercise to prove that the solution is exponentially decaying for  $x_1 \rightarrow -\infty$  since here the potential is constant and the calculations may be carried out explicitly. We also point out that the derivative of  $\varphi_2$  has a jump at  $x_1 = 0$ . This is due to the fact that by the third equation (5.3) one has  $-i\varphi_2' = k\varphi_1 + \omega\psi_3$ , a sum of a discontinuous and a continuous quantity.



**Fig. 1.** (a) The dielectric function  $\epsilon_1(x_1) = 1\chi_{\mathbb{R}_-} + (1 + e^{-x_1})\chi_{\mathbb{R}_+}$ . (b) The respective solution  $w(x_1) = (\varphi_1(x_1), \varphi_2(x_1), \psi_3(x_1))^T$  of the linear problem (5.3) for  $\omega_0 = 3$  (cf. Remark 10). A corresponding eigenvalue  $k_0 = k_0(\omega_0)$  produced by the numerics is  $k_0 \approx 3.4352$ . We plot  $\epsilon_1\varphi_1$  instead of  $\varphi_1$  to show the fact all interface conditions in (5.3) are satisfied. (c) Demonstration of the exponential decay of  $\epsilon_1\varphi_1$ ,  $\varphi_2$ , and  $\varphi_3$  for  $x_1 \rightarrow +\infty$ . (For colour figures see the web version of this article.)

### 5.2. Application of Theorems 2.1-2.2

Let us choose  $\omega_0 \in \mathbb{R}$  and  $k_0 = k(\omega_0)$  as one of the corresponding eigenvalues. The corresponding eigenfunction as introduced in Section 5.1 is  $(\varphi_1, \varphi_2, \psi_3)^T$ . By our choice of TM-modes, the vector  $m_{\mathcal{U}}$  in the ansatz (1.7) is defined as  $m_{\mathcal{U}}(x_1) := (\varphi_1(x_1), \varphi_2(x_1), 0, 0, 0, \psi_3(x_1))^T$  and, dropping the third vanishing component, we may redefine the vector  $m_{\mathcal{E}}$  in (1.8) as  $m_{\mathcal{E}}(x_1) := (\varphi_1(x_1), \varphi_2(x_1))^T$ . Hence our given vector field  $U_0 = (U_{0,1}, U_{0,2})^T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , with which one begins the nonlinear analysis, is

$$U_0(x_1, x_2) = \varepsilon \mathcal{A}(\varepsilon x_2) m_{\mathcal{E}}(x_1) e^{ik_0 x_2} + c.c., \tag{5.10}$$

where  $U_0$  needs to satisfy assumptions of Theorems 2.1-2.2, which is the goal of this section. In the amplitude approximation of the wavepacket (1.7),  $\mathcal{A}$  is the initial condition  $\mathcal{A}(\cdot, 0)$ . We do not study the effective Schrödinger equation for  $\mathcal{A}$  in this paper, see [21,16,35,29].

We underline the fact that  $U_0$  scales differently in  $x_1$  and  $x_2$  with respect to  $\varepsilon$ . It is easy to see that

$$\|U_0\|_2 = \mathcal{O}(\varepsilon^{1/2})$$

provided  $m_{\mathcal{E}} \in L^2(\mathbb{R})^2$  and  $\mathcal{A} \in L^2(\mathbb{R})$ . The loss of the  $\frac{1}{2}$ -power of  $\varepsilon$  is due to the scaling in the second variable. Hence, in order for the solution  $\nabla\phi$  of (1.12a)-(1.12b) to be a correction term of  $U_0$ , its  $L^2$ -norm should be  $\mathcal{O}(\varepsilon^{1/2})$ . The reason why we may expect such a behaviour for  $\nabla\phi$  is that, if we isolate the term  $b = \nabla \cdot (\epsilon_1 U_0)$  as in (1.12a), we get

$$\begin{aligned} b(x_1, x_2) &= \partial_1 (\epsilon_1 U_{0,1}) + \epsilon_1 \partial_2 U_{0,2} \\ &= [\epsilon \mathcal{A}(\epsilon x_2) ((\epsilon_1 m_{\mathcal{E},1})' + \epsilon_1 i k_0 m_{\mathcal{E},2}) + \epsilon^2 \epsilon_1 \mathcal{A}'(\epsilon x_2) m_{\mathcal{E},2}] e^{i k_0 x_2} + \text{c.c.} . \end{aligned}$$

In the first term we may recognise  $e^{i\omega t} \nabla \cdot \mathcal{D}_{lin}$  as in (5.5), so this term vanishes and hence

$$b(x_1, x_2) = \epsilon^2 \epsilon_1(x_1) \mathcal{A}'(\epsilon x_2) m_{\mathcal{E},2}(x_1) e^{i k_0 x_2} + \text{c.c.} , \tag{5.11}$$

which implies

$$\|b\|_2 = \mathcal{O}(\epsilon^{3/2}), \tag{5.12}$$

provided  $m_{\mathcal{E},2} \in L^2(\mathbb{R})$  and  $\mathcal{A}' \in L^2(\mathbb{R})$ . Notice that if we were in a smooth bounded domain, e.g. with Dirichlet boundary conditions, then the estimate  $\|\phi\|_{H^2} \leq C\epsilon^{3/2}$  would follow from the weak formulation (2.4) directly using Poincaré’s inequality. On the other hand, if one looks at the weak formulation (2.4) with  $\eta = \phi$  and naively tries to estimate  $\|\nabla\phi\|_2^2$  by the absolute value of the right-hand side and then using the Cauchy-Schwarz inequality, one ends up only with the unsatisfactory estimate  $\|\nabla\phi\|_2 \leq C\epsilon^{1/2}$ , because  $\|\epsilon_1 U_0\|_2 = \mathcal{O}(\epsilon^{1/2})$ . This is the reason why it is important that the right-hand sides of the estimates of Theorems 2.1-2.2 are not dependent on  $\|U_0\|_2$ , as briefly observed in Remark 6. With our choice of a cubic  $\mathcal{D}$ -field as in (1.5), one may first apply Theorem 2.1 with  $\alpha = p = 3$  (since there are no further nonlinear terms in  $\mathcal{D}$ ) and get, in particular,

$$\|\nabla\phi\|_2^2 \lesssim \|U_0\|_4^4 + \|b\|_2^2 + \|b\|_{L^1(\log)}^2, \tag{5.13}$$

where we recall that  $\|b\|_{L^1(\log)} = \int |b(x)| \log(2 + |x|) dx$ . The term  $\|b\|_2$  was estimated in (5.12). Next,

$$\int |U_0|^4 = \epsilon^4 \int_{\mathbb{R}} |\mathcal{A}(\epsilon x_2)|^4 dx_2 \int_{\mathbb{R}} |m_{\mathcal{E}}(x_1)|^4 dx_1 \leq \epsilon^3 \|\mathcal{A}\|_4^4 \|m_{\mathcal{E}}\|_4^4, \tag{5.14}$$

so provided  $\mathcal{A} \in L^4(\mathbb{R})$  and  $m_{\mathcal{E}} \in L^4(\mathbb{R})^2$ , one infers

$$\|U_0\|_4 = \mathcal{O}(\epsilon^{3/4}). \tag{5.15}$$

For the last term in (5.13), we first note the simple equality

$$\log(s + t) = \log(s) + \log\left(1 + \frac{t}{s}\right), \quad s, t \in \mathbb{R}_+,$$

to obtain, for a suitable constant  $c > 0$ ,

$$\begin{aligned} \log(2 + |x|) &\leq \log(2 + c|x_1| + c|x_2|) = \log(2 + c|x_1|) + \log\left(1 + \frac{c|x_2|}{2 + c|x_1|}\right) \\ &\leq \log(2 + c|x_1|) + \log(2 + c|x_2|). \end{aligned}$$

Hence,

$$\begin{aligned} \|b\|_{L^1(\log)} &\leq \varepsilon^2 \|\epsilon_1\|_\infty \int |\mathcal{A}'(\varepsilon x_2)| |m_{\mathcal{E},2}(x_1)| \log(2 + |x|) dx \\ &\lesssim \varepsilon^2 \|m_{\mathcal{E},2}\|_{L^1(\log)} \int |\mathcal{A}'(\varepsilon x_2)| dx_2 + \varepsilon^2 \|m_{\mathcal{E},2}\|_1 \int |\mathcal{A}'(\varepsilon x_2)| \log(2 + c|x_2|) dx_2 \\ &\lesssim \varepsilon^2 \|m_{\mathcal{E},2}\|_{L^1(\log)} \int |\mathcal{A}'(\varepsilon x_2)| (1 + \log(2 + c|x_2|)) dx_2 \\ &\lesssim \varepsilon \|m_{\mathcal{E},2}\|_{L^1(\log)} \int |\mathcal{A}'(y)| \log\left(e + c\frac{|y|}{\varepsilon}\right) dy, \end{aligned}$$

where in the last inequality we made use of the variable transformation  $y = \varepsilon x_2$ . Finally, choosing any  $\gamma > 0$  one may estimate  $\log(e + ct) \lesssim 1 + t^\gamma$  for all  $t \geq 0$ . Therefore

$$\|b\|_{L^1(\log)} \lesssim \|m_{\mathcal{E},2}\|_{L^1(\log)} \left( \varepsilon \|\mathcal{A}'\|_1 + \varepsilon^{1-\gamma} \int |\mathcal{A}'(y)| |y|^\gamma dy \right) = \mathcal{O}(\varepsilon^{1-\gamma}), \tag{5.16}$$

provided  $m_{\mathcal{E},2} \in L^1(\log, \mathbb{R})$  and  $\mathcal{A}' \in L^1(\gamma, \mathbb{R}) := \{\varphi \in L^1(\mathbb{R}) \mid \int |\varphi(x)| (1 + |x|^\gamma) dx < \infty\}$ . In summary, from estimate (5.13), one infers by (5.12), (5.15), and (5.16) we infer

$$\|\nabla\phi\|_2 \leq C_1 \varepsilon^{3/2} + C_2 \varepsilon^{1-\gamma},$$

where  $\gamma > 0$  is arbitrary and the constants  $C_1$  and  $C_2$ , do not depend on  $\varepsilon$ . Hence, choosing  $\gamma \in (0, \frac{1}{2})$  one obtains the desired estimate of order  $\mathcal{O}(\varepsilon^{1/2})$ . For  $\gamma \in (0, \frac{1}{2})$  we get  $\|\nabla\phi\|_2 = \mathcal{O}(\varepsilon^{1-\gamma})$ .

Analogous estimates may be deduced for the first derivatives of  $\nabla\phi$  according to Theorem 2.2 if one prescribes further regularity on the functions  $\mathcal{A}$  and  $m_{\mathcal{E}}$ . For the sake of clarity, we collect the assumptions needed to estimate the quantities  $\|U_0\|_4, \|b\|_2, \|b\|_{L^1(\log)}$  in Table 1. For  $n, m \in \mathbb{N} \setminus \{0\}$  we denote  $W_{\pm}^{1,p}(\mathbb{R}^n)^m := \{\varphi \in L^p(\mathbb{R}^n)^m \mid \partial_k \varphi \in L^p(\mathbb{R}^n)^m \text{ for } k \geq 2, \partial_1 \varphi \in L^p(\mathbb{R}_{\pm}^n)^m\}$  and define the norm accordingly as  $\|\varphi\|_{W_{\pm}^{1,p}} := \sum_{k=2}^n \|\partial_k \varphi\|_p + \sum_{\pm} \|\partial_1 \varphi\|_{L^p(\mathbb{R}_{\pm}^n)}$ .

**Table 1**  
Sufficient regularity assumptions on  $\mathcal{A}$  and  $m_{\mathcal{E}}$  to estimate the terms on the first line.

$\ U_0\ _4$	$\ b\ _2$	$\ b\ _{\log}$	$\ U_0\ _{W_{\pm}^{1,4}}$	$\ \partial_2 b\ _2$
$\mathcal{A} \in L^4$	$\mathcal{A}' \in L^2$	$\mathcal{A}' \in L^1(\gamma)$	$\mathcal{A} \in W_{\pm}^{1,4}$	$\mathcal{A}' \in H^1$
$m_{\mathcal{E}} \in L^4$	$m_{\mathcal{E},2} \in L^2$	$m_{\mathcal{E},2} \in L^1(\log)$	$m_{\mathcal{E}} \in W_{\pm}^{1,4}$	$m_{\mathcal{E},2} \in L^2$

Summarising our results, we are led to the following.

**Proposition 5.1.** *Let  $f, \epsilon_1$ , and  $\epsilon_f$  be as in (5.1) and satisfy (H0)-(H1). Let  $U_0$  be as in (5.10) with  $\omega_0 \in \mathbb{R}, k_0 = k(\omega_0)$ , and with  $m_{\mathcal{E}}$  being the first two components of the eigenfunction corresponding to  $k_0$ . Suppose moreover  $\mathcal{A}$  and  $m_{\mathcal{E}}$  satisfy all the conditions in Table 1. Then there exists a function  $\phi \in \mathfrak{D}_{2,4}$  such that  $V = \nabla\phi$  is a solution of the problem (1.9)-(1.10)-(1.11) for  $n = 2$  with  $\|V\|_{W_{\pm}^{1,2}} \leq C\varepsilon^{1-\gamma}$  for any  $\gamma > 0$ , where (1.10) and the third interface condition in (1.11) hold in the sense of Definitions 2.2 and (2.1) resp. and the remaining interface conditions hold pointwise.*

**Remark 11.** Due to  $\|\nabla\phi\|_{L^2} = \mathcal{O}(\varepsilon^{1/2})$  the function  $\nabla\phi$  may be interpreted as a correction of the initial value  $U_0 = \mathcal{E}_{ans}^{(0)}$ .

**Proof.** It just remains to show that  $[\mathcal{E}_2^{(0)}] = 0$  since the third component of the electric field vanishes for TM-modes. This condition is already satisfied by  $U_0$  because of the choice of  $m_{\mathcal{E}}$ , so one needs to show that  $V_2 = \partial_2\phi$  is continuous across the interface. However, since  $\phi \in \mathfrak{D}_{2,4} \subset D_0^{1,4}(\mathbb{R}^2)$ , by (2.8) we know that

$\phi \in W_{loc}^{1,4}(\mathbb{R}^2)$ , so in particular  $\phi \in C_{loc}^{0,\frac{1}{2}}(\mathbb{R}^2)$  by the Morrey embedding. This means that  $\phi$  is continuous across the interface and, further, that every tangential derivative with respect to the interface is continuous too, see e.g. [48, Sec 173-175].  $\square$

**Remark 12.** It clearly follows from the analysis above and Remark 7 that it is possible to obtain analogous bounds with the same order in  $\varepsilon$  for the  $H^k$ -norm of  $\nabla\phi$ , provided higher regularity of  $\mathcal{A}$  in  $\mathbb{R}$  and  $m_\varepsilon$  in  $\mathbb{R}_\pm$  is prescribed.

**Remark 13.** Note that the exponential decay of  $m_\varepsilon$  shown by the numerics (Fig. 1(c)) indicates that the assumptions on  $m_\varepsilon$  in Table 1 are satisfied. In the application where  $\mathcal{A}$  is the initial condition of the envelope  $A$ , and hence free to choose, the conditions on  $\mathcal{A}$  in Table 1 are not critical.

**Remark 14.** Continuing the parallel with the case of a bounded domain, in which it is clear that  $\|b\|_2 = \mathcal{O}(\varepsilon^{3/2})$  would imply  $\|\nabla\phi\|_{H^1} = \mathcal{O}(\varepsilon^{3/2})$ , and hence  $\nabla\phi$  is of the same order in  $\varepsilon$  as the right-hand side, we expect that the estimates provided by Theorems 2.1-2.2 are not optimal. In our case  $\alpha = p = 3$  we can see that this is due to the additional term  $\|b\|$ , which produces the logarithmic term as  $n = 2$ , while all other terms on the right-hand side are indeed of order  $\mathcal{O}(\varepsilon^{3/2})$ . Thus it would be desirable to be able to estimate the term  $b$  differently than by Proposition 2.3.

### 5.3. Numerical test

To confirm the analytic results of Propositions 5.1, we calculate a numerical solution of problem (1.9)-(1.10)-(1.11) by applying a fixed point iteration to a finite element discretisation of the problem.

For a cubic nonlinearity, as described in the beginning of this chapter, we therefore have to solve

$$\begin{cases} -\nabla \cdot (\epsilon_1 \nabla \phi) = \mathfrak{f}(\phi) & \text{in } \mathbb{R}_\pm^2, \\ \llbracket \epsilon_1 \partial_1 \phi \rrbracket = \mathfrak{h}(\phi), \\ \llbracket \partial_2 \phi \rrbracket = 0, \end{cases} \tag{5.17}$$

with the  $\phi$ -dependent functions

$$\begin{aligned} \mathfrak{f}(\phi) &:= \nabla \cdot \left( \epsilon_1 U_0 + \epsilon_3 |U_0 + \nabla \phi|^2 (U_0 + \nabla \phi) \right), \\ \mathfrak{h}(\phi) &:= - \llbracket \epsilon_3 |U_0 + \nabla \phi|^2 (U_{0,1} + \partial_1 \phi) \rrbracket. \end{aligned}$$

To find a solution of (5.17), we rewrite the problem as a system of two coupled Neumann boundary value problems, in which we have to determine the functions  $\phi : \mathbb{R}_\pm^2 \rightarrow \mathbb{R}$  and  $\mathfrak{g} : \Gamma \rightarrow \mathbb{R}$  such that

$$\begin{cases} -\nabla \cdot (\epsilon_1 \nabla \phi) = \mathfrak{f}(\phi) & \text{in } \mathbb{R}_\pm^2, \\ (\epsilon_1 \partial_1 \phi)_- = \mathfrak{g}, \\ (\epsilon_1 \partial_1 \phi)_+ = \mathfrak{h}(\phi) + \mathfrak{g}, \\ \llbracket \partial_2 \phi \rrbracket = 0, \end{cases} \tag{5.18}$$

where we recall  $u_\pm(x) := \lim_{h \rightarrow 0^\pm} u(x + he_1)$ , for  $x \in \Gamma$ . Note that a solution  $\phi$  of (5.18) is also a solution of (5.17).

We now approximate the solution of the nonlinear problem (5.18) with the help of a fixed point iteration. We select an initial guess  $\phi_0$  and solve

$$\begin{cases} -\nabla \cdot (\epsilon_1 \nabla \phi_{n+1}) = \mathfrak{f}(\phi_n) & \text{in } \mathbb{R}_{\pm}^2, \\ (\epsilon_1 \partial_1 \phi_{n+1})_- = \mathfrak{g}_{n+1}, \\ (\epsilon_1 \partial_1 \phi_{n+1})_+ = \mathfrak{h}(\phi_n) + \mathfrak{g}_{n+1}, \\ \llbracket \partial_2 \phi_{n+1} \rrbracket = 0 \end{cases} \tag{5.19}$$

iteratively for  $n \geq 0$ . The weak formulation of the problem is given by

$$\begin{cases} \int_{\mathbb{R}_{\pm}^2} \epsilon_1 \nabla \phi_{n+1} \cdot \nabla \eta + \int_{\Gamma} \mathfrak{g}_{n+1} \eta = \int_{\mathbb{R}_{\pm}^2} \mathfrak{f}(\phi_n) \eta - \int_{\Gamma} \mathfrak{h}(\phi_n) \eta, & \eta \in H^1(\mathbb{R}_{\pm}^2), \\ \int_{\mathbb{R}_{\pm}^2} \epsilon_1 \nabla \phi_{n+1} \cdot \nabla \eta - \int_{\Gamma} \mathfrak{g}_{n+1} \eta = \int_{\mathbb{R}_{\pm}^2} \mathfrak{f}(\phi_n) \eta, & \eta \in H^1(\mathbb{R}_{\mp}^2), \\ \llbracket \partial_2 \phi_{n+1} \rrbracket = 0. \end{cases} \tag{5.20}$$

To solve (5.20) numerically, we use the finite element method. First we replace  $\mathbb{R}_{\pm}^2$  and  $\Gamma$  by suitable bounded domains  $\Omega_{\pm} \subset \overline{\mathbb{R}_{\pm}^2}$  and  $\tilde{\Gamma} := \overline{\Omega}_+ \cap \overline{\Omega}_- \subset \Gamma$ , respectively. Furthermore we substitute  $H^1(\mathbb{R}_{\pm}^2)$  with the following  $N$ -dimensional subspaces  $\mathcal{V}_{\pm} := \text{span}\{\eta_k^{\pm} \mid k = 1, \dots, N\}$ ,  $N \in \mathbb{N}$ , where the shape functions  $\eta_k^{\pm} \in H^1(\Omega_{\pm})$  are the standard piecewise linear hat functions, which are linearly independent. Then we look for solutions of the form

$$\begin{aligned} \phi_{n+1}(x) &= \begin{cases} \sum_k \Phi_{n+1,k}^+ \eta_k^+(x) & \text{if } x \in \Omega_+, \\ \sum_k \Phi_{n+1,k}^- \eta_k^-(x) & \text{if } x \in \Omega_-, \end{cases} \\ \mathfrak{g}_{n+1}(x) &= \sum_k G_{n+1,k} \eta_k^+(x)|_{\tilde{\Gamma}}, \end{aligned}$$

where the coefficients  $\Phi_{n+1,k}^{\pm}, G_{n+1,k}$  are the solutions of the following system

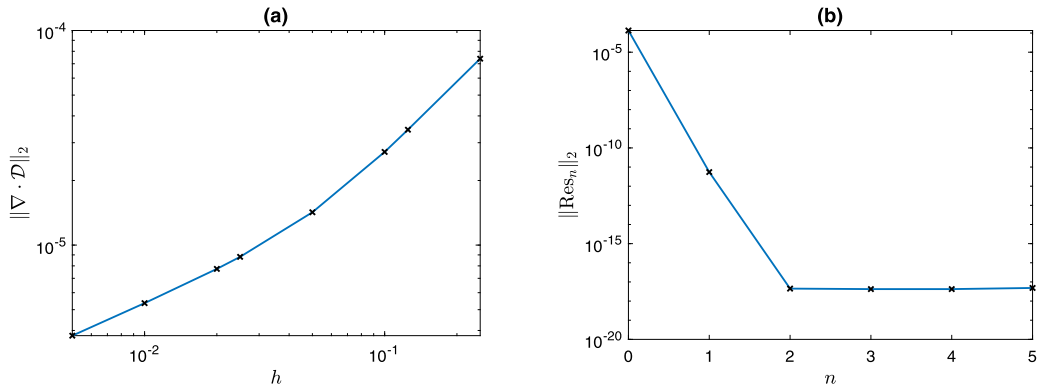
$$\begin{cases} \sum_k \left( \Phi_{n+1,k}^+ \int_{\Omega_+} \epsilon_1 \nabla \eta_j^+ \cdot \nabla \eta_k^+ + G_{n+1,k} \int_{\tilde{\Gamma}} \eta_j^+ \eta_k^+ \right) = \int_{\Omega_+} \mathfrak{f}(\phi_n) \eta_j^+ - \int_{\tilde{\Gamma}} \mathfrak{h}(\phi_n) \eta_j^+, & j = 1, \dots, N, \\ \sum_k \left( \Phi_{n+1,k}^- \int_{\Omega_-} \epsilon_1 \nabla \eta_j^- \cdot \nabla \eta_k^- - G_{n+1,k} \int_{\tilde{\Gamma}} \eta_j^- \eta_k^+ \right) = \int_{\Omega_-} \mathfrak{f}(\phi_n) \eta_j^-, & j = 1, \dots, N, \\ \sum_k \left( \Phi_{n+1,k}^- \partial_2 \eta_k^-|_{\tilde{\Gamma}} - \Phi_{n+1,k}^+ \partial_2 \eta_k^+|_{\tilde{\Gamma}} \right) = 0. \end{cases}$$

Due to the Neumann boundary conditions and the fact that  $\mathfrak{f}$  and  $\mathfrak{h}$  depend only on the gradient of  $\phi$ , the solution is unique only up to an additive constant. To get uniqueness we additionally demand  $\int_{\Omega_{\pm}} \phi_{n+1} = 0$ . Therefore, we extend the finite element formulation by

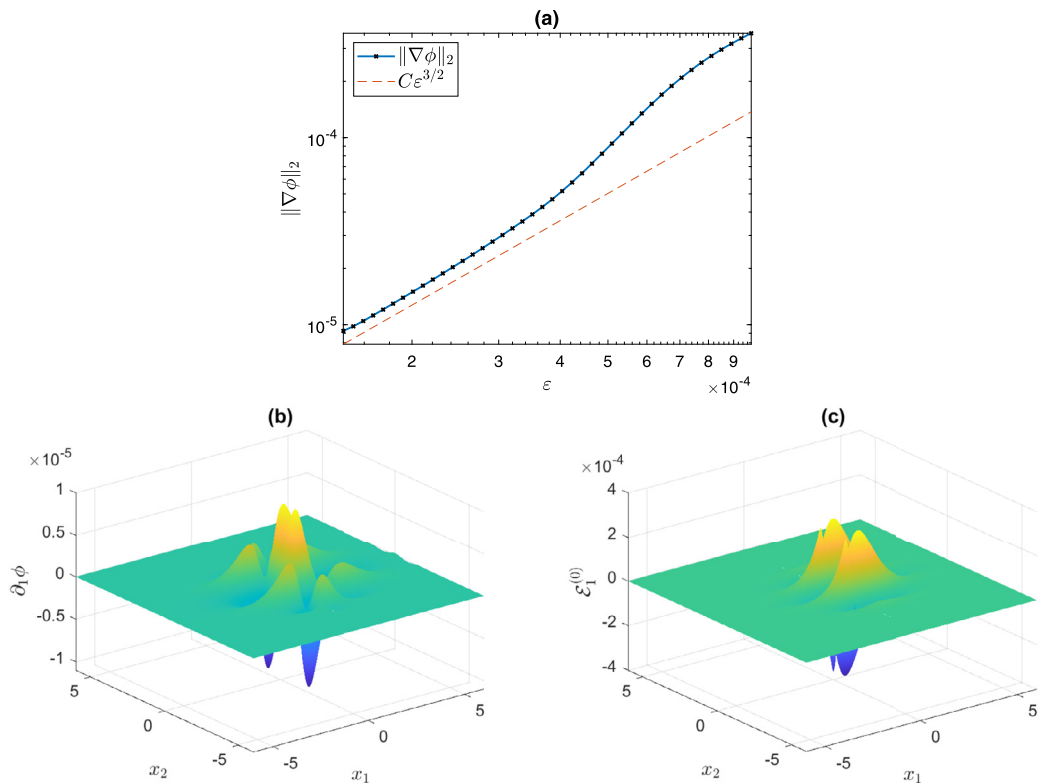
$$\sum_k \Phi_{n+1,k}^{\pm} \int_{\Omega_{\pm}} \eta_k^{\pm} = 0.$$

Let us now consider the example introduced in Section 5.1 with the main goal of verifying Proposition 5.1 and in particular the conjectured convergence rate  $\mathcal{O}(\varepsilon^{3/2})$  for  $\|\nabla \phi\|_2$  in Remark 14. Due to the fact that we choose piecewise linear shape functions we can only study the convergence of the  $L^2$ -norm of  $\nabla \phi$  although Proposition 5.1 estimates even the  $H^1$ -norm. For our numerical test we choose  $\epsilon_1(x_1) = 1\chi_{\mathbb{R}_-} + (1 + e^{-x_1})\chi_{\mathbb{R}_+}$  and  $m_{\varepsilon}$  accordingly (see Fig. 1)  $\epsilon_3 \equiv 1$ , and we set  $\mathcal{A}(x_2) = e^{-5 \cdot 10^6 \varepsilon^2 x_2^2}$ . Note that with such choices of (exponentially decaying)  $m_{\varepsilon}$  and  $\mathcal{A}$ , the conditions for  $U_0$  and  $b$  in Table 1 are satisfied. For the discretisation we select  $\overline{\Omega}_- = [-6, 0] \times [-6, 6]$  and  $\overline{\Omega}_+ = [0, 6] \times [-6, 6]$  and choose a regular triangulation of step size  $h$  together with standard hat functions for  $\eta_k^{\pm}$ . At those boundaries of  $\Omega_{\pm}$  which are not part of  $\tilde{\Gamma}$  we enforce homogeneous Neumann boundary conditions.

For the fixed point iteration we start with  $\phi_0 \equiv 0$  as an initial guess. Let us first check the convergence of the discretisation (in  $h$  and in the iteration  $n$ ). For Fig. 2 (a) we fixed  $\varepsilon = 3 \cdot 10^{-4}$  and calculated  $\|\nabla \cdot \mathcal{D}\|_2$



**Fig. 2.** (a)  $\|\nabla \cdot \mathcal{D}\|_2$  in dependence on the step size  $h$ . (b) Plot of the  $L^2$ -norm of the residual in the fixed point iteration for each step.



**Fig. 3.** (a) The graph shows that numerically  $\|\nabla\phi\|_2 = \mathcal{O}(\varepsilon^{3/2})$ . (b) Plot of  $\partial_1\phi$  for  $\varepsilon = 3 \cdot 10^{-4}$ . (c) The first component of the corrected initial value  $\mathcal{E}_1^{(0)} = U_{0,1} + \partial_1\phi$  for  $\varepsilon = 3 \cdot 10^{-4}$ . Note the different scales in (b) and (c).

for different step sizes  $h$  ranging from 0.25 to 0.005. For Fig. 2 (b) we also fixed  $h = 0.005$  and calculated the  $L^2$ -norm of the residual  $\text{Res}_n := -\nabla \cdot (\varepsilon_1 \nabla \phi_n) - f(\phi_n)$  in each step of the fixed point iteration. We see the numerical convergence in both plots.

Finally, we study the  $\varepsilon$ -convergence of  $\|\nabla\phi\|_2$ . For the fixed step size  $h = 0.005$  and  $\varepsilon$  ranging from  $10^{-4}$  to  $10^{-3}$  we obtain the desired rate of convergence, see Fig. 3 (a). For  $\varepsilon = 3 \cdot 10^{-4}$  Figs. 3 (b) and (c) show the first components the computed solutions  $\nabla\phi$  and  $\mathcal{E}^{(0)}$ , respectively.



## 6. Generalisations and open problems

Let us finally discuss possible generalisations and questions which are left unanswered by our analysis and which we believe to be of interest.

**Curved interfaces** The interface  $\Gamma$  separating the two materials was chosen straight merely for reasons of simplicity. The results can be directly generalised also for a smooth curved interface separating  $\mathbb{R}^n$  into two unbounded subsets.

**Positivity of  $\epsilon_1, \epsilon_f$**  The positivity imposed on the coefficients  $\epsilon_1, \epsilon_f$  in (H0) is a common assumptions in the analysis literature because the operator involved in equation (1.12a) is then of elliptic type. We are not aware of results in the literature dealing with quasilinear operators with coefficients of varying sign.

However, in the electromagnetic context materials with negative material functions exist. For the non-linear coefficient  $\epsilon_f$  it is the case, e.g., for Gallium compounds [44], and for metamaterials both  $\epsilon_1$  and  $\epsilon_f$  may be negative. We have also seen in Section 5.1 that in the simplest case of a piecewise constant  $\epsilon_1$ , in order to obtain the existence of a TM eigenfunction, opposite signs of  $\epsilon_1$  in the two subspaces need to be imposed, see the dispersion relation (5.7). Hence, if the analytical results could be extended to the case of nonpositive coefficients, one may obtain a result similar to Proposition 5.1 for a wider range of materials.

We believe that overcoming this sign problem would be of great interest both for the analysis and the physics application.

**Suboptimality of the asymptotics** As we already mentioned in Remark 14, we expect that the correction  $\nabla\phi$  of  $U_0$  is asymptotically of order  $\mathcal{O}(\varepsilon^{3/2})$  in the  $H^1$ -norm. For its  $L^2$ -norm, this behaviour has been also confirmed by our numerical test described in Section 5.3, see Fig. 3(a). The suboptimality of the asymptotic estimate in Proposition 5.1 ultimately relies on the term  $\|b\|$  in (2.5), since it produces the logarithmic term that we estimate by (5.16). All other terms produced when applying Theorem 2.1 to our Maxwell context are of the expected order  $\mathcal{O}(\varepsilon^{3/2})$ , see Section 5.2. In order to get a sharp result we would need to get rid of  $\|b\|$  in (2.5) and therefore to be able to deal with the product  $\int b\phi$  differently than by (3.5), i.e. than by means of Proposition 2.3.

Note that, once such an issue is solved for  $\|\nabla\phi\|_2$ , the problem of improving also the asymptotic estimates for the higher Sobolev norms does not arise. Indeed, the latter involve only terms which already bound  $\|\nabla\phi\|_2$ , and terms which are asymptotically of order  $\mathcal{O}(\varepsilon^{3/2})$ .

## Acknowledgments

This research is supported by the German Research Foundation, DFG grant No. DO 1467/4-1.

## References

- [1] N. Aronszajn, E. Gagliardo, Interpolation spaces and interpolation methods, *Ann. Mat.* 68 (1965) 51–117.
- [2] C. Atkinson, C.R. Champion, A Mode III crack at the interface between two nonlinear materials, *Proc. R. Soc. Lond. Ser. A, Math. Phys. Sci.* 429 (1876) (1990) 247–257.
- [3] A. Azzollini, P. d’Avenia, A. Pomponio, Quasilinear elliptic equations in  $\mathbb{R}^N$  via variational methods and Orlicz-Sobolev embeddings, *Calc. Var. Partial Differ. Equ.* 49 (1–2) (2014) 197–213.
- [4] M. Badiale, G. Citti, Concentration compactness principle and quasilinear elliptic equations in  $\mathbb{R}^n$ , *Commun. Partial Differ. Equ.* 16 (11) (1991) 1795–1818.
- [5] S. Barile, G.M. Figueiredo, Existence of least energy positive, negative and nodal solutions for a class of  $p$ -& $q$ -problems with potentials vanishing at infinity, *J. Math. Anal. Appl.* 427 (2) (2015) 1205–1233.
- [6] R. Bartolo, A.M. Candela, A. Salvatore, On a class of superlinear  $(p, q)$ -Laplacian type equations on  $\mathbb{R}^N$ , *J. Math. Anal. Appl.* 438 (1) (2016) 29–41.
- [7] M. Bathory, M. Bulíček, O. Souček, Existence and qualitative theory for nonlinear elliptic systems with a nonlinear interface condition used in electrochemistry, *Z. Angew. Math. Phys.* 71 (3) (2020) 74, 24 pp.
- [8] V. Benci, D. Fortunato, L. Pisani, Soliton like solutions of a Lorentz invariant equation in dimension 3, *Rev. Math. Phys.* 10 (3) (1998) 315–344.

- [9] N. Benouhiba, Z. Belyacine, A class of eigenvalue problems for the  $(p, q)$ -Laplacian in  $\mathbb{R}^N$ , *Int. J. Pure Appl. Math.* 80 (5) (2012) 727–737.
- [10] H. Berestycki, F. Hamel, L. Roques, Analysis of the periodically fragmented environment model. I. Species persistence, *J. Math. Biol.* 51 (1) (2005) 75–113.
- [11] M. Borsuk, *Transmission Problems for Elliptic Second-Order Equations in Non-smooth Domains*, *Frontiers in Mathematics*, Birkhäuser/Springer, Basel AG, Basel, 2010.
- [12] M. Brown, T. Dohnal, M. Plum, I. Wood, Spectrum of the Maxwell equations for the flat interface between homogeneous dispersive media, in preparation.
- [13] Zh. Chen, J. Zou, Finite element methods and their convergence for elliptic and parabolic interface problems, *Numer. Math.* 79 (2) (1998) 175–202.
- [14] P. Colli Franzone, L. Guerri, S. Rovida, Wavefront propagation in an activation model of the anisotropic cardiac tissue: asymptotic analysis and numerical simulations, *J. Math. Biol.* 28 (1990) 121–176.
- [15] M. Costabel, M. Dauge, S. Nicaise, Singularities of Maxwell interface problems, *Math. Model. Numer. Anal.* 33 (3) (1999) 627–649.
- [16] T. Dohnal, D. Rudolf, NLS approximation for wavepackets in periodic cubically nonlinear wave problems in  $\mathbb{R}^d$ , *Appl. Anal.* 99 (10) (2020) 1685–1723.
- [17] C. Ebmeyer, J. Frehse, M. Kassmann, Boundary regularity for nonlinear elliptic systems: applications to the transmission problem, in: S. Hildebrandt, H. Karcher (Eds.), *Geometric Analysis and Nonlinear Partial Differential Equations*, Springer Verlag, 2002, pp. 505–517.
- [18] R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures on Physics. Vol. 2: Mainly Electromagnetism and Matter*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1964.
- [19] G.P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations. Steady-State Problems*, second edition, *Springer Monographs in Mathematics*, Springer, New York, 2011.
- [20] E. Giusti, *Direct Methods in the Calculus of Variations*, World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
- [21] P. Kirrmann, G. Schneider, A. Mielke, The validity of modulation equations for extended systems with cubic nonlinearities, *Proc. R. Soc. Edinb., Sect. A* 122 (1–2) (1992) 85–91.
- [22] D. Knees, On the regularity of weak solutions of quasi-linear elliptic transmission problems on polyhedral domains, *Z. Anal. Anwend.* 23 (3) (2004) 509–546.
- [23] D. Knees, A.-M. Sändig, *Regularity of Elastic Fields in Composites. Multifield Problems in Solid and Fluid Mechanics*, *Lect. Notes Appl. Comput. Mech.*, vol. 28, Springer, Berlin, 2006, pp. 331–360.
- [24] T. Kurokawa, Riesz potentials, higher Riesz transforms and Beppo Levi spaces, *Hiroshima Math. J.* 18 (3) (1988) 541–597.
- [25] N. Kutev, P.-L. Lions, Nonlinear second-order elliptic equations with jump discontinuous coefficients. Part I, quasilinear equations, *Differ. Integral Equ.* 5 (6) (1992) 1201–1217.
- [26] O.A. Ladyzhenskaja, V.A. Solonnikov, N.N. Ural'tseva, *Linear and Quasilinear Equations of Parabolic Type*, *Translations of Mathematical Monographs*, vol. 23, American Mathematical Society, Providence, R.I., 1968 (in Russian). Translated from the Russian by S. Smith.
- [27] O.A. Ladyzhenskaja, V.Ja. Rivkind, N.N. Ural'tseva, Solvability of diffraction problems in the classical sense, *Tr. Mat. Inst. Steklova* 92 (1966) 116–146 (in Russian).
- [28] O.A. Ladyzhenskaya, N.N. Ural'tseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York-London, 1968.
- [29] G. Li, S.R. Seshadri, Weakly nonlinear surface polariton, *J. Opt. Soc. Am. B* 6 (1989) 1125–1137.
- [30] G. Li, X. Liang, The existence of nontrivial solutions to nonlinear elliptic equation of  $(p, q)$ -Laplacian type on  $\mathbb{R}^N$ , *Nonlinear Anal.* 71 (5–6) (2009) 2316–2334.
- [31] W.B. Liu, Degenerate quasilinear elliptic equations arising from bimaterial problems in elastic-plastic mechanics, *Nonlinear Anal., Theory Methods Appl.* 35 (4) (1999) 517–529.
- [32] A. Lorenzi, On elliptic equations with piecewise constant coefficients II, *Ann. Sc. Norm. Super. Pisa, Cl. Sci.* (3) 26 (4) (1972) 839–870.
- [33] P. Marcellini, Regularity and existence of solutions of elliptic equations with  $p, q$ -growth conditions, *J. Differ. Equ.* 90 (1) (1991) 1–30.
- [34] P. Marcellini, Growth conditions and regularity for weak solutions to nonlinear elliptic pdes, *J. Math. Anal. Appl.* 501 (1) (2021).
- [35] A. Marini, D.V. Skryabin, Ginzburg-Landau equation bound to the metal-dielectric interface and transverse nonlinear optics with amplified plasmon polaritons, *Phys. Rev. A* 81 (2010) 033850.
- [36] G. Mingione, Regularity of minima: an invitation to the dark side of the calculus of variations, *Appl. Math.* 51 (4) (2006) 355–426.
- [37] G. Mingione, V. Rădulescu, Recent developments in problems with nonstandard growth and nonuniform ellipticity, *J. Math. Anal. Appl.* 501 (1) (2021).
- [38] J.E. Muñoz Rivera, H. Portillo Oquendo, A transmission problem for thermoelastic plates, *Q. Appl. Math.* 62 (2) (2004) 273–293.
- [39] S. Nicaise, A.-M. Sändig, Transmission problems for the Laplace and elasticity operators. Regularity and boundary integral formulation, *Math. Methods Appl. Sci.* 9 (1999) 855–898.
- [40] O.A. Oleĭnik, Boundary-value problems for linear equations of elliptic parabolic type with discontinuous coefficients, *Izv. Akad. Nauk SSSR, Ser. Mat.* 25 (1961) 3–20 (in Russian).
- [41] C. Ortner, E. Süli, A note on linear elliptic systems on  $\mathbb{R}^d$ , preprint, arXiv:1202.3970, 2012.
- [42] C.G. Simader, The weak Dirichlet and Neumann problem for the Laplacian in  $L^q$  for bounded and exterior domains, in: M. Krbeč, A. Kufner, B. Opic, J. Rákosník (Eds.), *Nonlinear Analysis, Function Spaces and Applications*, in: *Teubner-Texte zur Mathematik*, vol. 4, Vieweg+Teubner Verlag, Wiesbaden, 1990.

- [43] R. Schnaubelt, M. Spitz, Local wellposedness of quasilinear Maxwell equations with conservative interface conditions, preprint, arXiv:1811.08714, 2018.
- [44] R.L. Sutherland, Handbook of Nonlinear Optics. Optical Science and Engineering, Taylor & Francis, 2003.
- [45] Q.J. Tan, Z.J. Leng, The method of upper and lower solutions for diffraction problems of quasilinear elliptic reaction-diffusion systems, *J. Math. Anal. Appl.* 380 (1) (2011) 363–376.
- [46] T. Tao, Nonlinear dispersive equations: local and global analysis, *Am. Math. Soc.* 106 (2006).
- [47] D.P. Tietz, Justification of the Nonlinear Schrödinger Equation for Interface Wave Packets in Maxwell's Equations with 2D Localization, Ph.D. Dissertation, Martin-Luther Universität, Halle-Wittenberg, in preparation.
- [48] C. Truesdell, R. Toupin, The classical field theories, in: S. Flügge (Ed.), *Handbuch der Physik*, Bd. III/1, Springer, Berlin, 1960. With an appendix on tensor fields by J. L. Ericksen.
- [49] W-Ch. Wang, A jump condition capturing finite difference scheme for elliptic interface problems, *SIAM J. Sci. Comput.* 25 (5) (2004) 1479–1496.
- [50] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, *Trans. Am. Math. Soc.* 36 (1934) 63–89.