

# Rigorous Envelope Approximation for Interface Wave-Packets in Maxwell's Equations with 2D Localization

Tomáš Dohnal

Institute of Mathematics, Martin Luther University Halle-Wittenberg,  
06099 Halle (Saale), Germany. Email: tomas.dohnal@mathematik.uni-halle.de

Roland Schnaubelt

Department of Mathematics, Karlsruhe Institute of Technology,  
76128 Karlsruhe, Germany. Email: schnaubelt@kit.edu

Daniel P. Tietz

Institute of Mathematics, Martin Luther University Halle-Wittenberg,  
06099 Halle (Saale), Germany. Email: daniel.tietz@mathematik.uni-halle.de

June 1, 2022

## Abstract

We study transverse magnetic (vector valued) wave-packets in the time dependent Kerr nonlinear Maxwell's equations at the interface of two inhomogeneous dielectrics with an instantaneous material response. The resulting model is quasilinear. The problem is solved on each side of the interface and the fields are coupled via natural interface conditions. The wave-packet is localized at the interface and propagates in the tangential direction. For a slowly modulated envelope approximation the nonlinear Schrödinger equation is formally derived as an amplitude equation for the envelope. We rigorously justify the approximation in a Sobolev space norm on the corresponding asymptotically large time intervals. The well-posedness result for the quasilinear Maxwell problem builds on the local theory of [R. Schnaubelt und M. Spitz, Local wellposedness of quasilinear Maxwell equations with conservative interface conditions, *Commun. Math. Sci.*, accepted, 2022] and extends this to asymptotically large time intervals for small data using an involved bootstrapping argument.

**2020 MSC:** 35Q61, 35C07, 35L50.

**Key words:** Maxwell's equations, Kerr nonlinearity, quasilinear, interface, envelope approximation, traveling pulse.

## 1 Introduction

Propagation of electromagnetic wave-packets at interfaces is of interest for applications in modern and future optical components. A typical example are surface plasmons (SPs) at the interface of a conductor and a dielectric. From an engineering point of view their advantage is mainly in the higher level of localization compared to wave-packets in bulk media. If the involved media feature a nonlinear material response, new phenomena and thus new functionality of SPs are produced [17]. Also interfaces of two dielectrics, e.g. photonic crystal waveguides [14, 19, 29], are interesting from the applied point of view.

We consider wave-packets at the interface of two generally inhomogeneous and Kerr nonlinear dielectrics. The problem is modeled by time dependent quasilinear Maxwell's equations in two spatial dimensions (assuming homogeneity of the material in the third direction). We study spatio-temporal wave-packets propagating in the direction tangential to the interface. They are broad in the propagation direction, have a small amplitude and are slowly modulated in time. In particular, we analyze their approximation via a slowly varying envelope. The equation governing the envelope dynamics is the one dimensional nonlinear Schrödinger equation (NLS).

The NLS is well known to approximate the dynamics of wave-packets in dispersive problems with a single carrier wave [15, 18]. The formal derivation of the NLS for spatio-temporal wave-packets in quasilinear Maxwell's equations exists for a number of scenarios ranging from pulses in optical fibers [1] over photonic crystals [2] to surface plasmons, e.g. [21, 10]. Spatial (time harmonic) surface plasmon wave-packets have been also formally approximated by the NLS (or more generally by the complex Ginzburg-Landau equation), e.g. in [9, 22, 7], where one of the spatial variables plays the role of an evolution variable.

Formally derived asymptotic models for wave-packet envelopes can fail to produce a valid approximation of the original nonlinear problem, see e.g. [24, 25]. Hence, a rigorous justification analysis with an error estimate must be performed. In the case of the semilinear wave equation with periodic coefficients this justification was carried out in [5] for the one dimensional case and in [12] in  $d$  dimensions. The time dependent Maxwell's equations for nonlinear materials are quasilinear and the NLS approximation has been justified only in the case of fields leading to a scalar equation. Examples are [20] and [26]. In [20] wave-packets in 2D photonic crystal waveguides are described by a quasilinear wave equation. In [26] an approximation via a complex Ginzburg-Landau equation is proved for the quasilinear wave equation in one spatial dimension and with a time delayed material response (memory effect modeled by the coupling to an ODE system).

We work with vector valued Maxwell's equations for TM polarized fields and reduce the problem to a system for the components  $E_1, E_2$  and  $H_3$ . We restrict here to the instantaneous material response (as relevant for dielectrics) and avoid time delayed terms (relevant for metals). To our knowledge there are no directly applicable results on long time existence in full quasilinear Maxwell's equations with time delay on unbounded domains. For the instantaneous case we use local well-posedness results from [23]. The working function space for each component is  $\mathcal{G}^3(\mathbb{R}^2 \times J) := \bigcap_{j=0}^3 C^j(\bar{J}, \mathcal{H}^{3-j}(\mathbb{R}^2))$ , where  $\mathcal{H}^s(\mathbb{R}^2)$  consists of functions which are  $H^s$  on each half-space defined by the interface and  $J$  is a time interval. This high regularity is enforced by the quasilinear term. The two half-spaces are coupled by interface conditions out of which only the continuity of  $E_2$  and  $H_3$  needs to be enforced in the time evolution. As the approximation result needs to be proved on asymptotically large time intervals, we use a sophisticated bootstrap argument to extend the local existence to such asymptotically long time scales for small data. Here one estimates space-time differentiated solutions of the problem. If no normal derivatives occur, we can apply higher-order energy inequalities. Otherwise, the boundary conditions are violated and we have to use the equation itself and the divergence condition iteratively to bound the differentiated solutions in Gronwall arguments, see Section 6.1.

Maxwell's equations in the whole space in the absence of free currents read

$$\begin{cases} \partial_t \mathcal{D} = \nabla \times \mathcal{H}, \\ \mu_0 \partial_t \mathcal{H} = -\nabla \times \mathcal{E}, \\ \nabla \cdot \mathcal{D} = \varrho_0, \\ \nabla \cdot \mathcal{H} = 0, \end{cases} \quad (1)$$

for  $\mathbf{x} \in \mathbb{R}^3$  and  $t > 0$ , where  $\mathcal{H} = \mathcal{H}(\mathbf{x}, t)$  is the magnetic field,  $\mathcal{E} = \mathcal{E}(\mathbf{x}, t)$  is the electric field,  $\varrho_0 = \varrho_0(\mathbf{x})$  is the volume charge density and  $\mu_0$  is the permeability of free space. We consider an electric displacement field  $\mathcal{D} = \mathcal{D}(\mathcal{E})$  given by the instantaneous material law

$$\mathcal{D}(\mathbf{x}, t) = \epsilon_0 \mathcal{E}(\mathbf{x}, t) + \mathcal{P}(\mathbf{x}, t). \quad (2)$$

Here  $\epsilon_0$  is the permittivity of free space and  $\mathcal{P} = \mathcal{P}(\mathcal{E})$  is the electric polarization modeling an  $x_1$ -

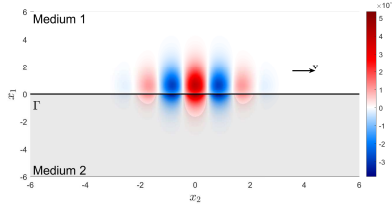


Figure 1: Schematic of a pulse propagating in the direction  $\mathbf{v} = (0, 1)^\top$ , i.e., along the interface.

dependent Kerr nonlinear material, i.e.,

$$\mathcal{P}(\mathbf{x}, t) = \epsilon_0 (\chi_1(x_1) \mathcal{E}(\mathbf{x}, t) + \chi_3(x_1) (\mathcal{E}(\mathbf{x}, t) \cdot \mathcal{E}(\mathbf{x}, t)) \mathcal{E}(\mathbf{x}, t)) \quad (3)$$

with the linear and cubic susceptibilities  $\chi_1, \chi_3 : \mathbb{R} \rightarrow \mathbb{R}$ . For simplicity, the susceptibilities are scalar quantities, meaning that the material is isotropic.

In order to describe an interface, we allow  $\chi_1$  and  $\chi_3$  to have a jump at  $x_1 = 0$  and denote

$$\chi_1(x_1) = \begin{cases} \chi_1^-(x_1), & x_1 < 0, \\ \chi_1^+(x_1), & x_1 > 0, \end{cases} \quad \chi_3(x_1) = \begin{cases} \chi_3^-(x_1), & x_1 < 0, \\ \chi_3^+(x_1), & x_1 > 0, \end{cases}$$

for  $\chi_1^\pm, \chi_3^\pm : \mathbb{R}_\pm := \{x_1 \in \mathbb{R} \mid \pm x_1 > 0\} \rightarrow \mathbb{R}$ . We also define

$$\epsilon_1 := \epsilon_0(1 + \chi_1), \quad \epsilon_3 := \epsilon_0\chi_3 \quad \text{and} \quad \epsilon_1^\pm := \epsilon_0(1 + \chi_1^\pm), \quad \epsilon_3^\pm := \epsilon_0\chi_3^\pm.$$

We investigate a two dimensional setting with all fields independent of  $x_3$ , i.e.,

$$(\mathcal{D}, \mathcal{E}, \mathcal{H}, \mathcal{P}) = (\mathcal{D}, \mathcal{E}, \mathcal{H}, \mathcal{P})(x_1, x_2, t).$$

Hence, the problem can be reduced to  $\mathbb{R}^2$ . From now on the variable  $\mathbf{x}$  lies in  $\mathbb{R}^2$ . The two resulting half-spaces are denoted by  $\mathbb{R}_-^2 := \{\mathbf{x} \in \mathbb{R}^2 : x_1 < 0\}$  and  $\mathbb{R}_+^2 := \{\mathbf{x} \in \mathbb{R}^2 : x_1 > 0\}$  and the interface is  $\Gamma := \{\mathbf{x} \in \mathbb{R}^2 : x_1 = 0\}$ . The aim of this paper is to describe the propagation of wave-packets localized near the interface  $\Gamma$  and propagating in the  $x_2$ -direction, see Figure 1.

We introduce also the (time independent) surface charge density  $\varrho_\Gamma : \Gamma \rightarrow \mathbb{R}$ . Using Maxwell's equations in integral form, one can formally derive the jump conditions for solutions

$$\begin{aligned} \llbracket \mathcal{D}_1 \rrbracket(\mathbf{x}, t) &= \varrho_\Gamma(\mathbf{x}), \quad \llbracket \mathcal{E}_2 \rrbracket(\mathbf{x}, t) = \llbracket \mathcal{E}_3 \rrbracket(\mathbf{x}, t) = 0, \quad \forall \mathbf{x} \in \Gamma, t \geq 0, \\ \llbracket \mathcal{H}_1 \rrbracket(\mathbf{x}, t) &= \llbracket \mathcal{H}_2 \rrbracket(\mathbf{x}, t) = \llbracket \mathcal{H}_3 \rrbracket(\mathbf{x}, t) = 0, \quad \forall \mathbf{x} \in \Gamma, t \geq 0, \end{aligned} \quad (4)$$

in the absence of surface currents, see also Section I.4.2.4 of [8]. Here for  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  continuous on  $\overline{\mathbb{R}_-^2}$  and  $\overline{\mathbb{R}_+^2}$  and for each  $\mathbf{x} \in \Gamma$  we define

$$\llbracket f \rrbracket(\mathbf{x}) := \lim_{\mathbf{y} \rightarrow \mathbf{x}, \mathbf{y} \in \overline{\mathbb{R}_+^2}} f(\mathbf{y}) - \lim_{\mathbf{y} \rightarrow \mathbf{x}, \mathbf{y} \in \overline{\mathbb{R}_-^2}} f(\mathbf{y}).$$

For  $\varrho_\Gamma = 0$  conditions (4) hold in trace sense for any  $\mathcal{E}(\cdot, t) \in H_{\text{curl}}(\mathbb{R}^2)$ ,  $\mathcal{D}(\cdot, t) \in H_{\text{div}}(\mathbb{R}^2)$  and  $\mathcal{H}(\cdot, t) \in H^1(\mathbb{R}^2)$ , see e.g. an appendix in [4].

We study specific solutions satisfying the reduction

$$\mathcal{E}_3 \equiv \mathcal{H}_1 \equiv \mathcal{H}_2 \equiv 0. \quad (5)$$

For the case that  $\epsilon_1^\pm$  is constant, all eigenfunctions of the linear eigenvalue problem for time harmonic fields (namely (15)) have to satisfy (5), see e.g. [4]. We study this reduced type of solutions also in the nonlinear case with non-constant  $\epsilon_1^\pm$ . Hence, we set

$$\mathbf{U} := (\mathcal{E}_1, \mathcal{E}_2, \mathcal{H}_3)^\top \quad (6)$$

and further define

$$\mathbf{U}^\pm := \mathbf{U}|_{\mathbf{x} \in \mathbb{R}_\pm^2}.$$

Throughout this paper we will always use  $f^\pm$  to indicate the restriction of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  to  $\mathbb{R}_\pm^n$  with  $n = 1, 2$ . We also write  $\mathbf{U}_E := (U_1, U_2, 0)^\top$  to denote the part of  $\mathbf{U}$  corresponding to the electric field. With (6) the first two equations in (1) reduce to a system of three instead of six scalar equations. Note that the problem is indeed compatible with this reduction since the form of the nonlinearity implies that  $\mathcal{D}_j \equiv 0$  if  $\mathcal{E}_j \equiv 0$ .

Regarding the interface conditions, note that for time-independent surface charges we have  $\llbracket \mathcal{D}_1 \rrbracket(\mathbf{x}, t) = \varrho_\Gamma(\mathbf{x})$  for all  $t > 0$  if  $\llbracket \mathcal{D}_1 \rrbracket(\mathbf{x}, 0) = \varrho_\Gamma(\mathbf{x})$ . This can be derived from the first component of the first equation in (1). Indeed, we get  $\partial_t \llbracket \mathcal{D}_1 \rrbracket(\mathbf{x}, t) = \partial_{x_2} \llbracket \mathcal{H}_3 \rrbracket(\mathbf{x}, t) = 0$  for all  $\mathbf{x} \in \Gamma$  and  $t > 0$ . Also the divergence condition  $\nabla \cdot \mathcal{D} = \varrho_0$  needs to be checked only at  $t = 0$  as follows from the first equation in (1). For our specific solutions  $\nabla \cdot \mathcal{H} = 0$  is always satisfied, since the only non-trivial component  $\mathcal{H}_3$  is independent of  $x_3$ . Therefore, the equations  $\llbracket \mathcal{D}_1 \rrbracket = \varrho_\Gamma$ ,  $\nabla \cdot \mathcal{D} = \varrho_0$ , and  $\nabla \cdot \mathcal{H} = 0$  play no role in our analysis. Only the fact that  $\nabla \cdot \mathcal{D}$  equals a time independent quantity, sufficiently smooth in each half space, is used in the bootstrapping argument in Section 6.

Let  $T^* > 0$ . From now on we study the initial value problem on the interval  $(0, T^*)$  with initial data  $\mathbf{U}^{(0)} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ . With the above reduction the Maxwell problem (1), (2), (3), and (4) becomes

$$\begin{pmatrix} \epsilon_1^\pm & 0 & 0 \\ 0 & \epsilon_1^\pm & 0 \\ 0 & 0 & \mu_0 \end{pmatrix} \partial_t \mathbf{U}^\pm + \epsilon_3^\pm \partial_t \begin{pmatrix} (U_1^{\pm 2} + U_2^{\pm 2}) U_1^\pm \\ (U_1^{\pm 2} + U_2^{\pm 2}) U_2^\pm \\ 0 \end{pmatrix} + \begin{pmatrix} -\partial_{x_2} U_3^\pm \\ \partial_{x_1} U_3^\pm \\ \partial_{x_1} U_2^\pm - \partial_{x_2} U_1^\pm \end{pmatrix} = \mathbf{0} \quad (7)$$

on  $\mathbb{R}_\pm^2 \times (0, T^*)$  with

$$\mathbf{U}^\pm(\cdot, 0) = \mathbf{U}^{(0), \pm} \quad \text{on } \mathbb{R}_\pm^2, \quad (8)$$

and the interface conditions

$$\llbracket U_2 \rrbracket = \llbracket U_3 \rrbracket = 0 \quad \text{on } \Gamma \times [0, T^*). \quad (9)$$

System (7), (8), and (9) is the problem treated by our approximation result.

If, in addition, the sought solutions are to fit a prescribed volume charge density  $\varrho_0$  and a prescribed surface charge density  $\varrho_\Gamma$ , then the initial condition  $\mathbf{U}_E^{(0)}$  must be chosen such that the divergence condition

$$\begin{aligned} \partial_{x_1} \mathcal{D}_1(\mathbf{U}_E^{(0), \pm}) + \partial_{x_2} \mathcal{D}_2(\mathbf{U}_E^{(0), \pm}) &= \partial_{x_1} \left( \epsilon_1^\pm U_1^{(0), \pm} + \epsilon_3^\pm (U_1^{(0), \pm 2} + U_2^{(0), \pm 2}) U_1^{(0), \pm} \right) \\ &\quad + \partial_{x_2} \left( \epsilon_1^\pm U_2^{(0), \pm} + \epsilon_3^\pm (U_1^{(0), \pm 2} + U_2^{(0), \pm 2}) U_2^{(0), \pm} \right) \\ &= \varrho_0 \quad \text{on } \mathbb{R}_\pm^2 \end{aligned} \quad (10)$$

and the interface condition

$$\llbracket \mathcal{D}_1(\mathbf{U}_E^{(0)}) \rrbracket = \llbracket \epsilon_1 U_1^{(0)} + \epsilon_3 (U_1^{(0)2} + U_2^{(0)2}) U_1^{(0)} \rrbracket = \varrho_\Gamma \quad \text{on } \Gamma \quad (11)$$

are satisfied.

We study wave-packets based on the carrier wave

$$\mathbf{m}(x_1) e^{i(k_0 x_2 - \nu_0 t)}, \quad (\mathbf{x}, t) \in (\mathbb{R}^2 \setminus \Gamma) \times [0, \infty),$$

which solves the linear Maxwell equations corresponding to (7), i.e., with  $\epsilon_3 = 0$ , and with  $\varrho_0 \equiv 0$  and  $\varrho_\Gamma \equiv 0$ . Here  $k_0 \in \mathbb{R}$  is a fixed wave-number and  $\mathbf{m}(x_1)$  is a localized profile (an eigenfunction) of the resulting eigenvalue problem in  $x_1$ , and  $\nu_0 \in \mathbb{R} \setminus \{0\}$  is the corresponding eigenvalue, see Section 2 for details. We use the classical formal asymptotic ansatz of a wave-packet

$$\mathbf{U}_{\text{ans}}(\mathbf{x}, t) = \begin{pmatrix} \mathcal{E}_{\text{ans},1}(\mathbf{x}, t) \\ \mathcal{E}_{\text{ans},2}(\mathbf{x}, t) \\ \mathcal{H}_{\text{ans},3}(\mathbf{x}, t) \end{pmatrix} := \varepsilon A(\varepsilon(x_2 - \nu_1 t), \varepsilon^2 t) \mathbf{m}(x_1) e^{i(k_0 x_2 - \nu_0 t)} + \text{c.c.} \quad (12)$$

with the complex envelope  $A : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{C}$ , a small parameter  $0 < \varepsilon \ll 1$ , the group velocity  $\nu_1 \in \mathbb{R}$  at the wave-number  $k_0$  as defined in (20), and c.c. denoting the complex conjugate of the previous term. The envelope travels with the velocity  $\nu_1$ , depends slowly on the moving frame variable  $x_2 - \nu_1 t$  and is modulated in time on an even slower scale. If  $A(\cdot, \varepsilon^2 t)$  is localized, then  $\mathbf{U}_{\text{ans}}$  describes a wave-packet localized in both  $x_1$  and  $x_2$  and propagating in the  $x_2$  direction. As explained in Section 3, the ansatz (after a suitable correction via higher-order terms) produces a small residual in the Maxwell system (7) only if  $A$  satisfies a nonlinear Schrödinger equation. This NLS is an effective macroscopic description of the wave-packet dynamics.

We make the assumptions

$$\epsilon_1^\pm \in C^3(\mathbb{R}_\pm) \cap W^{3,\infty}(\mathbb{R}_\pm), \quad \epsilon_1^\pm \geq \epsilon_{1,m}^\pm \in (0, \infty), \quad (\text{A1})$$

$$\epsilon_1^\pm(x_1) \rightarrow \epsilon_1^{\pm\infty} \in [\epsilon_{1,m}^\pm, \infty) \quad \text{as } x_1 \rightarrow \pm\infty, \quad (\text{A2})$$

$$\epsilon_3^\pm \in C^3(\mathbb{R}_\pm) \cap W^{3,\infty}(\mathbb{R}_\pm), \quad \epsilon_{3,m}^\pm \leq \epsilon_3^\pm \leq \epsilon_{3,M}^\pm \quad \text{with } \epsilon_{3,m}^\pm, \epsilon_{3,M}^\pm \in \mathbb{R}. \quad (\text{A3})$$

Our main result shows that the asymptotic wave-packet ansatz (12) is close to a true solution on a time interval of length  $\mathcal{O}(\varepsilon^{-2})$ .

**Theorem 1.1** (Approximation Theorem)

Let  $T_0 > 0$ . Assume (A1), (A2), (A3), and the conditions (A4), (A5), and (A6) stated in Section 2 and let  $A \in \bigcap_{k=0}^4 C^{4-k}([0, T_0], H^{3+k}(\mathbb{R}))$  be a solution of the effective nonlinear Schrödinger equation (38). Assume that the initial value  $\mathbf{U}^{(0)} := \mathbf{U}(\cdot, 0) \in \mathcal{H}^3(\mathbb{R}^2)^3$  satisfies the nonlinear compatibility conditions of order 3, see Definition 5.9. There exist constants  $\varepsilon_0 > 0$  and  $C > 0$  such that if  $\varepsilon \in (0, \varepsilon_0)$  and if  $\mathbf{U}^{(0)} := \mathbf{U}(\cdot, 0)$  fulfills

$$\left\| \mathbf{U}^{(0)} - \mathbf{U}_{\text{ans}}(\cdot, 0) \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3} \leq c\varepsilon^{\frac{3}{2}}, \quad (13)$$

with  $c > 0$ , then there exists a solution  $\mathbf{U} \in \mathcal{G}^3(\mathbb{R}^2 \times (0, T_0\varepsilon^{-2}))^3$  of (7), (8) and (9) such that

$$\|\mathbf{U} - \mathbf{U}_{\text{ans}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times (0, T_0\varepsilon^{-2}))^3} \leq C\varepsilon^{\frac{3}{2}-\delta} \quad (14)$$

for all  $\delta > 0$ . (The space  $\mathcal{G}^3$  is introduced below.) If, in addition,  $\mathbf{U}_E^{(0)}$  satisfies (10) and (11), then we have  $\nabla \cdot \mathcal{D}(\mathbf{U}_E) = \varrho_0$  on  $(\mathbb{R}^2 \setminus \Gamma) \times (0, T_0\varepsilon^{-2})$  and  $\llbracket \mathcal{D}_1(\mathbf{U}_E) \rrbracket = \varrho_\Gamma$  on  $\Gamma \times (0, T_0\varepsilon^{-2})$ .

**Remark 1.2**

1. The existence of initial data  $\mathbf{U}^{(0)}$  which satisfy (13) and the nonlinear compatibility conditions of order 3 is an open problem. Similarly, the existence of initial data  $\mathbf{U}^{(0)}$  which satisfy (13) as well as (10) and (11) for given  $\varrho_0$  and  $\varrho_\Gamma$  is an open problem.

For the case  $\varrho_0 = 0$  and  $\varrho_\Gamma = 0$  this problem was considered in [11], where initial data were found in the form  $\mathbf{U}^{(0)} = \mathbf{U}_{\text{ans}} + \nabla\phi$  with a correction function  $\phi$ , such that (13) holds with an exponent  $a < 1$  instead of  $\frac{3}{2}$ . Note that our ansatz  $\mathbf{U}_{\text{ans}}$  naturally fits the choice  $\varrho_0 = 0$  and  $\varrho_\Gamma = 0$  because  $\epsilon_1 m \mathbf{e}^{ik_0 x_2}$  is divergence free on  $\mathbb{R}_\pm^2$  and  $\epsilon_1 m_1$  is continuous at  $x_1 = 0$ , see Remark 2.1. As a result one can easily show that  $\|\nabla \cdot \mathcal{D}(\mathbf{U}_{\text{ans}})\|_{L^2(\mathbb{R}^2)} \leq c\varepsilon^{3/2}$  and  $\llbracket \mathcal{D}_1(\mathbf{U}_{\text{ans}}(\cdot, 0)) \rrbracket \leq c\varepsilon^3$  for any bounded continuous  $A$ .

2. Due to their high regularity the components  $\mathcal{E} := (U_1, U_2, 0)^\top$  and  $\mathcal{H} := (0, 0, U_3)^\top$  of the solution  $\mathbf{U}$  of Theorem 1.1 satisfy (1), (2), (3), and (4) on  $(\mathbb{R}^2 \setminus \Gamma) \times (0, T_0\varepsilon^{-2})$  in the classical sense.

3. In the case  $\varrho_\Gamma = 0$  the regularity of  $\mathbf{U}$  produced by Theorem 1.1 guarantees that we have  $\mathcal{E} \in H_{\text{curl}}(\mathbb{R}^2)$ ,  $\mathcal{D} \in H_{\text{div}}(\mathbb{R}^2)$  and  $\mathcal{H} \in H^1(\mathbb{R}^2)$  at each point in time. This is because functions  $\mathbf{f}$  with  $\mathbf{f}|_{\mathbb{R}_\pm^2} \in H_{\text{curl}}(\mathbb{R}_\pm^2)$  and with the tangential trace being continuous across the interface, are in  $H_{\text{curl}}(\mathbb{R}^2)$ . An analogous statement holds for  $H_{\text{div}}$  and the continuity of the normal trace, see [4].

For  $m, n \in \mathbb{N}$ ,  $p \in [1, \infty]$  and an interval  $J \subset \mathbb{R}$  we define

$$\begin{aligned} \mathcal{W}^{m,p}(\mathbb{R}^n) &:= \{u \in L^p(\mathbb{R}^n) \mid u^- \in W^{m,p}(\mathbb{R}_-^n), u^+ \in W^{m,p}(\mathbb{R}_+^n)\}, \\ \|u\|_{\mathcal{W}^{m,p}(\mathbb{R}^n)} &:= \|u^-\|_{W^{m,p}(\mathbb{R}_-^n)} + \|u^+\|_{W^{m,p}(\mathbb{R}_+^n)}, \end{aligned}$$

$$\begin{aligned} \mathcal{W}^{m,p}(\mathbb{R}^n \times J) &:= \left\{ u \in L^p(\mathbb{R}^n \times J) \mid u^- \in W^{m,p}(\mathbb{R}_-^n \times J), u^+ \in W^{m,p}(\mathbb{R}_+^n \times J) \right\}, \\ \|u\|_{\mathcal{W}^{m,p}(\mathbb{R}^n \times J)} &:= \|u^-\|_{W^{m,p}(\mathbb{R}_-^n \times J)} + \|u^+\|_{W^{m,p}(\mathbb{R}_+^n \times J)} \end{aligned}$$

with the usual Lebesgue spaces  $L^p$  and Sobolev spaces  $W^{m,p}$ . For  $\mathcal{H}^m := \mathcal{W}^{m,2}$  we also set

$$\mathcal{G}^m(\mathbb{R}^n \times J) := \bigcap_{j=0}^m C^j(\bar{J}, \mathcal{H}^{m-j}(\mathbb{R}^n)), \quad \|u\|_{\mathcal{G}^m(\mathbb{R}^n \times J)} := \max_{0 \leq j \leq m} \left\| \partial_t^j u \right\|_{L^\infty(J, \mathcal{H}^{m-j}(\mathbb{R}^n))}.$$

The norm  $\|\cdot\|_{\mathcal{G}^3}$  in Theorem 1.1 thus contains spatial and temporal derivatives of total degree three.

**Remark 1.3**

We will often extend functions  $f$  with  $f^\pm \in L^p(\mathbb{R}_\pm^n)$  to a function in  $L^p(\mathbb{R}^n)$ . In general, a function  $g \in \mathcal{H}^1(\mathbb{R}^n)$  does not belong to  $H^1(\mathbb{R}^n)$  as the weak partial derivatives only exist in the half-spaces, e.g.,  $\partial_{x_1} g^+ \in L^2(\mathbb{R}_+^n)$  and  $\partial_{x_1} g^- \in L^2(\mathbb{R}_-^n)$ . Nevertheless, we will often write  $\partial_{x_1} g \in L^2(\mathbb{R}^n)$  because the weak derivatives can be extended to a function defined on  $\mathbb{R}^n$  by an arbitrary extension on  $\mathbb{R}^n \setminus (\mathbb{R}_+^n \cup \mathbb{R}_-^n)$ .

The rest of the paper is organized as follows. In Section 2 the linear spatial eigenvalue problem is studied in order to construct a carrier wave for the wave-packet. Section 3 provides a formal derivation of the NLS as an amplitude equation. In Section 4 we estimate the residual of the asymptotic approximation. In Section 5 we rewrite the reduced quasilinear Maxwell system (7), (8), (9) in the form of a hyperbolic system and adapt the local existence results of [23] to this problem. The proof of the main approximation result (Theorem 1.1) is provided in Sec. 6. The proof is based on a bootstrapping argument which extends the local existence from [23] to the existence on time intervals of length  $\mathcal{O}(\varepsilon^{-2})$  for initial data close to the (small) asymptotic ansatz. The bootstrapping simultaneously provides the error bound (14). Finally in Appendix A we describe the numerical method for computing eigenvalues (and eigenfunctions) of the linear interface problem. In Appendix B the highest order residual terms are provided explicitly. Appendix C contains estimates on products of functions in the used function spaces.

## 2 Linear time-harmonic eigenvalue problem

### 2.1 Linear eigenvalue problem

We first study the linear part of equation (7) and (9), i.e., with  $\epsilon_3 = 0$ , assuming that  $\epsilon_1$  satisfies (A1) and (A2). Using the reduction (5), (6) and the ansatz

$$\mathbf{U}(\mathbf{x}, t) = e^{i(kx_2 - \omega t)} \mathbf{w}(x_1) + \text{c.c.}, \quad (\mathbf{x}, t) \in (\mathbb{R}^2 \setminus \Gamma) \times [0, \infty),$$

where  $k, \omega \in \mathbb{R}$  and  $\mathbf{w} : \mathbb{R} \rightarrow \mathbb{C}^3$ , one arrives at the eigenvalue problem

$$L(k) \mathbf{w}(x_1) + \omega \Lambda \mathbf{w}(x_1) = \mathbf{0}, \quad x_1 \in \mathbb{R} \setminus \{0\} \tag{15}$$

for the profile  $\mathbf{w}$ . Here for each  $k \in \mathbb{R}$  the operators  $L(k) : D(L(k)) \rightarrow L^2(\mathbb{R})^3$  and  $\Lambda : D(\Lambda) \rightarrow L^2(\mathbb{R})^3$  are given by

$$L(k) \mathbf{w} := \begin{pmatrix} kw_3 \\ i\partial_{x_1} w_3 \\ kw_1 + i\partial_{x_1} w_2 \end{pmatrix}, \quad \Lambda \mathbf{w} := \begin{pmatrix} \epsilon_1(x_1)w_1 \\ \epsilon_1(x_1)w_2 \\ \mu_0 w_3 \end{pmatrix}, \tag{16}$$

with the domains  $D(\Lambda) := L^2(\mathbb{R})^3$  and

$$D(L(k)) := \left\{ \mathbf{w} : \mathbb{R} \rightarrow \mathbb{C}^3 \mid w_1 \in L^2(\mathbb{R}), w_2, w_3 \in H^1(\mathbb{R}) \right\}.$$

We call  $\omega = \omega(k) \in \mathbb{R}$  an *eigenvalue* of (15) if there exists a function  $\mathbf{w} = \mathbf{w}(k) \in D(L(k)) \setminus \{\mathbf{0}\}$  such that (15) holds. For the eigenfunctions we choose the normalization

$$\|\Lambda \mathbf{w}\|_{L^2(\mathbb{R})^3}^2 = \sum_{j=1}^3 \|(\Lambda \mathbf{w})_j\|_{L^2(\mathbb{R})}^2 = 1. \tag{17}$$

Because the operator  $L(k)$  is self-adjoint and  $\Lambda$  is real and diagonal, all eigenvalues  $\omega$  (in fact the whole spectrum) are indeed real. The interface conditions for  $\mathbf{w}$  corresponding to (9) are

$$\llbracket w_2 \rrbracket_{\text{ID}} = \llbracket w_3 \rrbracket_{\text{ID}} = 0, \quad (18)$$

where we define  $\llbracket f \rrbracket_{\text{ID}} := \lim_{x_1 \rightarrow 0^+} f(x_1) - \lim_{x_1 \rightarrow 0^-} f(x_1)$  for  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Solutions  $\mathbf{w}$  of (15) fulfil these conditions, since  $w_2, w_3 \in H^1(\mathbb{R})$ .

**Remark 2.1**

Let us, in addition, discuss the (linear) divergence condition and the interface condition (11). Solutions  $\mathbf{w}$  of (15) with  $\omega \neq 0$  satisfy (11) since  $w_3 \in H^1(\mathbb{R})$  and so  $\epsilon_1 w_1$  is continuous because of  $k w_3 + \omega \epsilon_1 w_1 = 0$ , i.e., the first line in (15). The (linear) divergence condition with  $\varrho_0 = 0$ , i.e.,  $\partial_{x_1} \mathcal{D}_1(\mathbf{U}_E) + \partial_{x_2} \mathcal{D}_2(\mathbf{U}_E) = 0$  on  $(\mathbb{R}^2 \setminus \Gamma) \times [0, \infty)$  with  $\epsilon_3 = 0$  is also automatically satisfied. Indeed, we have (for  $\epsilon_3 = 0$ )

$$\partial_{x_1} \mathcal{D}_1(\mathbf{U}_E) + \partial_{x_2} \mathcal{D}_2(\mathbf{U}_E) = (\partial_{x_1}(\epsilon_1 w_1) + ik\epsilon_1 w_2) e^{i(kx_2 - \omega t)} + \text{c.c.}$$

and due to (15)

$$\partial_{x_1}(\epsilon_1 w_1) + ik\epsilon_1 w_2 = -\frac{k}{\omega} \partial_{x_1} w_3 - i\frac{k}{\omega} \epsilon_1 \left( \frac{i\partial_{x_1} w_3}{\epsilon_1} \right) = 0. \quad (19)$$

**Remark 2.2**

Note that the second assumption in (A1) eliminates the pathological case where each  $\omega \in \mathbb{C}$  is an eigenvalue of infinite multiplicity, which is caused by the fact that gradient fields belong in the kernel of the curl operator. Indeed, if e.g.  $\epsilon_1^+ = 0$ , then  $(\partial_{x_1} f, ikf, 0)^\top$  is an eigenfunction of (15) for any  $f \in C_c^\infty(\mathbb{R}_+)$ . Here, the electric field part  $(\partial_{x_1} f, ikf)^\top$  corresponds to a gradient field (in the selected ansatz).

For the construction of the wave-packet we need that near  $k_0$  there is a unique smooth eigenvalue curve  $k \mapsto \omega(k)$  and we set

$$\nu_0 := \omega(k_0), \quad \nu_1 := \partial_k \omega(k_0), \quad \nu_2 := \partial_k^2 \omega(k_0). \quad (20)$$

This eigenvalue curve defines the so-called dispersion relation. In addition, let us assume that the eigenvalue  $\nu_0$  is simple and denote the (normalized) eigenfunction by

$$\mathbf{m} := \mathbf{w}(k_0).$$

In other words, an eigenfunction exists (near  $(k_0, \nu_0)$ ) if and only if the dispersion relation is satisfied. Hence, we impose the following assumptions for Theorem 1.1.

$$\nu_0 = \omega(k_0) \text{ is a simple eigenvalue of (15) isolated from all other eigenvalues at } k = k_0. \quad (\text{A4})$$

In order to control the essential spectrum corresponding to (15), we also require

$$\omega(k_0)\epsilon_1 \neq 0, \quad \omega(3k_0)\epsilon_1 \neq 0 \quad \text{and} \quad k_0^2 > \omega(k_0)^2 \mu_0 \epsilon_1^{\pm\infty}. \quad (\text{A5})$$

In view of (A1) and (A2), the first two conditions in (A5) just say that  $\omega(k_0)$  and  $\omega(3k_0)$  do not vanish. As noted in Corollary 2.5, the above assumptions also guarantee that the eigenvalue  $\omega(k)$  and the eigenfunction  $\mathbf{w}(k)$  depend smoothly on  $k$  near  $k_0$ , which is needed below.

Given a solution  $\mathbf{w}(k)$ , also  $\tilde{\mathbf{w}} := (\bar{w}_1, -\bar{w}_2, \bar{w}_3)^\top$  solves (15). We can thus choose the eigenfunction  $\mathbf{w}$  with real valued  $w_1, w_3$  and imaginary valued  $w_2$ , which we do throughout the rest of the paper. With this choice the normalization (17) means that

$$\int_{\mathbb{R}} (\epsilon_1 (m_1^2 - m_2^2) + \mu_0 m_3^2) dx_1 = 1. \quad (21)$$

For the proof of Theorem 1.1, i.e., the justification of (12) as an asymptotic approximation of a solution of the cubically nonlinear problem, it will be necessary to assume the non-resonance condition

$$3\nu_0 \neq \omega(3k_0), \quad \text{i.e., } 3\nu_0 \text{ is not an eigenvalue of (15) at } k = 3k_0, \quad (\text{A6})$$

see (34) and the following arguments in Section 3.

Since  $\epsilon_1$  depends on  $x_1$  it is in general not possible to solve (15) explicitly. We therefore have to calculate solutions numerically and check if the Assumptions (A4), (A5) and (A6) are satisfied.

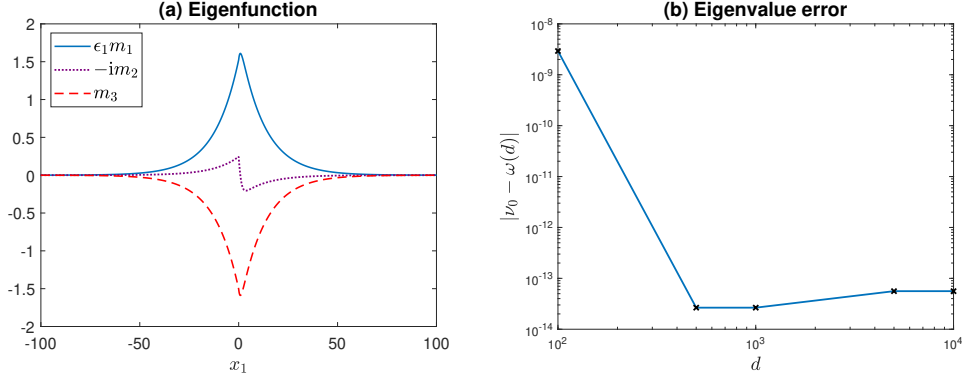


Figure 2: (for Example 2.3) (a) The eigenfunction  $\mathbf{m}$  of the linear problem (15) for  $k_0 = 0.5$ . (We plot  $\epsilon_1 m_1$  to show that the linear interface conditions are satisfied.) (b) Numerical convergence test for the eigenvalue  $\omega = \nu_0 \approx 0.494$  of  $T_{k_0, \omega} := L(k_0) + \omega \Lambda$  for  $k_0 = 0.5$  in dependence on the computational box size  $d$ .

### Example 2.3

For  $\epsilon_1(x_1) = 1\chi_{\mathbb{R}_-} + (1 + e^{-x_1})\chi_{\mathbb{R}_+}$  and  $\mu_0 = 1$ , we compute a numerical solution of (15) with the method described in Appendix A. We study the generalized eigenvalue problem  $L(k)\mathbf{w} = -\omega\Lambda\mathbf{w}$  for  $k = k_0$  on the interval  $[-d, d]$  and compute all eigenvalues in a neighborhood of  $\omega = \nu_0$ . For  $k_0 = 0.5$ , step size  $h = 0.01$  and interval length  $d = 5 \cdot 10^4$  we get the eigenvalue  $\omega(k_0) = \nu_0 \approx 0.494$  and no other eigenvalue in a neighborhood of  $\nu_0$ .

To check the effects of the boundary, we repeated the calculation for different intervals  $[-d, d]$  and get the eigenvalue  $\omega(d)$  closest to  $\nu_0$  in dependence on  $d$ . In Figure 2 (a) we see the calculated eigenfunction  $\mathbf{w}$  and Figure 2 (b) shows that the error in the calculation of  $\omega(d)$  converges to zero for increasing  $d$ .

Note that for this example we have  $\epsilon_{1,m}^{\pm} = \epsilon_1^{\pm\infty} = 1$  and one can numerically calculate that the eigenvalue closest to  $3\nu_0 \approx 1.481$  is given by  $\omega(3k_0) \approx 1.404$ . Therefore, Assumptions (A1), (A2), (A4), (A5), and (A6) appear to be satisfied.

## 2.2 Solution of the inhomogeneous problem

In Section 3 we also have to solve the inhomogeneous version of the eigenvalue problem

$$T_{k,\omega}\mathbf{v} := (L(k) + \omega\Lambda)\mathbf{v} = \begin{pmatrix} \omega\epsilon_1 & 0 & k \\ 0 & \omega\epsilon_1 & i\partial_{x_1} \\ k & i\partial_{x_1} & \omega\mu_0 \end{pmatrix} \mathbf{v} = \mathbf{f}, \quad (22)$$

with  $\mathbf{f} \in N(T_{k,\omega})^\perp$ , where  $N(T_{k,\omega}) \subset D(T_{k,\omega}) := \{\mathbf{v} \in L^2(\mathbb{R})^3 \mid v_2, v_3 \in H^1(\mathbb{R})\}$ .

### Lemma 2.4

Let  $\epsilon_1 \in W^{1,\infty}(\mathbb{R})$  satisfy (A2) and let  $k, \omega \in \mathbb{R}$  be such that  $k^2 > \omega^2 \epsilon_1^{\pm\infty} \mu_0$  and  $\omega\epsilon_1 \neq 0$ . Assume that we are in one of the cases

- i) 0 is a simple eigenvalue of  $T_{k,\omega}$  isolated from all other eigenvalues;
- ii) 0 is not an eigenvalue of  $T_{k,\omega}$ .

If  $\mathbf{f} \in N(T_{k,\omega})^\perp \subset L^2(\mathbb{R})^3$  ( $\mathbf{f} \in L^2(\mathbb{R})^3$  if  $N(T_{k,\omega}) = \{0\}$ ), then (22) has a solution  $\mathbf{v} \in D(T_{k,\omega})$ .

PROOF: Equation (22) splits for  $\omega\epsilon_1 \neq 0$  into the scalar equation

$$v_1 = \frac{1}{\omega\epsilon_1} (f_1 - kv_3)$$



and the reduced problem

$$\tilde{T}_{k,\omega} \tilde{\mathbf{v}} = \tilde{\mathbf{f}}$$

with

$$\tilde{T}_{k,\omega} := \begin{pmatrix} \omega\epsilon_1 & i\partial_{x_1} \\ i\partial_{x_1} & \omega\mu_0 - \frac{k^2}{\omega\epsilon_1} \end{pmatrix}, \quad D(\tilde{T}_{k,\omega}) := H^1(\mathbb{R})^2, \quad \tilde{\mathbf{v}} := \begin{pmatrix} v_2 \\ v_3 \end{pmatrix}, \quad \tilde{\mathbf{f}} := \begin{pmatrix} f_2 \\ f_3 - \frac{k}{\omega\epsilon_1} f_1 \end{pmatrix}.$$

Note that

$$(v_2, v_3)^\top \in N(\tilde{T}_{k,\omega}) \iff \left( -\frac{k}{\omega\epsilon_1} v_3, v_2, v_3 \right)^\top \in N(T_{k,\omega})$$

and hence

$$\tilde{\mathbf{f}} \in N(\tilde{T}_{k,\omega})^\perp \iff \mathbf{f} \in N(T_{k,\omega})^\perp.$$

We also obtain that  $0 \in \sigma(\tilde{T}_{k,\omega})$  if and only if  $0 \in \sigma(T_{k,\omega})$ .

Since  $\tilde{T}_{k,\omega}$  is self-adjoint, the result will follow from the closedness of the range of  $\tilde{T}_{k,\omega}$ . We check the closedness by showing that  $\tilde{T}_{k,\omega}$  is Fredholm. To this aim, we rewrite the problem as the linear ordinary differential equation

$$\partial_{x_1} \tilde{\mathbf{v}} = A(x_1) \tilde{\mathbf{v}} + \mathbf{g}$$

with

$$\mathbf{g} := -i \begin{pmatrix} f_3 - \frac{k}{\omega\epsilon_1} f_1 \\ f_2 \end{pmatrix}, \quad A(x_1) := \begin{pmatrix} 0 & i \left( \omega\mu_0 - \frac{k^2}{\omega\epsilon_1(x_1)} \right) \\ i\omega\epsilon_1(x_1) & 0 \end{pmatrix} =: \begin{cases} A_-(x_1), & x_1 < 0, \\ A_+(x_1), & x_1 > 0. \end{cases}$$

Theorem 1.2 in [3] says that  $\tilde{T}_{k,\omega}$  is Fredholm if and only if the ODEs

$$\partial_{x_1} \tilde{\mathbf{v}}^- = A_-(x_1) \tilde{\mathbf{v}}^-, \quad x_1 < 0, \quad (23)$$

$$\partial_{x_1} \tilde{\mathbf{v}}^+ = A_+(x_1) \tilde{\mathbf{v}}^+, \quad x_1 > 0, \quad (24)$$

have exponential dichotomies. We only show the dichotomy for (24) as (23) can be treated analogously. First, the problem

$$\partial_{x_1} \mathbf{w} = A_{+\infty} \mathbf{w}$$

with the constant coefficient matrix

$$A_{+\infty} := A(x_1 \rightarrow \infty) = \begin{pmatrix} 0 & i\omega\mu_0 - \frac{k^2}{\omega\epsilon_1^{+\infty}} \\ i\omega\epsilon_1^{+\infty} & 0 \end{pmatrix}$$

has an exponential dichotomy since the eigenvalues

$$\lambda_{1,2} = \pm \sqrt{k^2 - \omega^2 \epsilon_1^{+\infty} \mu_0}$$

of  $A_{+\infty}$  are real with different signs for  $k^2 > \omega^2 \epsilon_1^{+\infty} \mu_0$ . Then Proposition 1 in Chapter 4 and the discussion starting on page 13 of [6] imply that also

$$\partial_{x_1} \tilde{\mathbf{v}}^+ = A_+(x_1) \tilde{\mathbf{v}}^+ = (A_{+\infty} + (A_+(x_1) - A_{+\infty})) \tilde{\mathbf{v}}^+$$

has an exponential dichotomy, because  $A_+(x_1) - A_{+\infty}$  tends to 0 as  $x_1 \rightarrow \infty$ .  $\square$

Using the spectral information obtained above, we next show that the eigenvalues  $\omega(k)$  and the corresponding eigenfunctions  $w(k)$  are smooth in  $k$ .

### Corollary 2.5

Let (A1), (A2), (A4) and (A5) be true. Then for some  $\delta > 0$  the eigenvalues and the corresponding eigenfunctions of problem (15) satisfy

$$\omega \in C^\infty((k_0 - \delta, k_0 + \delta), \mathbb{R}) \quad \text{and} \quad \mathbf{w} \in C^\infty((k_0 - \delta, k_0 + \delta), L^2(\mathbb{R}) \times H^1(\mathbb{R}) \times H^1(\mathbb{R})).$$

PROOF: We translate our problem into standard perturbation theory of spectra, as discussed in [16]. By the assumptions,  $\omega_0$  is a simple eigenvalue of  $\Lambda^{-1}L(k_0)$  with eigenfunction  $\mathbf{w}(k_0)$ , and there are no other eigenvalues nearby. As shown in the proof of Lemma 2.4,  $\omega$  belongs to the resolvent set of  $\Lambda^{-1}L(k_0)$  if and only if  $\tilde{T}_{k_0,\omega}$  is invertible. For  $\omega \approx \omega_0$  we can write  $\tilde{T}_{k_0,\omega} = \tilde{T}_{k_0,\omega_0} + R$  with a perturbation  $R : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ , whose norm is bounded by  $c|\omega - \omega_0|$ . In the proof of Lemma 2.4 we have seen that  $\tilde{T}_{k_0,\omega_0}$  is a Fredholm operator, hence  $0 \notin \sigma_{\text{ess}}(\tilde{T}_{k_0,\omega_0})$  and the same is true for  $\tilde{T}_{k_0,\omega}$  if  $\omega$  is close to  $\omega_0$ . If  $0$  was an eigenvalue of  $\tilde{T}_{k_0,\omega}$ , the number  $\omega \neq \omega_0$  would be an eigenvalue of  $\Lambda^{-1}L(k_0)$  which is impossible in a small enough neighborhood of  $\omega_0$  by assumption (A4). As a result,  $0$  is contained in  $\rho(\tilde{T}_{k_0,\omega})$  and thus  $\omega_0$  is an isolated simple eigenvalue of  $\Lambda^{-1}L(k_0)$ .

For  $k \approx k_0$ , Theorem 1.8 in §VII.1 of [16] now shows that  $\Lambda^{-1}L(k)$  has a simple eigenvalue  $\omega(k)$  smoothly depending on  $k$ . Also the projection  $P(k)$  onto the eigenspace is smooth in  $k$ . Hence, the mapping  $k \mapsto P(k)\mathbf{w}(k_0)$  is a smooth family of eigenfunctions of (15) if  $k$  is close to  $k_0$ .  $\square$

The next lemma improves the regularity of solutions to (22) if the right-hand side is smooth enough.

### Lemma 2.6

Assume (A1) and (A5). Let  $k, \omega \in \mathbb{R}$ ,  $\mathbf{f} := (f_1, f_2, f_3)^\top$  with  $f_1 \in \mathcal{H}^3(\mathbb{R})$  and  $f_2, f_3 \in \mathcal{H}^2(\mathbb{R})$ . If  $\mathbf{v} \in L^2(\mathbb{R})^3$  is a solution of (22), then  $\mathbf{v} \in \mathcal{H}^3(\mathbb{R})^3$ .

PROOF: We start by showing that  $\mathbf{v} \in \mathcal{H}^1(\mathbb{R})$ . From (22) we know that

$$\begin{cases} i\partial_{x_1} v_2 = f_3 - kv_1 - \mu_0 \omega v_3, \\ i\partial_{x_1} v_3 = f_2 - \epsilon_1 \omega v_2. \end{cases} \quad (25)$$

The right-hand sides in (25) belong to  $L^2(\mathbb{R})$  and therefore  $v_2, v_3 \in \mathcal{H}^1(\mathbb{R})$ . The assumptions on  $\epsilon_1$  imply that  $\partial_{x_1}(\epsilon_1^{-1}) = -\epsilon_1^{-2}\partial_{x_1}\epsilon_1 \in L^\infty(\mathbb{R}_\pm)$ . Now  $v_1 \in \mathcal{H}^1(\mathbb{R})$  is a direct consequence of

$$v_1 = \frac{1}{\epsilon_1 \omega} (f_1 - kv_3). \quad (26)$$

We can now iterate this process since  $\epsilon_1^{-1} \in \mathcal{W}^{3,\infty}(\mathbb{R})$ . Equations (25) and (26) yield that  $\mathbf{v} \in \mathcal{H}^2(\mathbb{R})^3$  if one knows that  $\mathbf{v} \in \mathcal{H}^1(\mathbb{R})^3$ . This fact then implies that  $\mathbf{v} \in \mathcal{H}^3(\mathbb{R})^3$ .  $\square$

## 3 Envelope approximation of wave-packets; amplitude equation

The aim of this section is to make the residual

$$\mathbf{Res}(\mathbf{U}_{\text{ans}}) := \begin{pmatrix} \partial_t \mathcal{D}_1(\mathbf{U}_{\text{ans}}) - \partial_{x_2} U_{\text{ans},3} \\ \partial_t \mathcal{D}_2(\mathbf{U}_{\text{ans}}) + \partial_{x_1} U_{\text{ans},3} \\ -\partial_{x_2} U_{\text{ans},1} + \partial_{x_1} U_{\text{ans},2} + \mu_0 \partial_t U_{\text{ans},3} \end{pmatrix} \quad (27)$$

of (12) in the Maxwell problem (7) small enough for the subsequent justification of the asymptotics, i.e., for the proof of Theorem 1.1. As we will see, this requires an extension of the ansatz (12). In the propagation direction  $x_2$  we mostly work in Fourier variables applying the Fourier transform

$$\widehat{f}(k) = \mathcal{F}(f)(k) := (2\pi)^{-1/2} \int_{\mathbb{R}} f(x) e^{-ikx} dx.$$

The corresponding inverse transform is given by

$$\mathcal{F}^{-1}(f)(x) := (2\pi)^{-1/2} \int_{\mathbb{R}} f(k) e^{ikx} dk.$$

Indeed, it is  $\mathcal{F} \circ \mathcal{F}^{-1} = \mathcal{F}^{-1} \circ \mathcal{F} = \text{Id} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$  after the standard extension of the transforms from  $L^1(\mathbb{R})$  to  $L^2(\mathbb{R})$ .

For the wave-packet  $\mathbf{U}_{\text{ans}}$  we compute

$$\widehat{\mathbf{U}}_{\text{ans}}(x_1, k, t) := \begin{pmatrix} \widehat{\mathcal{E}}_{\text{ans},1}(x_1, k, t) \\ \widehat{\mathcal{E}}_{\text{ans},2}(x_1, k, t) \\ \widehat{\mathcal{H}}_{\text{ans},3}(x_1, k, t) \end{pmatrix} := \widehat{A}\left(\frac{k-k_0}{\varepsilon}, \varepsilon^2 t\right) \mathbf{m}(x_1) e^{-i(\nu_0+(k-k_0)\nu_1)t} + \widehat{\text{c.c.}}, \quad (28)$$

where  $\widehat{\text{c.c.}}(\widehat{f}) = \widehat{\text{c.c.}}(\widehat{f})$ . Maxwell's equations with the reduction (5) transform to

$$\begin{cases} \partial_t \widehat{\mathcal{D}}_1 - ik \widehat{\mathcal{H}}_3 = 0, \\ \partial_t \widehat{\mathcal{D}}_2 + \partial_{x_1} \widehat{\mathcal{H}}_3 = 0, \\ -ik \widehat{\mathcal{E}}_1 + \partial_{x_1} \widehat{\mathcal{E}}_2 + \mu_0 \partial_t \widehat{\mathcal{H}}_3 = 0, \end{cases} \quad (29)$$

with

$$\widehat{\mathcal{D}}(\boldsymbol{\mathcal{E}}) = \varepsilon_1 \widehat{\boldsymbol{\mathcal{E}}} + \varepsilon_3 ((\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{E}})^\wedge.$$

In what follows we use the notations  $E_1 := e^{-i(\nu_0+(k-k_0)\nu_1)t}$ ,  $F_1 := e^{i(k_0 x_2 - \nu_0 t)}$ ,  $K := \frac{k-k_0}{\varepsilon}$ ,  $T := \varepsilon^2 t$  and  $X_2 := \varepsilon(x_2 - \nu_1 t)$ , and we will suppress the arguments of  $\mathbf{m} = \mathbf{m}(x_1)$  and  $\widehat{A} = \widehat{A}(K, T)$  and their derivatives if they are obvious.

We start our formal asymptotic analysis by writing out the nonlinear term, where we employ the notation  $\mathcal{D} = \mathcal{D}_{\text{lin}} + \mathcal{D}_{\text{nl}}$  with

$$\mathcal{D}_{\text{lin}} := \varepsilon_1 \boldsymbol{\mathcal{E}}, \quad \mathcal{D}_{\text{nl}} := \varepsilon_3 (\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{E}}.$$

In the physical variables we get

$$\begin{aligned} \partial_t \mathcal{D}_{\text{nl},1}(\mathbf{U}_{\text{ans}}) &= -\varepsilon^3 3i\varepsilon_3 \nu_0 F_1^3 A^3 (m_1^3 + m_1 m_2^2) \\ &\quad -\varepsilon^3 i\varepsilon_3 \nu_0 F_1 |A|^2 A (3m_1^3 - m_2^2 m_1) + \mathcal{O}(\varepsilon^4) + \text{c.c.} \quad (\varepsilon \rightarrow 0), \end{aligned}$$

using that  $m_{1,3}$  are real and  $m_2$  is imaginary. As one easily checks,

$$\begin{aligned} \int_{\mathbb{R}} e^{i(k_0 x_2 - \nu_0 t)} |A(X_2, T)|^2 A(X_2, T) e^{-ikx_2} dx_2 &= (2\pi)^{-1/2} \varepsilon^{-1} E_1 \left( \widehat{A} * \widehat{A} * \widehat{A} \right) (K, T), \\ \int_{\mathbb{R}} e^{3i(k_0 x_2 - \nu_0 t)} A^3(X_2, T) e^{-ikx_2} dx_2 &= (2\pi)^{-1/2} \varepsilon^{-1} E_3 \left( \widehat{A} * \widehat{A} * \widehat{A} \right) (\widetilde{K}, T), \end{aligned}$$

with  $\widetilde{K} := \frac{k-3k_0}{\varepsilon}$ ,  $E_3 := e^{-i(3\nu_0+(k-3k_0)\nu_1)t}$ , and the convolution

$$(f * g)(K) = \int_{\mathbb{R}} f(K-s)g(s) ds.$$

Hence, as  $\varepsilon \rightarrow 0$  we have

$$\begin{aligned} \partial_t \widehat{\mathcal{D}}_{\text{nl},1}(\mathbf{U}_{\text{ans}}) &= -\varepsilon^2 (2\pi)^{-1} i\varepsilon_3 \nu_0 E_1 (3m_1^3 - m_2^2 m_1) \left( \widehat{A} * \widehat{A} * \widehat{A} \right) (K, T) \\ &\quad -\varepsilon^2 3(2\pi)^{-1} i\varepsilon_3 \nu_0 E_3 (m_1^3 + m_1 m_2^2) \left( \widehat{A} * \widehat{A} * \widehat{A} \right) (\widetilde{K}, T) + \mathcal{O}(\varepsilon^3) + \widehat{\text{c.c.}} \end{aligned} \quad (30)$$

The second component  $\partial_t \widehat{\mathcal{D}}_{\text{nl},2}(\mathbf{U}_{\text{ans}})$  is obtained from  $\partial_t \widehat{\mathcal{D}}_{\text{nl},1}(\mathbf{U}_{\text{ans}})$  by simply switching the indices 1 and 2 and the third component  $\widehat{\mathcal{D}}_{\text{nl},3}(\mathbf{U}_{\text{ans}})$  obviously vanishes.

Below we use the Taylor expansion of  $\omega(k)$ , of the corresponding eigenfunction  $\mathbf{w}(k)$ , see Corollary 2.5, and of the operator  $L(k)$ . Recalling (20) and  $k = k_0 + \varepsilon K$ , we obtain

$$\begin{aligned} \omega(k) &= \omega(k_0 + \varepsilon K) = \nu_0 + \varepsilon K \nu_1 + \frac{1}{2} \varepsilon^2 K^2 \nu_2 + \mathcal{O}(\varepsilon^3), \\ \mathbf{w}(k) &= \mathbf{w}(k_0 + \varepsilon K) = \mathbf{m} + \varepsilon K \partial_k \mathbf{w}(k_0) + \frac{1}{2} \varepsilon^2 K^2 \partial_k^2 \mathbf{w}(k_0) + \mathcal{O}(\varepsilon^3), \end{aligned}$$

$$L(k) = L(k_0 + \varepsilon K) = L(k_0) + \varepsilon K \partial_k L(k_0) = L_0 + \varepsilon K L_1$$

as  $\varepsilon \rightarrow 0$ , where

$$L_1 \mathbf{m} := (\partial_k L(k_0)) \mathbf{m} = \begin{pmatrix} m_3 \\ 0 \\ m_1 \end{pmatrix}$$

and all higher derivatives in  $k$  of  $L$  vanish since it is linear in  $k$ . Differentiation of the linear eigenvalue problem (15) then produces the equations

$$(L_0 + \nu_0 \Lambda) \mathbf{m} = \mathbf{0}, \quad (31)$$

$$(L_1 + \nu_1 \Lambda) \mathbf{m} + (L_0 + \nu_0 \Lambda) \partial_k \mathbf{w}(k_0) = \mathbf{0}, \quad (32)$$

$$K^2 \nu_2 \Lambda \mathbf{m} + 2K^2 (L_1 + \nu_1 \Lambda) \partial_k \mathbf{w}(k_0) + K^2 (L_0 + \nu_0 \Lambda) \partial_k^2 \mathbf{w}(k_0) = \mathbf{0}. \quad (33)$$

The residual is obtained by inserting (28) in the left-hand side of (29). We obtain at  $\mathcal{O}(\varepsilon^0)$  the expression  $-i \widehat{A} E_1 (L_0 + \nu_0 \Lambda) \mathbf{m}$  which vanishes due to (31). At  $\mathcal{O}(\varepsilon^1)$  we obtain  $-i K \widehat{A} E_1 (L_1 \mathbf{m} + \nu_1 \Lambda \mathbf{m})$ . In order to annihilate the residual also at  $\mathcal{O}(\varepsilon^1)$ , equation (32) dictates that we need to extend the ansatz  $\widehat{U}_{\text{ans}}$  by the term  $\varepsilon \widehat{A} K \partial_k \mathbf{w}(k_0) E_1$ .

At  $\mathcal{O}(\varepsilon^2)$  terms proportional to  $E_1$  and those proportional to  $E_3$  (as obtained in (30)) appear in the residual. The latter terms can be removed by introducing a further correction term to  $\widehat{U}_{\text{ans}}$ , namely  $\varepsilon^2 (2\pi)^{-1} (\widehat{A} * \widehat{A} * \widehat{A}) \mathbf{h}(x_1) E_3$ , where  $\mathbf{h}$  solves

$$(L(3k_0) + 3\nu_0 \Lambda) \mathbf{h} = \begin{pmatrix} -3\nu_0 \varepsilon_3 (m_1^3 + m_1 m_2^2) \\ -3\nu_0 \varepsilon_3 (m_2^3 + m_2 m_1^2) \\ 0 \end{pmatrix}. \quad (34)$$

The non-resonance assumption (A6) guarantees that a solution  $\mathbf{h}$  exists. Indeed,  $L(3k_0) + 3\nu_0 \Lambda$  is injective by (A6) and hence the closed range theorem implies  $R(L(3k_0) + 3\nu_0 \Lambda) = N(L(3k_0) + 3\nu_0 \Lambda)^\perp = L^2(\mathbb{R})^3$ , thanks to Lemma 2.4 (case ii) with  $k = 3k_0$  and  $\omega = 3\nu_0$ . Here Assumption (A5) is used.

In summary, the residual is in  $\mathcal{O}(\varepsilon^2)$  and contains only terms proportional to  $E_1$  (and their complex conjugates) if we modify  $\widehat{U}_{\text{ans}}$  to

$$\begin{aligned} \widehat{U}_{\text{mod}}(x_1, k, t) := & \widehat{A} \left( \frac{k - k_0}{\varepsilon}, \varepsilon^2 t \right) (\mathbf{m}(x_1) + \varepsilon K \partial_k \mathbf{w}(x_1, k_0)) e^{-i(\nu_0 + (k - k_0)\nu_1)t} \\ & + \varepsilon^2 (2\pi)^{-1} (\widehat{A} * \widehat{A} * \widehat{A}) \left( \frac{k - 3k_0}{\varepsilon}, \varepsilon^2 t \right) \mathbf{h}(x_1) e^{-i(3\nu_0 + (k - 3k_0)\nu_1)t} + \widehat{c.c.} \end{aligned} \quad (35)$$

As a result, the  $\mathcal{O}(\varepsilon^2)$ -terms in the residual of  $\widehat{U}_{\text{mod}}$  are

$$\varepsilon^2 \widehat{\mathbf{Res}}^{(2, E_1)} E_1 + \widehat{c.c.},$$

where

$$\begin{aligned} \widehat{\mathbf{Res}}^{(2, E_1)} := & -i \left[ K^2 \widehat{A} (L_1 + \nu_1 \Lambda) \partial_k \mathbf{w}(\cdot, k_0) + i \partial_T \widehat{A} \Lambda \mathbf{m} \right. \\ & \left. + \nu_0 \varepsilon_3 (2\pi)^{-1} (\widehat{A} * \widehat{A} * \widehat{A}) \begin{pmatrix} 3m_1^3 - m_1 m_2^2 \\ -3m_2^3 + m_1^2 m_2 \\ 0 \end{pmatrix} \right], \end{aligned} \quad (36)$$

again recalling that  $m_{1,3}$  are real and  $m_2$  is imaginary.

It will be sufficient that the  $L^2$ -projection  $P_m$  of  $\widehat{\mathbf{Res}}^{(2, E_1)}$  onto the kernel of  $L_0 + \nu_0 \Lambda$  vanishes, where

$$P_m \mathbf{f}(x_1) := \langle \mathbf{f}, \mathbf{m} \rangle_{L^2(\mathbb{R})^3} \mathbf{m}(x_1) = \int_{\mathbb{R}} \mathbf{f}(\xi_1) \cdot \overline{\mathbf{m}}(\xi_1) d\xi_1 \mathbf{m}(x_1).$$

This is equivalent to the condition that the envelope  $A$  satisfies a certain nonlinear Schrödinger equation, as we show now. First, we use (33) and replace  $K^2(L_1 + \nu_1\Lambda)\partial_k\mathbf{w}(k_0)$  in (36) by

$$-\frac{1}{2}K^2\nu_2\Lambda\mathbf{m} - \frac{1}{2}K^2(L_0 + \nu_0\Lambda)\partial_k^2\mathbf{w}(k_0).$$

The self-adjointness of  $L_0 + \nu_0\Lambda$  implies

$$P_m \left( -\frac{1}{2}K^2\nu_2\Lambda\mathbf{m} - \frac{1}{2}K^2(L_0 + \nu_0\Lambda)\partial_k^2\mathbf{w}(k_0) \right) = -\frac{1}{2}K^2\nu_2P_m(\Lambda\mathbf{m}) = -\frac{1}{2}K^2\nu_2\mathbf{m}$$

due to the normalization of  $\mathbf{m}$ . Altogether,  $P_m$  of (36) is zero if  $\widehat{A}$  satisfies

$$\begin{aligned} 0 = & \int_{\mathbb{R}} (\epsilon_1 (m_1^2 - m_2^2) + \mu_0 m_3^2) dx_1 \left( i\partial_T \widehat{A} - \frac{1}{2}K^2\nu_2\widehat{A} \right) \\ & + (2\pi)^{-1}\nu_0 \int_{\mathbb{R}} \epsilon_3 (3m_1^4 - 2m_1^2m_2^2 + 3m_2^4) dx_1 \left( \widehat{A} * \widehat{A} * \widehat{A} \right). \end{aligned} \quad (37)$$

In other words,  $A$  has to solve the nonlinear Schrödinger equation

$$i\partial_T A = -\frac{1}{2}\nu_2\partial_{X_2}^2 A + \kappa|A|^2 A \quad (38)$$

with

$$\kappa := -\nu_0 \int_{\mathbb{R}} \epsilon_3 (3m_1^4 - 2m_1^2m_2^2 + 3m_2^4) dx_1.$$

Here (21) has been used. Equation (38) is the so-called effective amplitude equation. Note that for Theorem 1.1 we need smooth solutions  $A \in \bigcap_{k=0}^4 C^{4-k}([0, T_0], H^{3+k}(\mathbb{R}))$  of (38). Such solutions are provided by Proposition 3.8 and Remark 3.9 in [28].

## 4 Estimation of the residual

In this section we estimate the residual in the  $\mathcal{H}^3(\mathbb{R}^2)^3$ -norm rigorously under the assumption that  $A$  solves (38). Here  $\mathbf{Res}$  is obtained from  $\widehat{\mathbf{Res}}$  by applying the inverse Fourier transformation. The modified wave-packet  $\mathbf{U}_{\text{mod}}$  from (35) has the residual  $\widehat{\mathbf{Res}}(\widehat{\mathbf{U}}_{\text{mod}}) = \varepsilon^2(I - P_m)\widehat{\mathbf{Res}}^{(2, E_1)} E_1 + \mathcal{O}(\varepsilon^3)$  as shown in Section 3. Since  $\widehat{\mathbf{Res}}^{(2, E_1)}$  only depends on  $K = \varepsilon^{-1}(k - k_0)$  (and not directly on  $k$ ), we have formally  $\mathbf{Res}^{(2, E_1)} = \mathcal{O}(\varepsilon)$ . The  $L^2$ -norm is bounded by

$$\|\mathbf{Res}(\mathbf{U}_{\text{mod}})(\cdot, \cdot, t)\|_{L^2(\mathbb{R}^2)^3} \leq C\varepsilon^{\frac{5}{2}}$$

due to the presence of functions depending on  $X_2 = \varepsilon(x_2 - \nu_1 t)$ .

It turns out that for the error analysis the residual has to be bounded in the  $\mathcal{H}^3(\mathbb{R}^2)^3$ -norm by  $C\varepsilon^{7/2}$  for all  $t \in [0, T_0\varepsilon^{-2}]$ . This requires a further extension of the ansatz. We introduce the final modification of the wave-packet by

$$\begin{aligned} \widehat{\mathbf{U}}_{\text{ext}}(x_1, k, t) := & \widehat{A} \left( \frac{k - k_0}{\varepsilon}, \varepsilon^2 t \right) (\mathbf{m}(x_1) + \varepsilon K \partial_k \mathbf{w}(x_1, k_0)) e^{-i(\nu_0 + (k - k_0)\nu_1)t} \\ & + \varepsilon^2 (2\pi)^{-1} \left( \widehat{A} * \widehat{A} * \widehat{A} \right) \left( \frac{k - 3k_0}{\varepsilon}, \varepsilon^2 t \right) \mathbf{h}(x_1) e^{-i(3\nu_0 + (k - 3k_0)\nu_1)t} \\ & + \varepsilon^2 \widehat{A} \left( \frac{k - k_0}{\varepsilon}, \varepsilon^2 t \right) \frac{1}{2} K^2 \partial_k^2 \mathbf{w}(x_1, k_0) e^{-i(\nu_0 + (k - k_0)\nu_1)t} \\ & + \varepsilon^2 (2\pi)^{-1} \left( \widehat{A} * \widehat{A} * \widehat{A} \right) \left( \frac{k - k_0}{\varepsilon}, \varepsilon^2 t \right) \mathbf{p}(x_1) e^{-i(\nu_0 + (k - k_0)\nu_1)t} + \widehat{c.c.}, \end{aligned}$$

where  $\mathbf{p}$  will be chosen such that formally  $\widehat{\mathbf{Res}}(\widehat{\mathbf{U}}_{\text{ext}}) = \mathcal{O}(\varepsilon^3)$ . To determine  $\mathbf{p}$ , we calculate

$$\begin{aligned} \widehat{\mathbf{Res}}(\widehat{\mathbf{U}}_{\text{ext}}) &= \varepsilon^2 E_1 (I - P_m) \widehat{\mathbf{Res}}^{(2, E_1)} - \frac{1}{2} i \varepsilon^2 E_1 K^2 (L_0 + \nu_0 \Lambda) \partial_k^2 \mathbf{w}(x_1, k_0) \widehat{A} \\ &\quad - i(2\pi)^{-1} \varepsilon^2 E_1 (L_0 + \nu_0 \Lambda) \mathbf{p} \left( \widehat{A} * \widehat{A} * \widehat{A} \right) + \text{c.c.} + \mathcal{O}(\varepsilon^3). \end{aligned}$$

Since  $\widehat{A}$  solves (37), we know that

$$\begin{aligned} (I - P_m) \widehat{\mathbf{Res}}^{(2, E_1)} &= \frac{1}{2} i K^2 (L_0 + \nu_0 \Lambda) \partial_k^2 \mathbf{w}(x_1, k_0) \widehat{A} \\ &\quad - i(2\pi)^{-1} \left( \kappa \begin{pmatrix} \varepsilon_1 m_1 \\ \varepsilon_1 m_2 \\ \mu_0 m_3 \end{pmatrix} + \varepsilon_3 \nu_0 \begin{pmatrix} 3m_1^3 - m_1 m_2^2 \\ -3m_2^3 + m_1^2 m_2 \\ 0 \end{pmatrix} \right) \left( \widehat{A} * \widehat{A} * \widehat{A} \right), \end{aligned}$$

where (33) was used again. Therefore the terms of order  $\varepsilon^2$  in  $\widehat{\mathbf{Res}}(\widehat{\mathbf{U}}_{\text{ext}})$  vanish if  $\mathbf{p}$  solves

$$(L_0 + \nu_0 \Lambda) \mathbf{p} = -\kappa \begin{pmatrix} \varepsilon_1 m_1 \\ \varepsilon_1 m_2 \\ \mu_0 m_3 \end{pmatrix} - \varepsilon_3 \nu_0 \begin{pmatrix} 3m_1^3 - m_1 m_2^2 \\ -3m_2^3 + m_1^2 m_2 \\ 0 \end{pmatrix}. \quad (39)$$

Such a function  $\mathbf{p}$  exists since the right-hand side in (39) is orthogonal to  $\mathbf{m}$  by the choice of  $\kappa$  and it therefore lies in the range of  $L_0 + \nu_0 \Lambda$  due to the closed range theorem. Here we use Lemma 2.4 (case i)) with  $k = k_0$  and  $\omega = \nu_0$ . Here Assumption (A5) is employed. We now apply the inverse Fourier transformation to obtain

$$\begin{aligned} \mathbf{U}_{\text{ext}}(x_1, x_2, t) &= \varepsilon A(X_2, T) \mathbf{m}(x_1) e^{i(k_0 x_2 - \nu_0 t)} \\ &\quad - \varepsilon^2 i \partial_{X_2} A(X_2, T) \partial_k \mathbf{w}(x_1, k_0) e^{i(k_0 x_2 - \nu_0 t)} \\ &\quad - \varepsilon^3 \frac{1}{2} \partial_{X_2}^2 A(X_2, T) \partial_k^2 \mathbf{w}(x_1, k_0) e^{i(k_0 x_2 - \nu_0 t)} \\ &\quad + \varepsilon^3 |A(X_2, T)|^2 A(X_2, T) \mathbf{p}(x_1) e^{i(k_0 x_2 - \nu_0 t)} \\ &\quad + \varepsilon^3 A^3(X_2, T) \mathbf{h}(x_1) e^{3i(k_0 x_2 - \nu_0 t)} + \text{c.c.}, \end{aligned} \quad (40)$$

recalling that  $X_2 = \varepsilon(x_2 - \nu_1 t)$  and  $T = \varepsilon^2 t$ . Since  $\widehat{\mathbf{Res}}(\widehat{\mathbf{U}}_{\text{ext}})$  is of order  $\varepsilon^3$  after transformation  $\mathbf{Res}(\mathbf{U}_{\text{ext}})$  is of order  $\varepsilon^4$  formally. The terms of order  $\varepsilon^4$  of  $\mathbf{Res}(\mathbf{U}_{\text{ext}})$  can be found in Appendix B.

#### Remark 4.1

Note that our residual incorporates neither the divergence condition on  $\mathcal{D}$  nor the interface conditions. This is because these quantities do not directly appear in the  $\mathcal{G}^3$ -norm which we use to estimate the error.

Nevertheless, for  $\mathcal{E}_{\text{ext}} := (U_{\text{ext},1}, U_{\text{ext},2}, 0)^\top$  one can show that  $\nabla \cdot \mathcal{D}(\mathcal{E}_{\text{ext}}) = \mathcal{O}(\varepsilon^2)$ . Indeed, for the divergence condition we have

$$\partial_{x_1} \mathcal{D}_1(\mathcal{E}_{\text{ext}}) + \partial_{x_2} \mathcal{D}_2(\mathcal{E}_{\text{ext}}) = \varepsilon F_1 (\partial_{x_1} (\varepsilon_1 m_1) + \varepsilon_1 i k_0 m_2) A + \text{c.c.} + \mathcal{O}(\varepsilon^2) = \mathcal{O}(\varepsilon^2)$$

because  $\partial_{x_1} (\varepsilon_1 m_1) + \varepsilon_1 i k_0 m_2 = 0$ , see (19) at  $k = k_0$ .

Regarding the interface conditions,  $\mathbf{U}_{\text{ans}}$  and  $\mathbf{U}_{\text{ext}}$  fulfill (9) exactly since the second and third components of  $\mathbf{m}$ ,  $\partial_k \mathbf{w}(k_0)$ ,  $\partial_k^2 \mathbf{w}(k_0)$ ,  $\mathbf{h}$  and  $\mathbf{p}$  belong to  $H^1(\mathbb{R})$  and are therefore continuous at the interface. Moreover, the jumps of  $\mathcal{D}_1(\mathcal{E}_{\text{ans}})$  and  $\mathcal{D}_1(\mathcal{E}_{\text{ext}})$  at  $x_1 = 0$  are of order  $\mathcal{O}(\varepsilon^3)$ , respectively  $\mathcal{O}(\varepsilon^4)$ . Indeed, at  $\mathcal{O}(\varepsilon)$  condition (11) holds exactly for  $\mathbf{U}_{\text{ans}}$  and  $\mathbf{U}_{\text{ext}}$  because  $\mathbf{m}$  satisfies the interface conditions (18). At  $\mathcal{O}(\varepsilon^2)$  there are no contributions for  $\mathbf{U}_{\text{ans}}$ . For  $\mathbf{U}_{\text{ext}}$  only a linear term involving  $\partial_k \mathbf{w}(\cdot, k_0)$  appears. As (18) holds for each  $k$ , we have  $[[\varepsilon_1 \partial_k w_1(\cdot, k_0)]]_{ID} = 0$  and also  $[[\varepsilon_1 \partial_k^2 w_1(\cdot, k_0)]]_{ID} = 0$ , and hence  $\mathbf{U}_{\text{ext}}$  satisfies (11) at  $\mathcal{O}(\varepsilon^2)$ . Finally, (11) holds for  $\mathbf{U}_{\text{ext}}$  also at  $\mathcal{O}(\varepsilon^3)$  since  $\mathbf{h} \in D(L(3k_0, 3\nu_0))$  and  $\mathbf{p} \in D(L(k_0, \nu_0))$  solve (34) and (39), respectively. This fact and the continuity of  $p_1$  and  $h_1$  imply the jump equations

$$[[\varepsilon_1 p_1]]_{ID} = -[[\varepsilon_3 (3m_1^3 - m_1 m_2^2)]]_{ID} \quad \text{and} \quad [[\varepsilon_1 h_1]]_{ID} = -[[\varepsilon_3 (m_1^3 + m_1 m_2^2)]]_{ID},$$

implying that the first nonlinear contribution in (11) for  $\mathbf{U}_{\text{ext}}$  is cancelled.

As we explain next,  $\mathbf{Res}(\mathbf{U}_{\text{ext}})(\cdot, \cdot, t)$  lies in  $L^2(\mathbb{R}^2)^3$  for all  $t \in [0, T_0 \varepsilon^{-2}]$  if, e.g.,

$$A \in \bigcap_{k=0}^1 C^{1-k}([0, T_0], H^{2+k}(\mathbb{R})) \quad \text{and} \quad \mathbf{m}, \partial_k \mathbf{w}(\cdot, k_0), \partial_k^2 \mathbf{w}(\cdot, k_0), \mathbf{h}, \mathbf{p} \in L^2(\mathbb{R})^3 \cap L^\infty(\mathbb{R})^3. \quad (41)$$

So far we have used the  $\varepsilon$ -orders in a formal sense, i.e., without specifying the norm. To determine the asymptotic order of the  $L^2(\mathbb{R}^2)^3$ -norm of the residual, we note that the summands of  $\mathbf{Res}(\mathbf{U}_{\text{ext}})(\cdot, \cdot, t)$  of (the smallest) order  $\varepsilon^4$  have the form  $g(\mathbf{x}) := f_1(x_1)f_2(\varepsilon x_2)f_3(x_2)$  with  $f_1, f_2 \in L^2(\mathbb{R})$  and  $f_3 \in L^\infty(\mathbb{R})$ . Such products can be estimated by

$$\|g\|_{L^2(\mathbb{R}^2)} \leq \varepsilon^{-\frac{1}{2}} \|f_1\|_{L^2(\mathbb{R})} \|f_2\|_{L^2(\mathbb{R})} \|f_3\|_{L^\infty(\mathbb{R})}.$$

Terms of higher order in  $\varepsilon$  are of a similar form. The  $x_1$ -derivatives do not appear in the residual, since they only occur at low orders of  $\varepsilon$  and were canceled in the construction. Hence, in (41) we have to use Sobolev spaces only for  $A$ . In the residual the derivatives of highest order are  $\partial_{X_2}^3 A$  and  $\partial_T \partial_{X_2}^2 A$  which are bounded in  $L^2(\mathbb{R})^3$  uniformly in time according to (41). Due the embedding  $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$  the lower-order factors are bounded. The products appearing in nonlinear terms are estimated by (suppressing the time dependence)

$$\begin{aligned} \| |A|^2 \partial_{X_2} A \|_{L^2(\mathbb{R})} &\leq \|A\|_{L^\infty(\mathbb{R})}^2 \|\partial_{X_2} A\|_{L^2(\mathbb{R})} \leq C \|A\|_{H^1(\mathbb{R})}^3, \\ \| |\partial_{X_2}^2 A|^2 \partial_{X_2}^2 A \|_{L^2(\mathbb{R})} &\leq \|\partial_{X_2}^2 A\|_{L^\infty(\mathbb{R})}^2 \|\partial_{X_2}^2 A\|_{L^2(\mathbb{R})} \leq C \|A\|_{H^3(\mathbb{R})}^3, \\ \| |\partial_{X_2}^2 A|^2 \partial_T \partial_{X_2}^2 A \|_{L^2(\mathbb{R})} &\leq \|\partial_{X_2}^2 A\|_{L^\infty(\mathbb{R})}^2 \|\partial_T \partial_{X_2}^2 A\|_{L^2(\mathbb{R})} \leq C \|A\|_{H^3(\mathbb{R})}^2 \|\partial_T A\|_{H^2(\mathbb{R})}, \end{aligned}$$

for instance. Using these principles, we easily deduce

$$\|\mathbf{Res}(\mathbf{U}_{\text{ext}})(\cdot, \cdot, t)\|_{L^2(\mathbb{R}^2)^3} = \|\widehat{\mathbf{Res}}(\widehat{\mathbf{U}}_{\text{ext}})(\cdot, \cdot, t)\|_{L^2(\mathbb{R}^2)^3} \leq C \varepsilon^{\frac{7}{2}} \quad (42)$$

under condition (41). The constant  $C$  depends on the norms of  $A$ ,  $\mathbf{m}$ ,  $\partial_k \mathbf{w}(\cdot, k_0)$ ,  $\partial_k^2 \mathbf{w}(\cdot, k_0)$ ,  $\mathbf{h}$ , and  $\mathbf{p}$  appearing in (41).

For the error analysis in Section 6 we have to estimate the residual in  $\mathcal{G}^3(\mathbb{R}^2 \times [0, T_0 \varepsilon^{-2}])^3$  and not only in  $L^2(\mathbb{R}^2)^3$  for all  $t \in [0, T_0 \varepsilon^{-2}]$ . To do this, we impose the stricter conditions

$$A \in \bigcap_{k=0}^4 C^{4-k}([0, T_0], H^{2+k}(\mathbb{R})) \quad \text{and} \quad \mathbf{m}, \partial_k \mathbf{w}(\cdot, k_0), \partial_k^2 \mathbf{w}(\cdot, k_0), \mathbf{h}, \mathbf{p} \in \mathcal{H}^3(\mathbb{R})^3. \quad (43)$$

Since we want to estimate the derivatives up to order three of the residual, it is clear that the regularity of the envelope  $A$  has also to increase by three orders in space and time. Since no  $x_1$ -derivative appears in the residual we can use the algebra property of  $\mathcal{H}^3(\mathbb{R})$  to control the appearing nonlinear terms. With the same arguments as before it is now possible to bound the residual by

$$\|\mathbf{Res}(\mathbf{U}_{\text{ext}})\|_{\mathcal{G}^3(\mathbb{R}^2 \times (0, T_0 \varepsilon^{-2}))^3} \leq C \varepsilon^{\frac{7}{2}} \quad (44)$$

under the above conditions (43). The constant  $C$  depends on the norms of  $A$ ,  $\mathbf{m}$ ,  $\partial_k \mathbf{w}(\cdot, k_0)$ ,  $\partial_k^2 \mathbf{w}(\cdot, k_0)$ ,  $\mathbf{h}$ , and  $\mathbf{p}$  appearing in (43). Furthermore, under condition (43) we have

$$\mathbf{U}_{\text{ext}} \in \bigcap_{k=0}^4 C^{4-k}([0, T_0 \varepsilon^{-2}], \mathcal{H}^{\min\{3; k\}}(\mathbb{R}^2))^3.$$

In the bootstrapping argument of Section 6.1 we need somewhat stronger regularity properties of  $\mathbf{U}_{\text{ext}}$ . They follow from the structure of  $\mathbf{U}_{\text{ext}}$ , which is a sum of products of functions in  $(x_2, t)$  and in  $x_1$ , where the latter only appear linearly. So we can use the Sobolev embedding  $\mathcal{H}^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$  in both space dimensions separately, thus avoiding the less favorable embedding  $\mathcal{H}^2(\mathbb{R}^2) \hookrightarrow L^\infty(\mathbb{R}^2)$ . Take multi-indices

$\alpha = (\alpha_1, \alpha_2, \alpha_t)^\top$  with  $|\alpha| \leq 3$  and  $\alpha_1 \leq 2$ , as well as  $|\beta| = 3$  with  $\beta_1 \in \{1, 2\}$ . For  $\partial^\alpha := \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \partial_t^{\alpha_t}$  condition (43) implies that

$$\|\partial^\alpha \mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times (0, T_0 \varepsilon^{-2}))^3}, \|\partial^\beta \partial_t \mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times (0, T_0 \varepsilon^{-2}))^3} \leq C\varepsilon, \quad (45)$$

and

$$\int_{\mathbb{R}} \sup_{x_2 \in \mathbb{R}} |\partial_{x_1}^3 \partial_t^k \mathbf{U}_{\text{ext}}(x_1, x_2, t)|^2 dx_1 \leq C\varepsilon^2, \quad k \in \{0, 1\}, \quad (46)$$

for all  $t \in [0, T_0 \varepsilon^{-2}]$ . For  $|\gamma| = 3$  with  $\gamma_1 = 0$  we have  $\partial^\gamma \partial_t \mathbf{U}_{\text{ext}} = \mathcal{A} + \mathcal{B}$  with

$$\begin{aligned} \mathcal{A}(x_1, x_2, t) &:= \varepsilon A(X_2, T) \mathbf{m}(x_1) \partial^\gamma \partial_t \left( e^{i(k_0 x_2 - \nu_0 t)} \right), \\ \mathcal{B}(x_1, x_2, t) &:= \partial^\gamma \partial_t \mathbf{U}_{\text{ext}}(x_1, x_2, t) - \mathcal{A}(x_1, x_2, t), \end{aligned}$$

and one derives the estimates

$$\|\mathcal{A}\|_{L^\infty(\mathbb{R}^2 \times (0, T_0 \varepsilon^{-2}))^3} \leq C\varepsilon, \quad \int_{\mathbb{R}} \sup_{x_1 \in \mathbb{R}} |\mathcal{B}(x_1, x_2, t)|^2 dx_2 \leq C\varepsilon^2. \quad (47)$$

An application of Lemma 2.6 gives us the necessary regularity of  $\mathbf{m}$ ,  $\partial_k \mathbf{w}(\cdot, k_0)$ ,  $\partial_k^2 \mathbf{w}(\cdot, k_0)$ ,  $\mathbf{h}$ , and  $\mathbf{p}$  under our assumptions on  $\varepsilon_1$  and  $\varepsilon_3$ .

#### Lemma 4.2

Let  $\mathbf{m}, \partial_k \mathbf{w}(k_0), \partial_k^2 \mathbf{w}(k_0), \mathbf{h}, \mathbf{p} \in L^2(\mathbb{R})^3$  be defined as before. Assume that  $\varepsilon_1, \varepsilon_3 \in \mathcal{W}^{3, \infty}(\mathbb{R})$ . Then  $\mathbf{m}, \partial_k \mathbf{w}(k_0), \partial_k^2 \mathbf{w}(k_0), \mathbf{p}, \mathbf{h} \in \mathcal{H}^3(\mathbb{R})^3$ .

PROOF: Since  $(L(k_0) + \nu_0 \Lambda) \mathbf{m} = \mathbf{0}$ , Lemma 2.6 with  $\mathbf{f} = \mathbf{0}$  shows that  $\mathbf{m} \in \mathcal{H}^3(\mathbb{R})^3$ . Next, differentiating  $(L(k) + \omega \Lambda) \mathbf{w} = \mathbf{0}$  in  $k$ , we see that  $\partial_k \mathbf{w}(k_0)$  and  $\partial_k^2 \mathbf{w}(k_0)$  solve

$$\begin{aligned} (L(k_0) + \nu_0 \Lambda) \partial_k \mathbf{w}(k_0) &= -(\partial_k L(k_0) + \partial_k \omega(k_0) \Lambda) \mathbf{m}, \\ (L(k_0) + \nu_0 \Lambda) \partial_k^2 \mathbf{w}(k_0) &= -2(\partial_k L(k_0) + \partial_k \omega(k_0) \Lambda) \partial_k \mathbf{w}(k_0) - (\partial_k^2 L(k_0) + \partial_k^2 \omega(k_0) \Lambda) \mathbf{m}. \end{aligned}$$

As  $\mathbf{m} \in \mathcal{H}^3(\mathbb{R})^3$ , the functions

$$(\partial_k L(k_0) + \partial_k \omega(k_0) \Lambda) \mathbf{m} = \begin{pmatrix} \varepsilon_1 \nu_1 m_1 + m_3 \\ \varepsilon_1 \nu_1 m_2 \\ m_1 + \mu_0 \nu_1 m_3 \end{pmatrix}, \quad (\partial_k^2 L(k_0) + \partial_k^2 \omega(k_0) \Lambda) \mathbf{m} = \begin{pmatrix} \varepsilon_1 \nu_2 m_1 \\ \varepsilon_1 \nu_2 m_2 \\ \mu_0 \nu_2 m_3 \end{pmatrix}$$

belong to  $\mathcal{H}^3(\mathbb{R})^3$ . Therefore from Lemma 2.6 we infer  $\partial_k \mathbf{w}(k_0) \in \mathcal{H}^3(\mathbb{R})^3$ . This fact implies  $\partial_k^2 \mathbf{w}(k_0) \in \mathcal{H}^3(\mathbb{R})^3$  in the same way. To treat  $\mathbf{h}$  and  $\mathbf{p}$  from (34) and (39), we note that the right-hand sides in

$$\begin{aligned} (L(k_0) + \nu_0 \Lambda) \mathbf{p} &= -\kappa \begin{pmatrix} \varepsilon_1 m_1 \\ \varepsilon_1 m_2 \\ \mu_0 m_3 \end{pmatrix} - \begin{pmatrix} \varepsilon_3 \nu_0 (3m_1^3 - m_1 m_2^2) \\ \varepsilon_3 \nu_0 (-3m_2^3 + m_1^2 m_2) \\ 0 \end{pmatrix}, \\ (L(3k_0) + 3\nu_0 \Lambda) \mathbf{h} &= \begin{pmatrix} -i\varepsilon_3 (m_1^3 + m_1 m_2^2) \\ -i\varepsilon_3 (m_2^3 + m_2 m_1^2) \\ 0 \end{pmatrix} \end{aligned}$$

are also contained in  $\mathcal{H}^3(\mathbb{R})^3$  since  $\mathbf{m} \in \mathcal{H}^3(\mathbb{R})^3$ . Hence, the statement follows as before.  $\square$

## 5 Local existence theory

We employ local existence results of [23] for linear and quasilinear hyperbolic problems. We first define some additional function spaces. For any open  $\Omega \subset \mathbb{R}^2$  and  $J \subset \mathbb{R}$  we use

$$F^{m, n}(\Omega \times J) := \left\{ A \in W^{1, \infty}(\Omega \times J)^{n \times n} \mid \partial^\alpha A \in L^\infty(J, H^{m-|\alpha|}(\Omega))^{n \times n} + W^{m-|\alpha|, \infty}(\Omega \times J)^{n \times n} \right\}$$



for all  $\alpha \in \mathbb{N}_0^3$  with  $1 \leq |\alpha| \leq m$ ,

$$\|A\|_{F^{m,n}(\Omega \times J)} := \max \left\{ \|A\|_{W^{1,\infty}(\Omega \times J)^{n \times n}}, \max_{1 \leq |\alpha| \leq m} \|\partial^\alpha A\|_{L^\infty(J, H^{m-|\alpha|}(\Omega))^{n \times n} + W^{m-|\alpha|,\infty}(\Omega \times J)^{n \times n}} \right\},$$

$$\mathcal{F}^{m,n}(\mathbb{R}^2 \times J) := \{A \in \mathcal{W}^{1,\infty}(\mathbb{R}^2 \times J)^{n \times n} \mid A^- \in F^{m,n}(\mathbb{R}_-^2 \times J), A^+ \in F^{m,n}(\mathbb{R}_+^2 \times J)\},$$

$$\|A\|_{\mathcal{F}^{m,n}(\mathbb{R}^2 \times J)} := \max \left\{ \|A^-\|_{F^{m,n}(\mathbb{R}_-^2 \times J)}, \|A^+\|_{F^{m,n}(\mathbb{R}_+^2 \times J)} \right\},$$

with the usual definition

$$\|A\|_{L^\infty(J, H^{m-|\alpha|}(\Omega)) + W^{m-|\alpha|,\infty}(\Omega \times J)} := \inf \left\{ \|B\|_{L^\infty(J, H^{m-|\alpha|}(\Omega))} + \|C\|_{W^{m-|\alpha|,\infty}(\Omega \times J)} \mid A = B + C, \right. \\ \left. B \in L^\infty(J, H^{m-|\alpha|}(\Omega)), C \in W^{m-|\alpha|,\infty}(\Omega \times J) \right\}.$$

For a fixed time instant we use the spaces

$$F_0^{m,n}(\Omega) := \left\{ A \in L^\infty(\Omega)^{n \times n} \mid \partial^\alpha A \in H^{m-|\alpha|}(\Omega)^{n \times n} + W^{m-|\alpha|,\infty}(\Omega)^{n \times n} \right. \\ \left. \text{for all } \alpha \in \mathbb{N}_0^2 \text{ with } 1 \leq |\alpha| \leq m \right\},$$

$$\|A\|_{F_0^{m,n}(\Omega)} := \max \left\{ \|A\|_{L^\infty(\Omega)^{n \times n}}, \max_{1 \leq |\alpha| \leq m} \|\partial^\alpha A\|_{H^{m-|\alpha|}(\Omega)^{n \times n} + W^{m-|\alpha|,\infty}(\Omega)^{n \times n}} \right\},$$

$$\mathcal{F}_0^{m,n}(\mathbb{R}^2) := \{A \in L^\infty(\mathbb{R}^2)^{n \times n} \mid A^- \in F_0^{m,n}(\mathbb{R}_-^2), A^+ \in F_0^{m,n}(\mathbb{R}_+^2)\},$$

$$\|A\|_{\mathcal{F}_0^{m,n}(\mathbb{R}^2)} := \max \left\{ \|A^-\|_{F_0^{m,n}(\mathbb{R}_-^2)}, \|A^+\|_{F_0^{m,n}(\mathbb{R}_+^2)} \right\}.$$

Finally, we define subspaces of  $\mathcal{F}^{m,n}$  with the properties of positive definiteness, or convergence to a constant matrix for  $|\mathbf{x}, t| \rightarrow \infty$ . For  $\eta > 0$  we set

$$\mathcal{F}_\eta^{m,n}(\mathbb{R}^2 \times J) := \{A \in \mathcal{F}^{m,n}(\mathbb{R}^2 \times J) \mid A = A^\top, \mathbf{v}^\top A \mathbf{v} \geq \eta |\mathbf{v}|^2 \text{ for all } \mathbf{v} \in \mathbb{R}^n\},$$

$$\mathcal{F}_{\text{cv}}^{m,n}(\mathbb{R}^2 \times J) := \left\{ A \in \mathcal{F}^{m,n}(\mathbb{R}^2 \times J) \mid \exists \tilde{A} \in \mathbb{R}^{n \times n} : \lim_{|\mathbf{x}, t| \rightarrow \infty} A(\mathbf{x}, t) = \tilde{A} \right\},$$

$$\mathcal{F}_{\eta, \text{cv}}^{m,n}(\mathbb{R}^2 \times J) := \mathcal{F}_\eta^{m,n}(\mathbb{R}^2 \times J) \cap \mathcal{F}_{\text{cv}}^{m,n}(\mathbb{R}^2 \times J).$$

In the linear setting, as in (4.1) of [23] we rewrite our system as

$$\left\{ \begin{array}{l} A_t(\mathbf{x}, t) \partial_t \mathbf{U}^\pm + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{U}^\pm + M(\mathbf{x}, t) \mathbf{U}^\pm = \mathbf{f}^\pm, \quad \mathbf{x} \in \mathbb{R}_\pm^2, \quad t \in J, \\ B_\Gamma \begin{pmatrix} \mathbf{U}^+ \\ \mathbf{U}^- \end{pmatrix} = \mathbf{0}, \quad \mathbf{x} \in \Gamma, \quad t \in J, \\ \mathbf{U}(0) = \mathbf{U}^{(0)}, \quad \mathbf{x} \in \mathbb{R}^2 \end{array} \right. \quad (48)$$

on the interval  $J := (0, T')$  with some  $T' > 0$ , where  $M : \mathbb{R}^2 \times J \rightarrow \mathbb{R}^{3 \times 3}$ ,  $A_t : \mathbb{R}^2 \times J \rightarrow \mathbb{R}^{3 \times 3}$ ,  $A_t(\mathbf{x}, t)$  is symmetric for all  $(\mathbf{x}, t)$ , and

$$A_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 := \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad B_\Gamma := \begin{pmatrix} 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix},$$

Clearly,  $B_\Gamma(\mathbf{U}^+, \mathbf{U}^-)^\top$  encodes the interface conditions  $[[U_2]] = [[U_3]] = 0$  on  $\Gamma$ . Note that we are not going to use (48) in order to study the linear part of (7) but rather to study a fixed point problem in the bootstrapping argument for the nonlinear system in Section 6.1. Hence, we need the inhomogeneous term  $\mathbf{f}$  as well as the linear term  $M\mathbf{U}$  in (48).

**Definition 5.1** (Weak Solution of the Linear Hyperbolic Problem)

Under a weak solution of (48) we understand a function  $\mathbf{U} \in C(\bar{J}, L^2(\mathbb{R}^2))^3$  that satisfies

$$\int_J \int_{\mathbb{R}^2} \mathbf{f} \cdot \boldsymbol{\varphi} \, d\mathbf{x} \, dt = - \int_J \int_{\mathbb{R}^2} (\mathbf{U} \cdot \partial_t(A_t \boldsymbol{\varphi}) + \mathbf{U} \cdot \partial_{x_1}(A_1 \boldsymbol{\varphi}) + \mathbf{U} \cdot \partial_{x_2}(A_2 \boldsymbol{\varphi}) - \mathbf{U} \cdot M^\top \boldsymbol{\varphi}) \, d\mathbf{x} \, dt$$

for all test functions

$$\boldsymbol{\varphi} \in \mathcal{D}(\mathbb{R}^2 \times J)^3 := \{\boldsymbol{\varphi} \mid \boldsymbol{\varphi}^+ \in C^\infty(\mathbb{R}_+^2 \times J)^3, \boldsymbol{\varphi}^- \in C^\infty(\mathbb{R}_-^2 \times J)^3 \mid \text{supp } \boldsymbol{\varphi} \subset \mathbb{R}^2 \times J \text{ compact}\},$$

$$\text{Tr}_{\Gamma \times J} \left( B_\Gamma (\mathbf{U}^+, \mathbf{U}^-)^\top \right) = \mathbf{0}, \text{ and } \mathbf{U}(0) = \mathbf{U}^{(0)}.$$

For the trace in the above definition, note that for a weak solution  $\boldsymbol{\varphi}$  of (48) the space-time divergence of  $(A_1 \boldsymbol{\varphi}_1, A_2 \boldsymbol{\varphi}_2, A_t \boldsymbol{\varphi}_t)$  belongs to  $L^2(\mathbb{R}_\pm^2 \times J)^3$  and thus has traces in  $H^{-1/2}(\partial(\mathbb{R}_\pm^2 \times J))^3$ . This fact and the properties of  $A_j$  allow us to define the initial and interface conditions in  $H^{-1/2}(\mathbb{R}^2 \times \{0\})$  and  $H^{-1/2}(\Gamma \times J)^3$ , respectively. See Chapter 2.1 in [27] for an in depth discussion of this trace.

The following existence result is a consequence of Theorem 3.1 in [23].

**Theorem 5.2** (Existence result for the linear system (48))

Let  $\eta, T', r > 0$ ,  $m \in \{0, 1, 2, 3\}$ , and  $J = (0, T')$ . Take coefficients  $A_t \in \mathcal{F}_{\eta, \text{cv}}^{3,3}(\mathbb{R}^2 \times J)$  and  $M \in \mathcal{F}_{\text{cv}}^{3,3}(\mathbb{R}^2 \times J)$  with

$$\begin{aligned} \|A_t\|_{\mathcal{F}^{3,3}(\mathbb{R}^2 \times J)}, \|A_t(\cdot, 0)\|_{\mathcal{F}_0^{2,3}(\mathbb{R}^2)}, \left\| \partial_t^j A_t(\cdot, 0) \right\|_{\mathcal{H}^{2-j}(\mathbb{R}^2)^{3 \times 3}} &\leq r, \\ \|M\|_{\mathcal{F}^{3,3}(\mathbb{R}^2 \times J)}, \|M(\cdot, 0)\|_{\mathcal{F}_0^{2,3}(\mathbb{R}^2)}, \left\| \partial_t^j M(\cdot, 0) \right\|_{\mathcal{H}^{2-j}(\mathbb{R}^2)^{3 \times 3}} &\leq r, \end{aligned}$$

for all  $j \in \{1, 2\}$ . Choose  $\mathbf{f} \in \mathcal{H}^m(\mathbb{R}^2 \times J)^3$  and  $\mathbf{U}^{(0)} \in \mathcal{H}^m(\mathbb{R}^2)^3$  such that the linear compatibility conditions of order  $m$  are satisfied, see Definition 5.9.

Then there is a unique weak solution  $\mathbf{U}$  of (48) in  $\mathcal{G}^m(\mathbb{R}^2 \times J)^3$  and a constant  $C_m = C_m(\eta, r, T') \geq 1$  such that

$$\|\mathbf{U}\|_{\mathcal{G}^m(\mathbb{R}^2 \times J)^3}^2 \leq C_m \left( \left\| \mathbf{U}^{(0)} \right\|_{\mathcal{H}^m(\mathbb{R}^2)^3}^2 + \|\mathbf{f}\|_{\mathcal{H}^m(\mathbb{R}^2 \times J)^3}^2 + \sum_{j=0}^{m-1} \left\| \partial_t^j \mathbf{f}(\cdot, 0) \right\|_{\mathcal{H}^{m-1-j}(\mathbb{R}^2)^3}^2 \right) \quad (49)$$

where, as usual, the sum is empty if  $m = 0$ .

**Remark 5.3**

Theorem 3.1 of [23] deals with spatial domains in  $\mathbb{R}^3$  instead of  $\mathbb{R}^2$  and the solution vector takes values in  $\mathbb{R}^6$  instead of  $\mathbb{R}^3$ , but the above case can be treated in an analogous and simpler way. We do not need the lengthy localization process discussed in [23]. As in this paper, one reduces the interface problem on  $\mathbb{R}^2 \setminus \Gamma$  to a boundary value problem on  $\mathbb{R}_+^2$  and the latter can again be solved in  $L^2$  by means of general results from [13]. To obtain solutions in  $\mathcal{G}^m$  for appropriate data, one first needs a priori estimates. These can be shown as in Section 6.1 below which uses ideas from [23]. The regularity of solutions can then be shown by approximation arguments which are simplified versions of those in [23].

In the bootstrapping argument in Section 6.1 we need the next approximation result, involving the space

$$\mathcal{D}(\mathbb{R}^2)^3 := \{\boldsymbol{\varphi} \mid \boldsymbol{\varphi}^+ \in C^\infty(\mathbb{R}_+^2)^3, \boldsymbol{\varphi}^- \in C^\infty(\mathbb{R}_-^2)^3, \text{supp } \boldsymbol{\varphi} \subset \mathbb{R}^2 \text{ compact}\}.$$

**Lemma 5.4**

Let  $T' > 0$ ,  $J = (0, T')$ ,  $\mathbf{U}^{(0)} \in L^2(\mathbb{R}^2)^3$ ,  $A_t \in \mathcal{F}_{\eta, \text{cv}}^{3,3}(\mathbb{R}^2 \times J)$ ,  $M \in \mathcal{F}_{\text{cv}}^{3,3}(\mathbb{R}^2 \times J)$  and  $\mathbf{f} \in \mathcal{G}^0(\mathbb{R}^2 \times J)^3$ . Take a weak solution  $\mathbf{U} \in \mathcal{G}^0(\mathbb{R}^2 \times J)^3$  of (48) for the data  $(\mathbf{f}, \mathbf{U}^{(0)})$ . Then the following statements are true.

i) There are sequences  $(\mathbf{U}_n^{(0)})_n \subset \mathcal{D}(\mathbb{R}^2)^3$  and  $(\mathbf{f}_n)_n \subset \mathcal{H}^1(\mathbb{R}^2 \times J)^3$  such that  $\mathbf{U}_n^{(0)} \rightarrow \mathbf{U}^{(0)}$  in  $L^2(\mathbb{R}^2)^3$  and  $\mathbf{f}_n \rightarrow \mathbf{f}$  in  $L^2(\mathbb{R}^2 \times J)^3$  for  $n \rightarrow \infty$  and the linear compatibility conditions of order 1 are satisfied, i.e.,  $B_\Gamma \mathbf{U}_n^{(0)} = \mathbf{0}$ .

ii) There exists a sequence  $(\mathbf{U}_n)_n \subset \mathcal{G}^1(\mathbb{R}^2 \times J)^3$  such that for all  $n \in \mathbb{N}$  the function  $\mathbf{U}_n$  solves (48) for the data  $(\mathbf{f}_n, \mathbf{U}_n^{(0)})$  and  $\mathbf{U}_n \rightarrow \mathbf{U}$  in  $\mathcal{G}^0(\mathbb{R}^2 \times J)^3$  for  $n \rightarrow \infty$ .

PROOF: For i) we use that  $C_c^\infty(\Omega)$  and  $H^1(\Omega)$  are dense in  $L^2(\Omega)$  for any domain  $\Omega$ . Therefore we can choose sequences  $(\mathbf{w}_n)_n \subset \mathcal{D}(\mathbb{R}^2)^3$  and  $(\mathbf{f}_n)_n \subset \mathcal{H}^1(\mathbb{R}^2 \times J)^3$  with  $\mathbf{w}_n \rightarrow \mathbf{U}^{(0)}$  in  $L^2(\mathbb{R}^2)^3$  and  $\mathbf{f}_n \rightarrow \mathbf{f}$  in  $L^2(\mathbb{R}^2 \times J)^3$  for  $n \rightarrow \infty$ . To guarantee the compatibility condition, we introduce the characteristic function  $\chi_{M_n}$  with  $M_n := \mathbb{R}^2 \setminus \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 \in [-\frac{1}{n}, \frac{1}{n}]\}$ . By dominated convergence, one easily sees that  $\mathbf{w}_n \chi_{M_n} \rightarrow \mathbf{U}^{(0)}$  in  $L^2(\mathbb{R}^2)^3$ . Since  $\Gamma \cap M_n = \emptyset$  we also get that  $\text{Tr}_\Gamma(B_\Gamma \mathbf{w}_n \chi_{M_n}) = \mathbf{0}$ . Now we mollify  $\mathbf{w}_n \chi_{M_n}$  to produce functions  $\mathbf{U}_n^{(0)} \in \mathcal{D}(\mathbb{R}^2)^3$  with the stated properties.

The existence of  $(\mathbf{U}_n)_n$  in assertion ii) is a direct consequence of Theorem 5.2. To show the convergence we use that (48) is a linear problem, consequently  $\mathbf{U}_n - \mathbf{U}$  is a weak solution of (48) for the data  $(\mathbf{f}_n - \mathbf{f}, \mathbf{U}_n^{(0)} - \mathbf{U}^{(0)})$ . Estimate (49) thus yields

$$\|\mathbf{U}_n - \mathbf{U}\|_{\mathcal{G}^0(\mathbb{R}^2 \times J)^3} \leq C \left( \left\| \mathbf{U}_n^{(0)} - \mathbf{U}^{(0)} \right\|_{L^2(\mathbb{R}^2)^3} + \|\mathbf{f}_n - \mathbf{f}\|_{L^2(\mathbb{R}^2 \times J)^3} \right).$$

The convergence properties of  $(\mathbf{f}_n)_n$  and  $(\mathbf{U}_n^{(0)})_n$  complete the proof.  $\square$

For the quasilinear case we have to define spaces for functions whose range is not the full  $\mathbb{R}^3$ , namely

$$\begin{aligned} \mathcal{ML}^{m,k}(\mathbb{R}^2, \Omega_\pm) &:= \left\{ S : (\mathbb{R}_+^2 \times \Omega_+) \cup (\mathbb{R}_-^2 \times \Omega_-) \rightarrow \mathbb{R}^{k \times k} \mid S^\pm \in C^m(\mathbb{R}_\pm^2 \times \Omega_\pm, \mathbb{R}^{k \times k}), \right. \\ &\quad \left. \sup_{(\mathbf{x}, \mathbf{u}) \in \mathbb{R}_\pm^2 \times \Omega_\pm} |\partial^\alpha S(\mathbf{x}, \mathbf{u})| < \infty \text{ for all compact } \mathcal{U}_\pm \subset \Omega_\pm \text{ and } \alpha \in \mathbb{N}_0^5 \text{ with } |\alpha| \leq m \right\}, \end{aligned}$$

where  $\Omega_\pm \subset \mathbb{R}^3$  are open and  $S^+$  and  $S^-$  are the restrictions of  $S$  to  $\mathbb{R}_+^2 \times \Omega_+$  and  $\mathbb{R}_-^2 \times \Omega_-$ , respectively. As for the spaces  $\mathcal{F}$  we use the subscripts  $\eta$  and  $cv$  to denote the additional conditions that the matrices in  $\mathcal{ML}^{m,k}$  are symmetric and positive definite respectively convergent.

The reduced nonlinear Maxwell system (7), (8), (9) is a special case of equation (1.7) in [23] and can be written as

$$\begin{cases} \tilde{S}(\mathbf{x}, \mathbf{U}^\pm) \partial_t \mathbf{U}^\pm + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{U}^\pm = \mathbf{0}, & \mathbf{x} \in \mathbb{R}_\pm^2, t \in J, \\ B_\Gamma \begin{pmatrix} \mathbf{U}^+ \\ \mathbf{U}^- \end{pmatrix} = \mathbf{0}, & \mathbf{x} \in \Gamma, t \in J, \\ \mathbf{U}(0) = \mathbf{U}^{(0)}, & \mathbf{x} \in \mathbb{R}^2, \end{cases} \quad (50)$$

where for  $\mathbf{v} \in \mathbb{R}^n$  we set

$$\begin{aligned} \tilde{S}(\mathbf{x}, \mathbf{v}) &:= \Lambda(x_1) + \epsilon_3(x_1) \theta(\mathbf{v}), \\ \Lambda(x_1) &= \begin{pmatrix} \epsilon_1(x_1) & 0 & 0 \\ 0 & \epsilon_1(x_1) & 0 \\ 0 & 0 & \mu_0 \end{pmatrix}, \quad \theta(\mathbf{v}) := \begin{pmatrix} 3v_1^2 + v_2^2 & 2v_1 v_2 & 0 \\ 2v_1 v_2 & v_1^2 + 3v_2^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (51)$$

**Definition 5.5** (Solution of the Nonlinear Hyperbolic Problem)

Under a solution of (50) we understand a function  $\mathbf{U} \in \mathcal{G}^1(\mathbb{R}^2 \times J)^3 \cap L^\infty(\mathbb{R}^2 \times J)^3$  with  $\overline{\text{im } \mathbf{U}^\pm} \subset \Omega_\pm$  that satisfies

$$\tilde{S}(\mathbf{x}, \mathbf{U}) \partial_t \mathbf{U} + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{U} = \mathbf{0},$$

for almost all  $\mathbf{x} \in \mathbb{R}^2 \setminus \Gamma$  and for all  $t \in J$ ,  $\text{Tr}_{\Gamma \times J} \left( B_\Gamma \begin{pmatrix} \mathbf{U}^+ \\ \mathbf{U}^- \end{pmatrix} \right) = \mathbf{0}$ , and  $\mathbf{U}(0) = \mathbf{U}^{(0)}$ .

**Remark 5.6**

Note that a solution  $\mathbf{U}$  of (50) in  $\mathcal{G}^3(\mathbb{R}^2 \times J)^3$  is a classical solution of (50) because of the Sobolev embedding  $H^3(\mathbb{R}_\pm^2) \hookrightarrow C^1(\mathbb{R}_\pm^2)$ .

The following local existence result for (50) follows from Proposition 7.1 and Theorem 7.1 of [23].

**Theorem 5.7 (Nonlinear Existence Result)**

Let  $\eta > 0$ ,  $\Omega_\pm \subset \mathbb{R}^3$ , and  $\tilde{S} \in \mathcal{ML}_{\eta, \text{cv}}^{3,3}(\mathbb{R}^2, \Omega_\pm)$ . Assume that  $\mathbf{U}^{(0)} \in \mathcal{H}^3(\mathbb{R}^2)^3$  satisfies the nonlinear compatibility conditions of order 3, see (56), and  $\text{im } \mathbf{U}^{(0), \pm} \subset \Omega_\pm$  with

$$\text{dist} \left( \overline{\text{im } \mathbf{U}^{(0), \pm}}, \partial\Omega_\pm \right) > \kappa \quad (52)$$

for some  $\kappa > 0$ . Then the following statements are true.

- i) There exists a unique solution  $\mathbf{U} \in \mathcal{G}^3(\mathbb{R}^2 \times (0, t_M))^3$  of (50), where  $t_M > 0$  is the maximal existence time.
- ii) If  $t_M < \infty$ , then  $\lim_{t \nearrow t_M} \|\mathbf{U}(\cdot, t)\|_{\mathcal{H}^3(\mathbb{R}^2)^3} = \infty$  or  $\liminf_{t \nearrow t_M} \text{dist} \left( \overline{\text{im } \mathbf{U}^\pm(t)}, \partial\Omega_\pm \right) = 0$ .
- iii) Let  $t^* \in (0, t_M)$ . Then there is a constant  $C > 0$  such that

$$\|\mathbf{U}\|_{\mathcal{G}^3(\mathbb{R}^2 \times (0, t^*))^3} \leq C \left\| \mathbf{U}^{(0)} \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3}.$$

**Remark 5.8**

As already explained in Remark 5.3, the results of [23] treat a somewhat different but more difficult situation. Moreover, the above result does not contain the full local well-posedness and a refined blow-up condition shown in [23]. So Theorem 5.7 follows from Theorem 5.2 by rather standard arguments, compare the proof of Theorem 6.1 of [23].

**Compatibility conditions**

For a smooth solution  $\mathbf{U} \in \mathcal{G}^3(\mathbb{R}^2 \times J)^3$  of (50) with  $J := (0, t_M)$  we can differentiate (50) two times in time and get new equations that are still satisfied for all  $t \in J$ . By continuity these new equations have to be satisfied at  $t = 0$ . This gives us necessary conditions on the initial values for  $\mathbf{U} \in \mathcal{G}^3(\mathbb{R}^2 \times J)^3$ .

If  $\tilde{S}(\mathbf{U})$  is positive definite, then  $\tilde{S}(\mathbf{U})$  is invertible and system (50) implies

$$\begin{aligned} \partial_t \mathbf{U} &= -\tilde{S}(\mathbf{U})^{-1} \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{U} =: \tilde{\mathbf{V}}^{(1)}(\mathbf{U}), \\ \llbracket U_2 \rrbracket &= \llbracket U_3 \rrbracket = 0. \end{aligned} \quad (53)$$

Differentiation in time gives us the following new equations:

$$\partial_t^2 \mathbf{U} = -\tilde{S}(\mathbf{U})^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \partial_t \mathbf{U} + \partial_t \tilde{S}(\mathbf{U}) \partial_t \mathbf{U} \right) =: \tilde{\mathbf{V}}^{(2)}(\mathbf{U}, \partial_t \mathbf{U}), \quad (54)$$

$$\llbracket \partial_t U_2 \rrbracket = \llbracket \partial_t U_3 \rrbracket = 0,$$

$$\partial_t^3 \mathbf{U} = -\tilde{S}(\mathbf{U})^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \partial_t^2 \mathbf{U} + 2\partial_t \tilde{S}(\mathbf{U}) \partial_t^2 \mathbf{U} + \partial_t^2 \tilde{S}(\mathbf{U}) \partial_t \mathbf{U} \right) =: \tilde{\mathbf{V}}^{(3)}(\mathbf{U}, \partial_t \mathbf{U}, \partial_t^2 \mathbf{U}), \quad (55)$$

$$\llbracket \partial_t^2 U_2 \rrbracket = \llbracket \partial_t^2 U_3 \rrbracket = 0.$$

We can now iteratively define  $\mathbf{V}^{(1)}(\mathbf{U}) := \tilde{\mathbf{V}}^{(1)}(\mathbf{U})$ ,  $\mathbf{V}^{(2)}(\mathbf{U}) := \tilde{\mathbf{V}}^{(2)}(\mathbf{U}, \mathbf{V}^{(1)}(\mathbf{U}))$  and  $\mathbf{V}^{(3)}(\mathbf{U}) := \tilde{\mathbf{V}}^{(3)}(\mathbf{U}, \mathbf{V}^{(1)}(\mathbf{U}), \mathbf{V}^{(2)}(\mathbf{U}))$  to get operators  $\mathbf{V}^{(j)}$  that only contain space derivatives and no time derivatives. For the sake of completeness we also define  $\mathbf{V}^{(0)}(\mathbf{U}) := \mathbf{U}$ . The equations above imply that

$$\partial_t^j \mathbf{U}(\cdot, 0) = \mathbf{V}^{(j)}(\mathbf{U}(\cdot, 0)) \quad \text{and} \quad \left[ \left[ V_2^{(j-1)}(\mathbf{U}(\cdot, 0)) \right] \right] = \left[ \left[ V_3^{(j-1)}(\mathbf{U}(\cdot, 0)) \right] \right] = 0$$

for  $j \in \{1, 2, 3\}$ . Hence, the initial values  $\mathbf{U}^{(0)}$  have to satisfy the necessary conditions

$$\left[ \left[ V_2^{(j)}(\mathbf{U}^{(0)}) \right] \right] = \left[ \left[ V_3^{(j)}(\mathbf{U}^{(0)}) \right] \right] = 0 \quad (56)$$

for  $j \in \{0, 1, 2\}$ . Note that for higher regularity additional compatibility conditions are necessary, but we will focus our analysis on solutions in  $\mathcal{G}^3(\mathbb{R}^2 \times J)^3$ .

**Definition 5.9** (Nonlinear Compatibility Conditions)

Let  $m \in \{1, 2, 3\}$ . We say that an initial value  $\mathbf{U}^{(0)} \in \mathcal{H}^m(\mathbb{R}^2)^3$  satisfies the nonlinear compatibility conditions of order  $m$  for (50) if (56) is true for  $j \in \{0, \dots, m-1\}$ .

**Remark 5.10**

The compatibility conditions for the linear problem (48) can be derived analogously. In comparison to (53), (54) and (55) we have to replace  $\tilde{S}(\mathbf{U})$  by  $A_t$  and include the additional terms  $M(\mathbf{x}, t)\mathbf{U}$ ,  $\mathbf{f}$  and their time derivatives:

$$\begin{aligned} \tilde{\mathbf{V}}_{\text{lin}}^{(1)}(\mathbf{U}) &= -A_t^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{U} + M\mathbf{U} - \mathbf{f} \right), \\ \tilde{\mathbf{V}}_{\text{lin}}^{(2)}(\mathbf{U}, \partial_t \mathbf{U}) &= -A_t^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \partial_t \mathbf{U} + \partial_t A_t \partial_t \mathbf{U} + \partial_t (M\mathbf{U}) - \partial_t \mathbf{f} \right), \\ \tilde{\mathbf{V}}_{\text{lin}}^{(3)}(\mathbf{U}, \partial_t \mathbf{U}, \partial_t^2 \mathbf{U}) &= -A_t^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \partial_t^2 \mathbf{U} + \partial_t^2 A_t \partial_t \mathbf{U} + 2\partial_t A_t \partial_t^2 \mathbf{U} + \partial_t^2 (M\mathbf{U}) - \partial_t^2 \mathbf{f} \right). \end{aligned}$$

## 6 Proof of Theorem 1.1.

Let  $a > 1$ . We start by expressing the equation for the error  $\varepsilon^a \mathbf{R} := \mathbf{U} - \mathbf{U}_{\text{ext}}$ . Substituting

$$\mathbf{U} = \mathbf{U}_{\text{ext}} + \varepsilon^a \mathbf{R} \quad (57)$$

in (50), one obtains

$$\left\{ \begin{array}{l} S(\mathbf{x}, t, \mathbf{R}^\pm) \partial_t \mathbf{R}^\pm + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{R}^\pm + W(\mathbf{x}, t, \mathbf{R}^\pm) \mathbf{R}^\pm = -\varepsilon^{-a} \mathbf{Res}, \quad \mathbf{x} \in \mathbb{R}_\pm^2, t \in J, \\ B_\Gamma \begin{pmatrix} \mathbf{R}^+ \\ \mathbf{R}^- \end{pmatrix} = \mathbf{0}, \quad \mathbf{x} \in \Gamma, t \in J, \\ \mathbf{R}(0) = \mathbf{R}^{(0)}, \quad \mathbf{x} \in \mathbb{R}^2, \end{array} \right. \quad (58)$$

where we set  $\mathbf{R}^{(0)} := \varepsilon^{-a}(\mathbf{U}^{(0)} - \mathbf{U}_{\text{ext}}(\cdot, 0))$  and, recalling (51),

$$\mathbf{Res} := \mathbf{Res}(\mathbf{U}_{\text{ext}}) = \tilde{S}(\cdot, \mathbf{U}_{\text{ext}}) \partial_t \mathbf{U}_{\text{ext}} + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{U}_{\text{ext}},$$

$$\begin{aligned}
S(\mathbf{x}, t, \mathbf{R}) &:= \tilde{S}(\mathbf{x}, \mathbf{U}_{\text{ext}}(\mathbf{x}, t) + \varepsilon^a \mathbf{R}) = \Lambda(\mathbf{x}) + \varepsilon_3(x_1) \varepsilon^{2a} \theta(\mathbf{R}) + \varphi(\mathbf{R}), \\
\varphi(\mathbf{R}) &:= \varepsilon_3(x_1) \varepsilon^a \begin{pmatrix} 6U_{\text{ext},1}R_1 + 2U_{\text{ext},2}R_2 & 2U_{\text{ext},1}R_2 + 2U_{\text{ext},2}R_1 & 0 \\ 2U_{\text{ext},1}R_2 + 2U_{\text{ext},2}R_1 & 2U_{\text{ext},1}R_1 + 6U_{\text{ext},2}R_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \varepsilon_3(x_1) \theta(\mathbf{U}_{\text{ext}}), \\
W(\mathbf{x}, t, \mathbf{R}) \mathbf{R} &:= \varepsilon^{-a} (S(\mathbf{x}, t, \mathbf{R}) - \tilde{S}(\mathbf{x}, \mathbf{U}_{\text{ext}}(\mathbf{x}, t))) \partial_t \mathbf{U}_{\text{ext}}(\mathbf{x}, t).
\end{aligned}$$

One can check that the entries of the  $3 \times 3$  matrix  $W$  are

$$\begin{aligned}
W_{1,1} &= \varepsilon_3 (\partial_t (3U_{\text{ext},1}^2 + U_{\text{ext},2}^2) + 3\varepsilon^a R_1 \partial_t U_{\text{ext},1} + 2\varepsilon^a R_2 \partial_t U_{\text{ext},2}), \\
W_{1,2} &= \varepsilon_3 (\varepsilon^a R_2 \partial_t U_{\text{ext},1} + 2\partial_t (U_{\text{ext},1} U_{\text{ext},2})), \quad W_{2,1} = \varepsilon_3 (\varepsilon^a R_1 \partial_t U_{\text{ext},2} + 2\partial_t (U_{\text{ext},1} U_{\text{ext},2})), \\
W_{2,2} &= \varepsilon_3 (\partial_t (U_{\text{ext},1}^2 + 3U_{\text{ext},2}^2) + 3\varepsilon^a R_2 \partial_t U_{\text{ext},2} + 2\varepsilon^a R_1 \partial_t U_{\text{ext},1}), \\
W_{1,3} &= W_{2,3} = W_{3,1} = W_{3,2} = W_{3,3} = 0.
\end{aligned}$$

The interface condition of (58) is a consequence of  $B_\Gamma(\mathbf{U}_{\text{ext}}^+, \mathbf{U}_{\text{ext}}^-)^\top = \mathbf{0}$ , as explained in Remark 4.1.

For a fixed  $\mathbf{U}_{\text{ext}}$ , systems (50) and (58) are, of course, equivalent provided (57) holds. Our rough strategy is to use the local existence Theorem 5.7 for (50) in order to get the existence of  $\mathbf{R}$  on the time interval  $(0, t_M)$  and then apply a bootstrapping argument on (58) to show that  $t_M = \mathcal{O}(\varepsilon^{-2})$  and that the desired estimate on  $\mathbf{R}$  holds for all small enough  $\varepsilon > 0$ .

For the application of Theorem 5.7 to (50) we need to find  $\Omega_\pm \subset \mathbb{R}^3$  such that  $\tilde{S} \in \mathcal{ML}_{\eta, \text{cv}}^{3,3}(\mathbb{R}^2, \Omega_\pm)$ . For this, firstly,  $\varepsilon_1^\pm, \varepsilon_3^\pm \in C^3(\mathbb{R}_\pm) \cap W^{3,\infty}(\mathbb{R}_\pm)$  is needed and  $\varepsilon_1, \varepsilon_3$  have to converge for  $|x_1| \rightarrow \infty$ , as we have assumed in (A1)–(A3). Secondly, the symmetric matrix  $\tilde{S}(\mathbf{x}, \mathbf{v})$  has to be positive definite for all  $\mathbf{x} \in \mathbb{R}_\pm^2$  and  $\mathbf{v} \in \Omega_\pm$ , respectively. It is easy to verify that  $\tilde{S}(\mathbf{v})$  has the three eigenvalues  $\lambda_1 = \mu_0$ ,  $\lambda_2 = \varepsilon_1 + \varepsilon_3 (v_1^2 + v_2^2)$ , and  $\lambda_3 = \varepsilon_1 + 3\varepsilon_3 (v_1^2 + v_2^2)$ . We now have to check when  $\lambda_{1,2,3} \geq \eta > 0$ .

Recall the bounds on  $\varepsilon_1$  and  $\varepsilon_3$  in (A1) and (A3). If  $\varepsilon_{3,m}^\pm \geq 0$ , then clearly  $\lambda_{2,3} > \varepsilon_{1,m}^\pm$  and the choice  $\eta := \min\{\mu_0, \varepsilon_{1,m}^+, \varepsilon_{1,m}^-\}$  and  $\Omega_\pm := \mathbb{R}^3$  is possible (and (52) trivially holds). If  $\varepsilon_{3,m}^\pm < 0$ , we impose

$$\varepsilon_{1,m}^\pm + 3\varepsilon_{3,m}^\pm (v_1^2 + v_2^2) > \eta > 0 \quad \text{for all } \mathbf{v} \in \Omega_\pm.$$

Choosing  $\eta \in (0, \min\{\mu_0, \varepsilon_{1,m}^+, \varepsilon_{1,m}^-\})$  and

$$\Omega_\pm := \begin{cases} \left\{ \mathbf{v} \in \mathbb{R}^3 \mid v_1^2 + v_2^2 < \frac{\eta - \varepsilon_{1,m}^\pm}{3\varepsilon_{3,m}^\pm} \right\}, & \varepsilon_{3,m}^\pm < 0, \\ \mathbb{R}^3, & \varepsilon_{3,m}^\pm \geq 0, \end{cases}$$

we infer  $\tilde{S}(\cdot, t, \cdot) \in \mathcal{ML}_{\eta, \text{cv}}^{3,3}(\mathbb{R}^2, \Omega_\pm)$  for all  $t \in [0, T_0 \varepsilon^{-2}]$ .

We now take a solution  $A \in \bigcap_{k=0}^4 C^{4-k}([0, T_0], H^{2+k}(\mathbb{R}))$  of the effective nonlinear Schrödinger equation (38) for some  $T_0 > 0$ . Choose  $\mathbf{R}^{(0)} \in \mathcal{H}^3(\mathbb{R}^2)^3$  and  $\varepsilon_* > 0$  small enough such that  $\mathbf{U}^{(0)} := \mathbf{U}_{\text{ext}}(\cdot, 0) + \varepsilon_*^a \mathbf{R}^{(0)}$  satisfies  $\overline{\text{im } \mathbf{U}^{(0), \pm}} \subset \Omega_\pm$  (which implies (52)) and the nonlinear compatibility conditions of order 3, see Definition 5.9. Then the local existence Theorem 5.7 yields a maximal existence time  $t_M > 0$  and a solution  $\mathbf{U} \in \mathcal{G}^3(\mathbb{R}^2 \times (0, t_M))^3$  of (50). For  $t \in [0, t_M)$  we set

$$z(t) := \sum_{k=0}^3 \|\partial_t^k \mathbf{R}(\cdot, t)\|_{\mathcal{H}^{3-k}(\mathbb{R}^2)^3}^2.$$

We have  $\|\mathbf{R}(\cdot, t)\|_{L^\infty(\mathbb{R}^2)^3} \leq c_S z(t)^{1/2}$  for a constant  $c_S \geq 1$ . Take  $\varpi > 0$  with

$$\varpi^2 < \frac{1}{c_S^2} \min \left\{ \frac{\eta - \varepsilon_{1,m}^-}{3 \min\{\varepsilon_{3,m}^-, 0\}}, \frac{\eta - \varepsilon_{1,m}^+}{3 \min\{\varepsilon_{3,m}^+, 0\}} \right\}, \quad (59)$$

where  $\frac{\varepsilon}{\rho} := +\infty$ . The major part of the rest of the proof of Theorem 1.1 is a bootstrapping argument to prove the statement

$$\left\{ \begin{array}{l} \exists 1 \geq \rho > \rho_0 > 0 \exists \varepsilon_0 = \varepsilon_0(\rho) \in (0, \varepsilon_*) \exists t^* \in (0, t_M) \text{ such that for all } \varepsilon \in (0, \varepsilon_0) \text{ we have} \\ \varepsilon^a \rho + \|\mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times [0, t^*])} \leq \varpi \text{ and if } z(0) \leq \rho_0^2 \text{ and } t^* \leq T_0 \varepsilon^{-2}, \text{ then } z(t) \leq \rho^2 \text{ for all } t \in [0, t^*]. \end{array} \right. \quad (60)$$

Note that the first condition in (60), i.e., the smallness of  $\varepsilon^a \rho + \|\mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times [0, T_0 \varepsilon^{-2}])^3}$ , can be achieved by simply choosing  $\varepsilon_0 = \varepsilon_0(\rho)$  small enough.

To establish (60), we define for  $1 \geq \rho > \rho_0 > 0$

$$\begin{aligned} T_{\rho_0, \varepsilon_0} &:= \sup \left\{ t^* \geq 0 \mid \varepsilon^a \rho + \|\mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times [0, t^*])} \leq \varpi, \forall t \in [0, t^*] : z(t) \leq \rho^2, z(0) = \rho_0^2 \right\}, \quad (61) \\ J_{\rho_0, \varepsilon_0} &:= [0, T_{\rho_0, \varepsilon_0}). \end{aligned}$$

On the time interval  $J_{\rho_0, \varepsilon_0}$  the conditions

$$\forall t \in J_{\rho_0, \varepsilon_0} : \quad \text{dist} \left( \overline{\text{im } \mathbf{U}^\pm(\cdot, t)}, \partial\Omega_\pm \right) > \kappa > 0, \quad \|\mathbf{U}(\cdot, t)\|_{\mathcal{H}^3(\mathbb{R}^2)^3} \leq C < \infty$$

are guaranteed.

We will prove in Section 6.1 that  $z(t) \leq \rho^2/2$  for  $t \in J_{\rho_0, \varepsilon_0}$  and suitable  $\varepsilon_0$  and  $\rho_0$  and hence  $T_{\rho_0, \varepsilon_0} = T_0 \varepsilon^{-2}$  and (60) is true. This yields the estimate

$$\|\mathbf{U} - \mathbf{U}_{\text{ext}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times (0, T_0 \varepsilon^{-2}))^3} \leq \rho \varepsilon^a. \quad (62)$$

Finally, to obtain (14), it will only remain to show  $\|\mathbf{U}_{\text{ans}} - \mathbf{U}_{\text{ext}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times (0, T_0 \varepsilon^{-2}))^3} \leq C \varepsilon^a$ , which is straightforward, see Section 6.2.

## 6.1 Bootstrapping argument (proof of (60))

We use the multi-index  $\beta := (\beta_1, \beta_2, \beta_t)^\top \in \mathbb{N}_0^3$ ,  $|\beta| \leq 3$ , and the abbreviation  $\mathbf{r}_\beta := \partial^\beta \mathbf{R}$ . Applying  $\partial^\beta = \partial_{x_1}^{\beta_1} \partial_{x_2}^{\beta_2} \partial_t^{\beta_t}$  to (58) yields

$$\left\{ \begin{array}{l} S(\mathbf{x}, t, \mathbf{R}) \partial_t \mathbf{r}_\beta + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{r}_\beta \\ \quad = \mathbf{s}_\beta(\mathbf{x}, t, \mathbf{R}) + \mathbf{w}_\beta(\mathbf{x}, t, \mathbf{R}) - \varepsilon^{-a} \partial^\beta \mathbf{Res}(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^2 \setminus \Gamma, t \in J_{\rho_0, \varepsilon_0}, \\ \mathbf{r}_\beta(\mathbf{x}, 0) = \mathbf{r}_\beta^{(0)}(\mathbf{x}) := \partial^\beta \mathbf{R}(\mathbf{x}, 0), \quad \mathbf{x} \in \mathbb{R}^2 \end{array} \right. \quad (63)$$

with

$$\begin{aligned} \mathbf{s}_\beta(\mathbf{x}, t, \mathbf{R}) &:= - \sum_{\mathbf{0} < \gamma \leq \beta} \binom{\beta}{\gamma} \partial^\gamma S(\mathbf{x}, t, \mathbf{R}) \partial^{\beta-\gamma} \partial_t \mathbf{R}, \\ \mathbf{w}_\beta(\mathbf{x}, t, \mathbf{R}) &:= -\partial^\beta (W(\mathbf{x}, t, \mathbf{R}) \mathbf{R}). \end{aligned}$$

Note that the interface conditions cannot be simply differentiated for all  $\beta_1 \neq 0$ , therefore we will discuss them separately. The time derivatives  $\partial_t^k \mathbf{R}(\cdot, 0)$  have to be interpreted as right-sided derivatives that satisfy

$$\partial_t^j \mathbf{R}(\cdot, 0) = \mathbf{V}^{(j)}(\mathbf{R}(\cdot, 0)) = \mathbf{V}^{(j)}(\mathbf{R}^{(0)}), \quad (64)$$

with  $\mathbf{V}^{(j)}$  as defined in Section 5. Testing (63) with  $\mathbf{r}_\beta$  produces

$$\begin{aligned} &\int_0^t \int_{\mathbb{R}^2} \left( S(\mathbf{R}) \partial_t \mathbf{r}_\beta \cdot \mathbf{r}_\beta + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{r}_\beta \cdot \mathbf{r}_\beta \right) \mathrm{d}\mathbf{x} \mathrm{d}s \\ &= \int_0^t \int_{\mathbb{R}^2} (\mathbf{w}_\beta(\mathbf{R}) \cdot \mathbf{r}_\beta + \mathbf{s}_\beta(\mathbf{R}) \cdot \mathbf{r}_\beta) \mathrm{d}\mathbf{x} \mathrm{d}s - \varepsilon^{-a} \int_0^t \int_{\mathbb{R}^2} \partial^\beta \mathbf{Res} \cdot \mathbf{r}_\beta \mathrm{d}\mathbf{x} \mathrm{d}s. \end{aligned} \quad (65)$$

The main steps of our bootstrapping argument are:

- I. Employ (64) to estimate  $\left\| \mathbf{r}_\beta^{(0)} \right\|_{L^2(\mathbb{R}^2)^3}$  for all  $\beta \in \mathbb{N}_0^3$  with  $|\beta| \leq 3$ .
- II. Based on (65), estimate  $\sum_{|\gamma| \leq 3, \gamma_1=0} \|\partial^\gamma \mathbf{R}(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2$  using that  $\int_{\mathbb{R}^2} \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} = 0$  if  $\beta_1 = 0$ .
- III. Rewrite (63) to analyze  $\partial^\beta R_2$  and  $\partial^\beta R_3$  for  $\beta_1 = 1$ , and then iterate the process for  $\beta_1 = 2$  and  $\beta_1 = 3$ .
- IV. Use  $\nabla \cdot \partial_t \mathcal{D}(\mathbf{U}) = 0$  to estimate  $\partial^\beta R_1$  for  $\beta_1 = 1$ , where we start with  $\beta = (1, 0, 0)^\top$  and then iterate to increase  $\beta_t$  and  $\beta_2$ . Finally we have to iterate the process again for  $\beta_1 = 2$  and  $\beta_1 = 3$ .

Our basic strategy follows the proof of the local a priori estimates in [23]. The main difference is that, using the structure of our ansatz, we can derive the estimates on a large time interval  $[0, T_0 \varepsilon^{-2})$  with the desired dependence on  $\varepsilon$ . We let  $t \in J_{\rho_0, \varepsilon_0}$ .

### Step I: Estimates of the initial values

In this section we estimate  $\|\mathbf{r}_\beta(\cdot, 0)\|_{L^2(\mathbb{R}^2)^3}$  for all  $\beta \in \mathbb{N}_0^3$  with  $|\beta| \leq 3$ . For  $\beta_t = 0$  we already have by assumption that

$$\|\mathbf{R}(\cdot, 0)\|_{\mathcal{H}^3(\mathbb{R}^2)^3} = \|\mathbf{R}^{(0)}\|_{\mathcal{H}^3(\mathbb{R}^2)^3} < \rho_0.$$

If  $\beta_t \neq 0$  we will use (64) to estimate  $\partial_t^p \mathbf{R}(\cdot, 0)$  in  $\mathcal{H}^{3-p}(\mathbb{R}^2)^3$  for  $p \in \{1, 2, 3\}$ .

Since  $\mathbf{U}^{(0)}$  satisfies the nonlinear compatibility conditions of order 3 we know from Section 5 that (suppressing the  $\mathbf{x}$ -dependence)

$$\partial_t^j \mathbf{U}(t) = \mathbf{V}^{(j)}(\mathbf{U}(t))$$

for all  $t \in [0, t_M)$  and  $j \in \{0, 1, 2\}$ . With  $\mathbf{U} = \varepsilon^a \mathbf{R} + \mathbf{U}_{\text{ext}}$ , as in (58) we rewrite these three equations as

$$\begin{aligned} \partial_t \mathbf{R} &= -\tilde{S}(\mathbf{U})^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{R} + \varepsilon^{-a} \mathbf{Res} + \varepsilon^{-a} \left( \tilde{S}(\mathbf{U}) - \tilde{S}(\mathbf{U}_{\text{ext}}) \right) \partial_t \mathbf{U}_{\text{ext}} \right), \\ \partial_t^2 \mathbf{R} &= -\tilde{S}(\mathbf{U})^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \partial_t \mathbf{R} + \partial_t \tilde{S}(\mathbf{U}) \partial_t \mathbf{R} + \varepsilon^{-a} \partial_t \mathbf{Res} \right) \\ &\quad - \varepsilon^{-a} \tilde{S}(\mathbf{U})^{-1} \left( \partial_t \left( \tilde{S}(\mathbf{U}) - \tilde{S}(\mathbf{U}_{\text{ext}}) \right) \partial_t \mathbf{U}_{\text{ext}} + \left( \tilde{S}(\mathbf{U}) - \tilde{S}(\mathbf{U}_{\text{ext}}) \right) \partial_t^2 \mathbf{U}_{\text{ext}} \right), \\ \partial_t^3 \mathbf{R} &= -\tilde{S}(\mathbf{U})^{-1} \left( \sum_{j=1}^2 A_j \partial_{x_j} \partial_t^2 \mathbf{R} + 2 \partial_t \left( \tilde{S}(\mathbf{U}) \right) \partial_t^2 \mathbf{R} + \partial_t^2 \left( \tilde{S}(\mathbf{U}) \right) \partial_t \mathbf{R} + \varepsilon^{-a} \partial_t^2 \mathbf{Res} \right) \\ &\quad - \varepsilon^{-a} \tilde{S}(\mathbf{U})^{-1} \left( \partial_t^2 \left( \tilde{S}(\mathbf{U}) - \tilde{S}(\mathbf{U}_{\text{ext}}) \right) \partial_t \mathbf{U}_{\text{ext}} + 2 \partial_t \left( \tilde{S}(\mathbf{U}) - \tilde{S}(\mathbf{U}_{\text{ext}}) \right) \partial_t^2 \mathbf{U}_{\text{ext}} \right. \\ &\quad \left. + \left( \tilde{S}(\mathbf{U}) - \tilde{S}(\mathbf{U}_{\text{ext}}) \right) \partial_t^3 \mathbf{U}_{\text{ext}} \right). \end{aligned} \tag{66}$$

The following lemma collects some properties of the matrix function  $\tilde{S}$ . It can be shown as Lemma 2.23, Lemma 7.1 and Corollary 7.2 in [27].

#### Lemma 6.1

Let  $T', \eta_0, R > 0, \Omega_\pm \subset \mathbb{R}^3$  and  $\tilde{S} \in \mathcal{ML}_{\eta_0, \text{cv}}^{3,3}(\mathbb{R}^2, \Omega_\pm)$ . Then for all  $\mathbf{U}, \mathbf{V} \in B_R(0) \subset \mathcal{G}^3(\mathbb{R}^2 \times [0, T'])^3$  with  $\text{im } \mathbf{U}^\pm, \text{im } \mathbf{V}^\pm \subset \Omega_\pm$  there exists  $C > 0$  such that

$$i) \left\| \tilde{S}(\mathbf{U}(t))^{-1} \right\|_{\mathcal{W}^{2,\infty}(\mathbb{R}^2)^{3 \times 3} + \mathcal{H}^2(\mathbb{R}^2)^{3 \times 3}} \leq C,$$



$$ii) \left\| \partial_t^k \tilde{S}(\mathbf{U}(t)) \right\|_{\mathcal{W}^{3-k, \infty}(\mathbb{R}^2)^{3 \times 3} + \mathcal{H}^{3-k}(\mathbb{R}^2)^{3 \times 3}} \leq C,$$

$$iii) \left\| \partial_t^k \left( \tilde{S}(\mathbf{U}(t)) - \tilde{S}(\mathbf{V}(t)) \right) \right\|_{\mathcal{H}^{2-k}(\mathbb{R}^2)^{3 \times 3}} \leq C \sum_{j=0}^k \left\| \partial_t^j \mathbf{U}(t) - \partial_t^j \mathbf{V}(t) \right\|_{\mathcal{H}^{2-k}(\mathbb{R}^2)^3}$$

for all  $k \in \{0, 1, 2\}$  and  $t \in [0, T']$ .

We can now go back to equations (66) and use Lemma C.1, Lemma 6.1 and  $\varepsilon^a \mathbf{R} = \mathbf{U} - \mathbf{U}_{\text{ext}}$  to show

$$\begin{aligned} & \left\| \partial_t \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} \\ & \leq C \left\| \tilde{S}(\mathbf{U}(0))^{-1} \right\|_{\mathcal{W}^{2, \infty}(\mathbb{R}^2)^{3 \times 3} + \mathcal{H}^2(\mathbb{R}^2)^{3 \times 3}} \left( \left\| \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3} + \varepsilon^{-a} \left\| \mathbf{Res}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} \right) \\ & \quad + C \varepsilon^{-a} \left\| \varepsilon^a \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} \left\| \partial_t \mathbf{U}_{\text{ext}}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} \\ & \leq C \left( \left\| \mathbf{R}^{(0)} \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3} + \varepsilon^{-a} \left\| \mathbf{Res}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} \right). \end{aligned}$$

The remaining two estimates follow analogously:

$$\begin{aligned} \left\| \partial_t^2 \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^1(\mathbb{R}^2)^3} & \leq C \left( \left\| \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3} + \left\| \partial_t \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} + \varepsilon^{-a} \left\| \partial_t \mathbf{Res}(\cdot, 0) \right\|_{\mathcal{H}^1(\mathbb{R}^2)^3} \right), \\ \left\| \partial_t^3 \mathbf{R}(\cdot, 0) \right\|_{L^2(\mathbb{R}^2)^3} & \leq C \left( \left\| \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3} + \left\| \partial_t \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^2(\mathbb{R}^2)^3} + \left\| \partial_t^2 \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^1(\mathbb{R}^2)^3} \right. \\ & \quad \left. + \varepsilon^{-a} \left\| \partial_t^2 \mathbf{Res}(\cdot, 0) \right\|_{L^2(\mathbb{R}^2)^3} \right). \end{aligned}$$

Finally we use the recursive structure of the estimates and obtain

$$\left\| \partial_t^p \mathbf{R}(\cdot, 0) \right\|_{\mathcal{H}^{3-p}(\mathbb{R}^2)^3} \leq C \left( \left\| \mathbf{R}^{(0)} \right\|_{\mathcal{H}^3(\mathbb{R}^2)^3} + \varepsilon^{-a} \sum_{j=0}^{p-1} \left\| \partial_t^j \mathbf{Res}(\cdot, 0) \right\|_{\mathcal{H}^{2-j}(\mathbb{R}^2)^3} \right)$$

for all  $p \in \{1, 2, 3\}$ . With our estimate for the residual, see (44), we infer

$$\left\| \mathbf{r}_\beta^{(0)} \right\|_{L^2(\mathbb{R}^2)^3} \leq C \left( \rho_0 + \varepsilon^{\frac{7}{2}-a} \right) \quad (67)$$

for all  $\beta \in \mathbb{N}_0^3$  with  $|\beta| \leq 3$ .

## Step II: Estimates of the $\beta$ -derivatives of $\mathbf{R}(\cdot, t)$ with $\beta_1 = 0, |\beta| \leq 3$

We first show an energy estimate for the  $t$ - and  $x_2$ -derivatives of  $\mathbf{R}$ .

### Lemma 6.2

Let  $\mathbf{R} \in \mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  be a solution of (58) and let  $\beta \in \mathbb{N}_0^3, |\beta| \leq 3, \beta_1 = 0$ . Then  $\mathbf{r}_\beta = \partial^\beta \mathbf{R}$  satisfies

$$\begin{aligned} \frac{\eta}{2} \left\| \mathbf{r}_\beta(\cdot, t) \right\|_{L^2(\mathbb{R}^2)^3}^2 & \leq C \left\| \mathbf{r}_\beta^{(0)} \right\|_{L^2(\mathbb{R}^2)^3}^2 + \int_0^t \int_{\mathbb{R}^2} \left( \mathbf{w}_\beta(\mathbf{R}) \cdot \mathbf{r}_\beta + \mathbf{s}_\beta(\mathbf{R}) \cdot \mathbf{r}_\beta \right. \\ & \quad \left. + \frac{1}{2} \partial_t S(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta - \varepsilon^{-a} \partial^\beta \mathbf{Res} \cdot \mathbf{r}_\beta \right) dx ds \end{aligned} \quad (68)$$

for every  $t \in J_{\rho_0, \varepsilon_0}$ .

PROOF: Step 1. Let us first study the case  $|\beta| < 3$ . Since  $\mathbf{R} \in \mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$ , we have  $\mathbf{r}_\beta = \partial^\beta \mathbf{R} \in \mathcal{G}^1(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$ . To employ (65), we compute

$$\int_0^t \int_{\mathbb{R}^2} S(\mathbf{R}) \partial_t \mathbf{r}_\beta \cdot \mathbf{r}_\beta dx ds = \frac{1}{2} \int_0^t \partial_t \left( \int_{\mathbb{R}^2} S(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta dx \right) ds - \frac{1}{2} \int_0^t \int_{\mathbb{R}^2} \partial_t S(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta dx ds.$$

Using that  $S(\mathbf{R})$  is positive definite, we estimate  $\int_{\mathbb{R}^2} S(\mathbf{R})(t) \mathbf{r}_\beta(t) \cdot \mathbf{r}_\beta(t) \, d\mathbf{x} \geq \eta \|\mathbf{r}_\beta(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2$ . Moreover, we have  $\int_{\mathbb{R}^2} S(\mathbf{R})(0) \mathbf{r}_\beta(0) \cdot \mathbf{r}_\beta(0) \, d\mathbf{x} \leq \|S(\mathbf{R}^{(0)})\|_{L^\infty(\mathbb{R}^2)^{3 \times 3}} \|\mathbf{r}_\beta^{(0)}\|_{L^2(\mathbb{R}^2)^3}^2$ . Since  $\mathbf{R}^{(0)}, \mathbf{U}_{\text{ext}}^{(0)} \in L^\infty(\mathbb{R}^2)^3$ , this leads to

$$\frac{\eta}{2} \|\mathbf{r}_\beta(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2 \leq C \|\mathbf{r}_\beta^{(0)}\|_{L^2(\mathbb{R}^2)^3}^2 + \int_0^t \int_{\mathbb{R}^2} S(\mathbf{R}) \partial_t \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds + \frac{1}{2} \int_0^t \int_{\mathbb{R}^2} \partial_t S(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds.$$

An integration by parts yields

$$\int_{\mathbb{R}^2} \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} = \int_{\mathbb{R}^2} (-\partial_{x_2} r_{\beta,3} r_{\beta,1} + \partial_{x_1} r_{\beta,3} r_{\beta,2} - \partial_{x_2} r_{\beta,1} r_{\beta,3} + \partial_{x_1} r_{\beta,2} r_{\beta,3}) \, d\mathbf{x} = 0,$$

employing differentiated interface conditions

$$[[r_{\beta,2}]] = [[r_{\beta,3}]] = 0$$

in the  $x_1$ -integral. The interface conditions can be differentiated since  $\beta_1 = 0$ . Now (68) is a consequence of (65) and the above formulas.

Step 2. Next, we consider the remaining case  $|\beta| = 3, \beta_1 = 0$ . Let  $\mathbf{f} := \mathbf{s}_\beta(\mathbf{R}) + \mathbf{w}_\beta(\mathbf{R}) - \varepsilon^{-a} \partial^\beta \mathbf{R} \mathbf{e}_s$ . The differential equation in (63) becomes

$$S(\mathbf{R}) \partial_t \mathbf{r}_\beta + \sum_{j=1}^2 A_j \partial_{x_j} \mathbf{r}_\beta = \mathbf{f}, \quad \mathbf{x} \in \mathbb{R}^2 \setminus \Gamma, \, t \in J_{\rho_0, \varepsilon_0}, \quad (69)$$

Since  $S(\mathbf{R}) \in \mathcal{F}_{\eta, \text{cv}}^{3,3}(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})$  and  $\mathbf{f} \in \mathcal{G}^0(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  (which can easily be shown), we can apply Lemma 5.4 to (69) (setting  $A_t := S(\mathbf{R})$  and  $M := 0$ ). Because  $\mathbf{r}_\beta$  is a weak solution of (69) with the initial and interface conditions from (63), the lemma provides sequences  $(\mathbf{r}_{\beta,n}^{(0)})_n \subset \mathcal{D}(\mathbb{R}^2)^3$ ,  $(\mathbf{f}_n)_n \subset \mathcal{H}^1(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  and  $(\mathbf{r}_{\beta,n})_n \subset \mathcal{G}^1(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  with  $\mathbf{r}_{\beta,n}^{(0)} \rightarrow \mathbf{r}_\beta^{(0)}$  in  $L^2(\mathbb{R}^2)^3$ ,  $\mathbf{f}_n \rightarrow \mathbf{f}$  in  $L^2(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  and  $\mathbf{r}_{\beta,n} \rightarrow \mathbf{r}_\beta$  in  $\mathcal{G}^0(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  for  $n \rightarrow \infty$ , and  $\mathbf{r}_{\beta,n}$  is a weak solution of (69) with data  $(\mathbf{f}_n, \mathbf{r}_{\beta,n}^{(0)})$  for all  $n \in \mathbb{N}$ . The same calculation as in Step 1 shows that

$$\frac{\eta}{2} \|\mathbf{r}_{\beta,n}(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2 \leq C \|\mathbf{r}_{\beta,n}^{(0)}\|_{L^2(\mathbb{R}^2)^3}^2 + \int_0^t \int_{\mathbb{R}^2} \left( \mathbf{f}_n \cdot \mathbf{r}_{\beta,n} + \frac{1}{2} \partial_t S(\mathbf{R}) \mathbf{r}_{\beta,n} \cdot \mathbf{r}_{\beta,n} \, d\mathbf{x} \right) \, ds.$$

The Cauchy-Schwarz inequality then implies the statement.  $\square$

We now have to estimate each part of the right-hand side in (68). Let

$$\tilde{z}(t) := \sum_{\substack{\boldsymbol{\gamma} \in \mathbb{N}_0^3 \\ |\boldsymbol{\gamma}| \leq 3, \gamma_1 = 0}} \|\partial^\boldsymbol{\gamma} \mathbf{R}(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2.$$

From Step I we know that

$$\|\mathbf{r}_\beta^{(0)}\|_{L^2(\mathbb{R}^2)^3} \leq C \left( \rho + \varepsilon^{\frac{7}{2}-a} \right).$$

For  $\int_0^t \int_{\mathbb{R}^2} \partial_t S(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds$  we first have

$$\varepsilon^{2a} \varepsilon_3 \int_0^t \int_{\mathbb{R}^2} \frac{1}{2} \partial_t \theta(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds \leq C \varepsilon^{2a} \int_0^t \|\mathbf{r}_\beta(\cdot, s)\|_{L^2(\mathbb{R}^2)^3}^2 \, ds \leq C \varepsilon^{2a} \int_0^t \tilde{z}(s) \, ds$$

since  $\mathbf{R}, \partial_t \mathbf{R} \in L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$ . Similarly, using also that

$$\|\mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3}, \|\partial_t \mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \leq C \varepsilon,$$

cf. (45), we derive

$$\int_0^t \int_{\mathbb{R}^2} \frac{1}{2} \partial_t \varphi(\mathbf{R}) \mathbf{r}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds \leq C (\varepsilon^2 + \varepsilon^{1+a}) \int_0^t \|\mathbf{r}_\beta(\cdot, s)\|_{L^2(\mathbb{R}^2)^3}^2 \, ds \leq C \varepsilon^2 \int_0^t \tilde{z}(s) \, ds.$$

For the residual term, (44) yields  $\|\partial^\beta \mathbf{Res}(\cdot, t)\|_{L^2(\mathbb{R}^2)^3} \leq C \varepsilon^{7/2}$ . The Cauchy-Schwarz inequality and  $\|\mathbf{r}_\beta(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2 \leq z(t) \leq \rho \leq 1$  for  $t \in J_{\rho_0, \varepsilon_0}$  then give us

$$\int_0^t \int_{\mathbb{R}^2} \varepsilon^{-a} \partial^\beta \mathbf{Res} \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds \leq \int_0^t \varepsilon^{-a} \|\partial^\beta \mathbf{Res}(\cdot, s)\|_{L^2(\mathbb{R}^2)^3} \|\mathbf{r}_\beta(\cdot, s)\|_{L^2(\mathbb{R}^2)^3} \, ds \leq C \varepsilon^{\frac{3}{2}-a}.$$

The remaining terms  $\int_0^t \int_{\mathbb{R}^2} (\mathbf{w}_\beta \cdot \mathbf{r}_\beta + \mathbf{s}_\beta \cdot \mathbf{r}_\beta) \, d\mathbf{x} \, ds$  mainly consist of integrals of the type

$$I_1 := \int_0^t \int_{\mathbb{R}^2} \partial^{\mathbf{a}} f(\mathbf{x}, s) \partial^{\mathbf{b}} g(\mathbf{x}, s) \partial^{\mathbf{c}} h(\mathbf{x}, s) k(\mathbf{x}, s) \, d\mathbf{x} \, ds, \quad (70)$$

where  $f, g, h \in \mathcal{G}^3(\mathbb{R}^2 \times J)$ ,  $k \in \mathcal{G}^0(\mathbb{R}^2 \times J)$  and  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{N}_0^3$  with  $|\mathbf{a}|, |\mathbf{b}|, |\mathbf{c}| < 4$  and  $s := |\mathbf{a}| + |\mathbf{b}| + |\mathbf{c}| \leq 4$ . For  $s = 4$  we only have integrals where at least one time-derivative is present, i.e.,  $a_t = b_t = c_t = 0$  is not possible. The case where four derivatives fall on a component of  $\mathbf{U}_{\text{ext}}$  also appears and will be discussed separately. The following two classes of estimates are needed.

- i)  $|\mathbf{a}| \leq 3, |\mathbf{b}| \leq 1, |\mathbf{c}| \leq 1$ : Here  $\partial^{\mathbf{a}} f(\cdot, t), k(\cdot, t) \in L^2(\mathbb{R}^2)$  and  $\partial^{\mathbf{b}} g(\cdot, t), \partial^{\mathbf{c}} h(\cdot, t) \in L^\infty(\mathbb{R}^2)$ . With the Cauchy-Schwarz inequality we obtain

$$I_1 \leq C \int_0^t \|\partial^{\mathbf{a}} f(\cdot, s)\|_{L^2(\mathbb{R}^2)} \|\partial^{\mathbf{b}} g(\cdot, s)\|_{L^\infty(\mathbb{R}^2)} \|\partial^{\mathbf{c}} h(\cdot, s)\|_{L^\infty(\mathbb{R}^2)} \|k(\cdot, s)\|_{L^2(\mathbb{R}^2)} \, ds.$$

- ii)  $|\mathbf{a}| \leq 2, |\mathbf{b}| \leq 2, |\mathbf{c}| = 0$ : Now  $k(\cdot, t) \in L^2(\mathbb{R}^2)$ ,  $\partial^{\mathbf{c}} h(\cdot, t) \in L^\infty(\mathbb{R}^2)$ , and  $\partial^{\mathbf{a}} f(\cdot, t), \partial^{\mathbf{b}} g(\cdot, t) \in L^p(\mathbb{R}^2)$  for all  $p \in [1, \infty)$ . This follows from the Sobolev embedding  $H^1(\mathbb{R}_\pm^2) \hookrightarrow L^p(\mathbb{R}_\pm^2)$  for all  $1 \leq p < \infty$ . Hölder's inequality then yields

$$I_1 \leq C \int_0^t \|\partial^{\mathbf{a}} f(\cdot, s)\|_{L^3(\mathbb{R}^2)} \|\partial^{\mathbf{b}} g(\cdot, s)\|_{L^6(\mathbb{R}^2)} \|\partial^{\mathbf{c}} h(\cdot, s)\|_{L^\infty(\mathbb{R}^2)} \|k(\cdot, s)\|_{L^2(\mathbb{R}^2)} \, ds.$$

The role of the function  $k$  in (70) will always be played by a component of  $\mathbf{r}_\beta = \partial^\beta \mathbf{R}$ .

Recall that  $S(\mathbf{R}) = \Lambda + \varepsilon_3 \varepsilon^{2a} \theta(\mathbf{R}) + \varphi(\mathbf{R})$ . Hence, to estimate  $\int_0^t \int_{\mathbb{R}^2} \mathbf{s}_\beta \cdot \mathbf{r}_\beta \, d\mathbf{x} \, ds$ , we first analyze  $\varepsilon^{2a} \varepsilon_3 \partial^\gamma \theta(\mathbf{R}) \partial^{\beta-\gamma} \partial_t \mathbf{R} \cdot \mathbf{r}_\beta$  where  $\beta_1 = \gamma_1 = 0$ . This sum consists of terms of the form

$$C \varepsilon^{2a} \partial^{\gamma'} R_i \partial^{\gamma''} R_j \partial^{\beta-\gamma} \partial_t R_k \partial^\beta R_l$$

with  $\gamma = \gamma' + \gamma''$ ,  $\gamma'_1 = \gamma''_1 = 0$ , and  $i, j, k, l \in \{1, 2\}$ . The case i) above applies if  $|\beta - \gamma| = 0$ , where we may take  $|\gamma''| \leq 1$ . We then estimate

$$\begin{aligned} C \varepsilon^{2a} \int_0^t \int_{\mathbb{R}^2} \partial^{\gamma'} R_i \partial^{\gamma''} R_j \partial^{\beta-\gamma} \partial_t R_k \partial^\beta R_l \, d\mathbf{x} \, ds &\leq C \varepsilon^{2a} \int_0^t \|\partial^\beta R_l(\cdot, s)\|_{L^2(\mathbb{R}^2)} \|\partial^{\gamma'} R_i(\cdot, s)\|_{L^2(\mathbb{R}^2)} \, ds \\ &\leq C \varepsilon^{2a} \int_0^t \tilde{z}(s) \, ds. \end{aligned}$$

A representative of type ii) is any term with  $|\beta - \gamma| = 1, |\gamma'| = 2$  and  $|\gamma''| = 0$ , which is estimated via

$$C \varepsilon^{2a} \int_0^t \int_{\mathbb{R}^2} \partial^{\gamma'} R_i \partial^{\gamma''} R_j \partial^{\beta-\gamma} \partial_t R_k \partial^\beta R_l \, d\mathbf{x} \, ds$$

$$\begin{aligned}
&\leq C\varepsilon^{2a} \int_0^t \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left\| \partial^{\gamma'} R_i(\cdot, s) \right\|_{L^6(\mathbb{R}^2)} \left\| \partial^{\beta-\gamma} \partial_t R_k(\cdot, s) \right\|_{L^3(\mathbb{R}^2)} ds \\
&\leq C\varepsilon^{2a} \int_0^t \left( \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)}^2 + \left\| \partial^{\gamma'} R_i(\cdot, s) \right\|_{L^6(\mathbb{R}^2)}^2 \left\| \partial^{\beta-\gamma} \partial_t R_k(\cdot, s) \right\|_{L^3(\mathbb{R}^2)}^2 \right) ds \\
&\leq C\varepsilon^{2a} \int_0^t \left( \tilde{z}(s) + (z(s))^2 \right) ds \leq C\varepsilon^{2a} \int_0^t \tilde{z}(s) ds + C\rho^2 \varepsilon^{2a} t,
\end{aligned}$$

using again  $H^1(\mathbb{R}_\pm^2) \hookrightarrow L^p(\mathbb{R}_\pm^2)$  for  $1 \leq p < \infty$ . The remaining cases can be treated similarly.

Next, we study  $\partial^\gamma \varphi(\mathbf{R}) \partial^{\beta-\gamma} \partial_t \mathbf{R} \cdot \partial^\beta \mathbf{R}$  with  $|\beta - \gamma| \leq 2$  and  $\beta_1 = \gamma_1 = 0$ . We again use (45) which provides the inequality  $\|\partial^\alpha \mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \leq C\varepsilon$  for all  $|\alpha| \leq 3$  with  $\alpha_1 \leq 2$ . For terms quadratic in  $\mathbf{U}_{\text{ext}}$  case i) applies:

$$\begin{aligned}
\int_0^t \int_{\mathbb{R}^2} \partial^\gamma (U_{\text{ext},i} U_{\text{ext},j}) \partial^{\beta-\gamma} \partial_t R_k \partial^\beta R_l d\mathbf{x} ds &\leq C\varepsilon^2 \int_0^t \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left\| \partial^{\beta-\gamma} \partial_t R_k(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} ds \\
&\leq C\varepsilon^2 \int_0^t \tilde{z}(s) ds.
\end{aligned}$$

For terms linear in  $\mathbf{U}_{\text{ext}}$ , i.e.,  $I_2 := \varepsilon^a \int_0^t \int_{\mathbb{R}^2} \partial^\gamma U_{\text{ext},i} \partial^{\gamma''} R_j \partial^{\beta-\gamma} \partial_t R_k \partial^\beta R_l d\mathbf{x} ds$ , we distinguish the three cases  $|\beta - \gamma| = 0, 1, \text{ and } 2$ . For  $|\beta - \gamma| = 0$  we compute

$$I_2 \leq C\varepsilon^{1+a} \int_0^t \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left\| \partial^{\gamma''} R_j(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} ds \leq C\varepsilon^{2a} \int_0^t \tilde{z}(s) ds$$

by means of the estimate type i) and the fact that  $\partial_t R_k \in L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})$ . For  $|\beta - \gamma| = 1$  the estimate of type ii) applies and we have

$$\begin{aligned}
I_2 &\leq C\varepsilon^{1+a} \int_0^t \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left\| \partial^{\gamma''} R_j(\cdot, s) \right\|_{L^6(\mathbb{R}^2)} \left\| \partial^{\beta-\gamma} \partial_t R_k(\cdot, s) \right\|_{L^3(\mathbb{R}^2)} ds \\
&\leq C\varepsilon^{1+a} \int_0^t \tilde{z}(s) ds + C\varepsilon^{1+a} t
\end{aligned}$$

as  $\partial^{\gamma''} R_j, \partial^{\beta-\gamma} \partial_t R_k \in L^\infty(J_{\rho_0, \varepsilon_0}, \mathcal{H}^1(\mathbb{R}^2))$ . Finally, for  $|\beta - \gamma| = 2$  case i) again yields

$$I_2 \leq C\varepsilon^{1+a} \int_0^t \left\| \partial^{\beta-\gamma} \partial_t R_k(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} ds \leq C\varepsilon^{1+a} \int_0^t \tilde{z}(s) ds$$

where we have used  $\partial^{\gamma''} R_j \in L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})$  because  $|\gamma''| \leq 1$ .

At last, we treat  $\partial^\beta (W(\mathbf{R}) \mathbf{R}) \cdot \partial^\beta \mathbf{R}$ . Terms quadratic in  $\mathbf{U}_{\text{ext}}$  are estimated as follows, where  $\beta = \beta' + \beta''$ . If  $|\beta'| < 3$  or if not all three derivatives fall on  $\partial_t U_{\text{ext},j}$ , we obtain

$$\begin{aligned}
I_3 &:= \int_0^t \int_{\mathbb{R}^2} \partial^{\beta'} (U_{\text{ext},i} \partial_t U_{\text{ext},j}) \partial^{\beta''} R_k \partial^\beta R_l d\mathbf{x} ds \leq C\varepsilon^2 \int_0^t \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left\| \partial^{\beta''} R_k(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} ds \\
&\leq C\varepsilon^2 \int_0^t \tilde{z}(s) ds
\end{aligned}$$

as  $\|\partial^{\beta'} (U_{\text{ext},i} \partial_t U_{\text{ext},j})\|_{L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})} \leq C\varepsilon^2$  by (45). If  $|\beta'| = 3$  and  $\partial^{\beta'}$  is only applied to  $\partial_t U_{\text{ext},j}$ , we use (47) with  $\partial^{\beta'} \partial_t U_{\text{ext},j} = \mathcal{A}_j + \mathcal{B}_j$ . Sobolev's embedding for  $x_2 \mapsto R_k(x_1, x_2, s)$  implies that

$$\begin{aligned}
I_3 &\leq \left| \int_0^t \int_{\mathbb{R}^2} U_{\text{ext},i} \mathcal{A}_j \partial^{\beta''} R_k \partial^\beta R_l d\mathbf{x} ds \right| + \left| \int_0^t \int_{\mathbb{R}^2} U_{\text{ext},i} \mathcal{B}_j \partial^{\beta''} R_k \partial^\beta R_l d\mathbf{x} ds \right| \\
&\leq C\varepsilon^2 \int_0^t \tilde{z}(s) ds + \int_0^t \left\| \partial^\beta R_l(\cdot, s) \right\|_{L^2(\mathbb{R}^2)} \left[ \int_{\mathbb{R}^2} |U_{\text{ext},i} \mathcal{B}_j R_k|^2 d\mathbf{x} \right]^{\frac{1}{2}} ds
\end{aligned}$$

$$\begin{aligned}
&\leq C\varepsilon^2 \int_0^t \tilde{z}(s) \, ds + C\varepsilon^2 \int_0^t \|\partial^\beta R_l(\cdot, s)\|_{L^2(\mathbb{R}^2)} \left[ \int_{\mathbb{R}^2} \sup_{x_2 \in \mathbb{R}} |R_k(x_1, x_2, s)|^2 \, dx_1 \right]^{\frac{1}{2}} \, ds \\
&\leq C\varepsilon^2 \int_0^t \tilde{z}(s) \, ds + C\varepsilon^2 \int_0^t \|\partial^\beta R_l(\cdot, s)\|_{L^2(\mathbb{R}^2)} \left[ \int_{\mathbb{R}^2} (|R_k(x_1, x_2, s)|^2 + |\partial_{x_2} R_k(x_1, x_2, s)|^2) \, dx \right]^{\frac{1}{2}} \, ds \\
&\leq C\varepsilon^2 \int_0^t \tilde{z}(s) \, ds. \tag{71}
\end{aligned}$$

In the same way, we treat terms linear in  $\mathbf{U}_{\text{ext}}$ . Let  $\beta = \beta' + \beta'' + \beta'''$  and w.l.o.g.  $|\beta'''| \leq 1$ . If  $|\beta'| \leq 2$ , it follows

$$\begin{aligned}
I_4 &:= \varepsilon^a \int_0^t \int_{\mathbb{R}^2} \partial^{\beta'} \partial_t U_{\text{ext},i} \partial^{\beta''} R_j \partial^{\beta'''} R_k \partial^\beta R_l \, dx \, ds \\
&\leq C\varepsilon^{1+a} \int_0^t \|\partial^\beta R_l(\cdot, s)\|_{L^2(\mathbb{R}^2)} \|\partial^{\beta''} R_j(\cdot, s)\|_{L^2(\mathbb{R}^2)} \, ds \\
&\leq C\varepsilon^{1+a} \int_0^t \tilde{z}(s) \, ds,
\end{aligned}$$

using  $\partial^{\beta'''} R_k, \partial^{\beta'} \partial_t U_{\text{ext},i} \in L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})$ . If  $|\beta'| = 3$ , as above we estimate

$$\begin{aligned}
I_4 &\leq \varepsilon^a \left| \int_0^t \int_{\mathbb{R}^2} \mathcal{A}_i R_j R_k \partial^\beta R_l \, dx \, ds \right| + \varepsilon^a \left| \int_0^t \int_{\mathbb{R}^2} \mathcal{B}_i R_j R_k \partial^\beta R_l \, dx \, ds \right| \\
&\leq C\varepsilon^{1+a} \int_0^t \tilde{z}(s) \, ds + C\varepsilon^a \int_0^t \|\partial^\beta R_l(\cdot, s)\|_{L^2(\mathbb{R}^2)} \left( \int_{\mathbb{R}^2} |\mathcal{B}_i R_j R_k|^2 \, dx \right)^{\frac{1}{2}} \, ds \\
&\leq C\varepsilon^{1+a} \int_0^t \tilde{z}(s) \, ds + C\varepsilon^{1+a} \int_0^t \|\partial^\beta R_l(\cdot, s)\|_{L^2(\mathbb{R}^2)} \left( \int_{\mathbb{R}^2} \sup_{x_2 \in \mathbb{R}} |R_k(x_1, x_2, s)|^2 \, dx_1 \right)^{\frac{1}{2}} \, ds \\
&\leq C\varepsilon^{1+a} \int_0^t \tilde{z}(s) \, ds.
\end{aligned}$$

Note that these are the only cases where four derivatives can fall on one function in this step.

Collecting the above partial estimates, we finally get in (68)

$$\tilde{z}(t) \leq C \left( \rho_0^2 + \varepsilon^2 \int_0^t \tilde{z}(s) \, ds + \varepsilon^{1+a} t + \varepsilon^{\frac{3}{2}-a} + \varepsilon^{7-2a} \right).$$

If  $a \in (1, \frac{11}{2})$ , the Gronwall's inequality yields

$$\tilde{z}(t) \leq C \left( \rho_0^2 + \varepsilon^{\frac{3}{2}-a} + \varepsilon^{1+a} t \right) e^{C\varepsilon^2 t} \leq C \left( \rho_0^2 + \varepsilon^{\frac{3}{2}-a} + \varepsilon^{a-1} \right)$$

for all  $t \in J_{\rho_0, \varepsilon_0}$  if  $\|\mathbf{R}^{(0)}\|_{\mathcal{H}^3(\mathbb{R}^2)^3} \leq \rho_0$ .

### Step III: Analysis of $\partial^\beta \mathbf{R}_{2,3}$ for $|\beta| \leq 3, \beta_1 \neq 0$

We first consider  $\beta_1 = 1$ . Setting  $\alpha := (0, \beta_2, \beta_3)^\top$ , we have  $\beta = (1, 0, 0)^\top + \alpha$  and

$$\begin{cases} \partial^\beta R_2 = \partial_{x_2} \partial^\alpha R_1 - (S(\mathbf{R}) \partial_t \partial^\alpha \mathbf{R} + \mathbf{s}_\alpha(\mathbf{R}) + \mathbf{w}_\alpha(\mathbf{R}) + \varepsilon^{-a} \partial^\alpha \mathbf{Res})_3, \\ \partial^\beta R_3 = - (S(\mathbf{R}) \partial_t \partial^\alpha \mathbf{R} + \mathbf{s}_\alpha(\mathbf{R}) + \mathbf{w}_\alpha(\mathbf{R}) + \varepsilon^{-a} \partial^\alpha \mathbf{Res})_2. \end{cases} \tag{72}$$

Each term on the right-hand side has derivatives  $\partial^\gamma$  with  $|\gamma| \leq 3$  and  $\gamma_1 = 0$  and hence can be bounded by Step II, for instance,

$$\|(S(\mathbf{R}) \partial_t \partial^\alpha \mathbf{R})(\cdot, t)\|_{L^2(\mathbb{R}^2)^3} \leq \|S(\mathbf{R})(\cdot, t)\|_{L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})} \|(\partial_t \partial^\alpha \mathbf{R})(\cdot, t)\|_{L^2(\mathbb{R}^2)^3} \leq C \|\partial^\gamma \mathbf{R}\|_{L^2(\mathbb{R}^2)^3}.$$

In summary, we get

$$\|\partial^\beta R_{2,3}\|_{L^2(\mathbb{R}^2)^3}^2 \leq C\tilde{z}(t) \leq C\left(\rho_0^2 + \varepsilon^{\frac{3}{2}-a} + \varepsilon^{a-1}\right)$$

for all  $|\beta| \leq 3$ ,  $\beta_1 = 1$  and all  $t \in J_{\rho_0, \varepsilon_0}$  if  $a \in (1, \frac{11}{2})$  and  $\|\mathbf{R}^{(0)}\|_{\mathcal{H}^3(\mathbb{R}^2)^3} \leq \rho_0$ .

For larger values of  $\beta_1$  we iterate the process. For  $\beta_1 = 2$  we have (72) with  $\beta = (1, 0, 0)^\top + \alpha$  and  $\alpha := (1, \beta_2, \beta_t)^\top$  and using the previous step, all terms in the right-hand side can be estimated in  $L^2(\mathbb{R}^2)$ . For  $\beta_1 = 3$  the same process applies, with  $\beta = (1, 0, 0)^\top + \alpha$  and  $\alpha = (2, 0, 0)^\top$ . Altogether, we arrive at

$$\sum_{\substack{|\beta| \leq 3, \\ \beta_1 = 0}} \|\partial^\beta \mathbf{R}(\cdot, t)\|_{L^2(\mathbb{R}^2)^3}^2 + \sum_{|\beta| \leq 3} \left( \|\partial^\beta R_2(\cdot, t)\|_{L^2(\mathbb{R}^2)}^2 + \|\partial^\beta R_3(\cdot, t)\|_{L^2(\mathbb{R}^2)}^2 \right) \leq C\left(\rho_0^2 + \varepsilon^{\frac{3}{2}-a} + \varepsilon^{a-1}\right) \quad (73)$$

for all  $t \in J_{\rho_0, \varepsilon_0}$  if  $a \in (1, \frac{11}{2})$  and  $\|\mathbf{R}^{(0)}\|_{\mathcal{H}^3(\mathbb{R}^2)^3} \leq \rho_0$ .

#### Step IV: Analysis of $\partial^\beta \mathbf{R}_1$ , $|\beta| \leq 3$

In this final step we exploit the divergence equation  $\nabla \cdot \mathcal{D}(\mathbf{U}) = \nabla \cdot \mathcal{D}(\mathbf{U}^{(0)})$  to estimate  $\partial^\beta \mathbf{R}_1$ . First, for  $\alpha \in \mathbb{N}_0^3$ ,  $|\alpha| \leq 2$  and  $\mathbf{r}_\alpha = \partial^\alpha \mathbf{R}$  we have

$$\begin{aligned} \varepsilon^{-a} \partial^\alpha \partial_t \tilde{\mathcal{D}}(\varepsilon^a \mathbf{R} + \mathbf{U}_{\text{ext}}) &= \partial_t \left( (\varepsilon_1 + \varepsilon^{2a} \varepsilon_3 |\tilde{\mathbf{R}}|^2) \tilde{\mathbf{r}}_\alpha \right) + \partial^\alpha \left( \tilde{\varphi}(\mathbf{R}) \partial_t \tilde{\mathbf{R}} \right) + \partial^\alpha \left( \tilde{W}(\mathbf{R}) \tilde{\mathbf{R}} \right) \\ &+ \partial_t \left( \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^\gamma \left( \varepsilon_1 + \varepsilon^{2a} \varepsilon_3 |\tilde{\mathbf{R}}|^2 \right) \partial^{\alpha-\gamma} \tilde{\mathbf{R}} \right) + \varepsilon^{-a} \partial^\alpha \tilde{\mathbf{R}} \text{es} \quad (74) \\ &+ \varepsilon^{-a} \partial^\alpha \left( \begin{array}{c} \partial_{x_2} U_{\text{ext},3} \\ -\partial_{x_1} U_{\text{ext},3} \end{array} \right) \end{aligned}$$

on  $\mathbb{R}_\pm^2 \times J_{\rho_0, \varepsilon_0}$ , where  $\tilde{\cdot}$  of a  $3 \times 3$ -matrix denotes the restriction to the upper left  $(2 \times 2)$ -submatrix and  $\tilde{\cdot}$  of a vector in  $\mathbb{R}^3$  denotes the first two components of this vector. The calculation to obtain (74) uses that  $\varphi(\mathbf{R})$  and  $W(\mathbf{R})$  have a block structure and that the sum of the last two terms in (74) is  $\varepsilon^{-a} \partial_t \tilde{\mathcal{D}}(\mathbf{U}_{\text{ext}})$ .

An integration by parts yields

$$\begin{aligned} \int_0^t \partial^\alpha \left( \tilde{\varphi}(\mathbf{R}) \partial_t \tilde{\mathbf{R}} \right) ds &= \int_0^t \left( \tilde{\varphi}(\mathbf{R}) \partial^\alpha \partial_t \tilde{\mathbf{R}} + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^\gamma \tilde{\varphi}(\mathbf{R}) \partial^{\alpha-\gamma} \partial_t \tilde{\mathbf{R}} \right) ds \\ &= \int_0^t \left( -\partial_t \tilde{\varphi}(\mathbf{R}) \partial^\alpha \tilde{\mathbf{R}} + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^\gamma \tilde{\varphi}(\mathbf{R}) \partial^{\alpha-\gamma} \partial_t \tilde{\mathbf{R}} \right) ds + \left[ \tilde{\varphi}(\mathbf{R}) \partial^\alpha \tilde{\mathbf{R}} \right]_0^t. \end{aligned}$$

Integrating (74) in time, we then deduce

$$\begin{aligned} &\left[ \varepsilon^{-a} \partial^\alpha \tilde{\mathcal{D}}(\varepsilon^a \mathbf{R} + \mathbf{U}_{\text{ext}}) \right]_0^t \\ &= \left[ (\varepsilon_1 + \varepsilon^{2a} \varepsilon_3 |\tilde{\mathbf{R}}|^2) \tilde{\mathbf{r}}_\alpha + \tilde{\varphi}(\mathbf{R}) \tilde{\mathbf{r}}_\alpha + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^\gamma \left( \varepsilon_1 + \varepsilon^{2a} \varepsilon_3 |\tilde{\mathbf{R}}|^2 \right) \partial^{\alpha-\gamma} \tilde{\mathbf{R}} \right]_0^t \\ &+ \int_0^t \left( -\partial_t \tilde{\varphi}(\mathbf{R}) \tilde{\mathbf{r}}_\alpha + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^\gamma \tilde{\varphi}(\mathbf{R}) \partial^{\alpha-\gamma} \partial_t \tilde{\mathbf{R}} + \partial^\alpha \left( \tilde{W}(\mathbf{R}) \tilde{\mathbf{R}} \right) \right) ds \quad (75) \\ &+ \varepsilon^{-a} \int_0^t \left( \partial^\alpha \tilde{\mathbf{R}} \text{es} + \partial^\alpha \left( \begin{array}{c} \partial_{x_2} U_{\text{ext},3} \\ -\partial_{x_1} U_{\text{ext},3} \end{array} \right) \right) ds. \end{aligned}$$

Note that the divergence of the last term vanishes.

**Substep 1:**  $\beta_1 = 1$ . We write  $\beta = (1, 0, 0)^\top + \alpha$ , where  $\alpha = (0, \beta_2, \beta_t)^\top$ . We have that  $\nabla \cdot \partial^\alpha \tilde{\mathcal{D}}(\mathbf{U})$  is constant in time because

$$\nabla \cdot \partial^\alpha \tilde{\mathcal{D}}(\mathbf{U}) = \partial^\alpha (\nabla \cdot \tilde{\mathcal{D}}(\mathbf{U})) = \partial^\alpha \tilde{\varrho}, \quad \tilde{\varrho} := \nabla \cdot \tilde{\mathcal{D}}(\mathbf{U}^{(0)}).$$

Note that  $\tilde{\varrho} \in \mathcal{H}^2(\mathbb{R}^2)$  because of the algebra property of  $\mathcal{H}^2(\mathbb{R}^2)$  and  $\mathbf{U}^{(0)} \in \mathcal{H}^3(\mathbb{R}^2)^3$ . Hence, taking the divergence of (75), the first term vanishes and we have

$$\begin{aligned} & \left[ (\epsilon_1 + \varepsilon^{2a} \epsilon_3 |\tilde{\mathbf{R}}|^2) (\partial_{x_1} r_{\alpha,1} + \partial_{x_2} r_{\alpha,2}) + \nabla \left( \epsilon_1 + \varepsilon^{2a} \epsilon_3 |\tilde{\mathbf{R}}|^2 \right) \cdot \tilde{\mathbf{r}}_{\alpha} \right]_0^t \\ &= - \left[ \nabla \cdot \left( \tilde{\varphi}(\mathbf{R}) \partial^{\alpha} \tilde{\mathbf{R}} + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^{\gamma} (\epsilon_1 + \varepsilon^{2a} \epsilon_3 |\tilde{\mathbf{R}}|^2) \partial^{\alpha-\gamma} \tilde{\mathbf{R}} \right) \right]_0^t \\ & \quad - \int_0^t \nabla \cdot \left( -\partial_t \tilde{\varphi}(\mathbf{R}) \partial^{\alpha} \tilde{\mathbf{R}} + \partial^{\alpha} (\tilde{W}(\mathbf{R}) \tilde{\mathbf{R}}) + \varepsilon^{-a} \partial^{\alpha} \tilde{\mathbf{Res}} + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^{\gamma} \tilde{\varphi}(\mathbf{R}) \partial^{\alpha-\gamma} \partial_t \tilde{\mathbf{R}} \right) ds. \end{aligned}$$

Because of  $\tilde{\mathbf{R}} \in C(J_{\rho_0, \varepsilon_0}, L^{\infty}(\mathbb{R}^2))^2$ , there exists a number  $\vartheta > 0$  with  $(\epsilon_1 + \varepsilon^{2a} \epsilon_3 |\tilde{\mathbf{R}}|^2)(\mathbf{x}, t) \geq \vartheta$  for small enough  $\varepsilon$ , all  $t \in J_{\rho_0, \varepsilon_0}$  and almost all  $\mathbf{x} \in \mathbb{R}^2$ . Since  $\partial_{x_1} \mathbf{R}, \partial_{x_2} \mathbf{R} \in C(J_{\rho_0, \varepsilon_0}, L^{\infty}(\mathbb{R}^2))^2$ , we can also estimate  $\left\| \nabla \left( \epsilon_1 + \varepsilon^{2a} \epsilon_3 |\tilde{\mathbf{R}}|^2 \right) (\cdot, t) \right\|_{L^{\infty}(\mathbb{R}^2)^3} \leq C$  for all  $t \in J_{\rho_0, \varepsilon_0}$ .

These facts yield the central inequality of this step:

$$\begin{aligned} & \vartheta \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} \\ & \leq C \left( \|\partial_{x_2} r_{\alpha,2}(\cdot, t)\|_{L^2(\mathbb{R}^2)} + \|\nabla \cdot \tilde{\mathbf{r}}_{\alpha}(\cdot, 0)\|_{L^2(\mathbb{R}^2)} + \|\tilde{\mathbf{r}}_{\alpha}(\cdot, t)\|_{L^2(\mathbb{R}^2)^2} + \|\tilde{\mathbf{r}}_{\alpha}(\cdot, 0)\|_{L^2(\mathbb{R}^2)^2} \right) \\ & \quad + \left\| \left[ \nabla \cdot \left( \tilde{\varphi}(\mathbf{R}) \partial^{\alpha} \tilde{\mathbf{R}} + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^{\gamma} (\epsilon_1 + \varepsilon^{2a} \epsilon_3 |\tilde{\mathbf{R}}|^2) \partial^{\alpha-\gamma} \tilde{\mathbf{R}} \right) (\cdot, s) \right]_0^t \right\|_{L^2(\mathbb{R}^2)}^t \quad (76) \\ & \quad + \left\| \int_0^t \nabla \cdot \left( -\partial_t \tilde{\varphi}(\mathbf{R}) \partial^{\alpha} \tilde{\mathbf{R}} + \partial^{\alpha} (\tilde{W}(\mathbf{R}) \tilde{\mathbf{R}}) + \sum_{0 < \gamma \leq \alpha} \binom{\alpha}{\gamma} \partial^{\gamma} \tilde{\varphi}(\mathbf{R}) \partial^{\alpha-\gamma} \partial_t \tilde{\mathbf{R}} \right) (\cdot, s) ds \right\|_{L^2(\mathbb{R}^2)} \\ & \quad + \varepsilon^{-a} \left\| \int_0^t \nabla \cdot \partial^{\alpha} \tilde{\mathbf{Res}}(\cdot, s) ds \right\|_{L^2(\mathbb{R}^2)} \end{aligned}$$

We next iterate over  $\beta_t$  and  $\beta_2$ .

(i)  $\alpha = (0, 0, 0)^{\top}$ . Here (76) simplifies to

$$\begin{aligned} & \vartheta \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} \leq \varepsilon^{-a} \left\| \int_0^t \nabla \cdot \tilde{\mathbf{Res}}(\cdot, s) ds \right\|_{L^2(\mathbb{R}^2)} \\ & \quad + C \left( \|\partial_{x_2} r_{\alpha,2}(\cdot, t)\|_{L^2(\mathbb{R}^2)} + \|\nabla \cdot \tilde{\mathbf{r}}_{\alpha}(\cdot, 0)\|_{L^2(\mathbb{R}^2)} + \|\tilde{\mathbf{r}}_{\alpha}(\cdot, t)\|_{L^2(\mathbb{R}^2)^2} + \|\tilde{\mathbf{r}}_{\alpha}(\cdot, 0)\|_{L^2(\mathbb{R}^2)^2} \right) \quad (77) \\ & \quad + \left\| \left[ \nabla \cdot \left( \tilde{\varphi}(\mathbf{R}) \tilde{\mathbf{R}} \right) (\cdot, s) \right]_0^t \right\|_{L^2(\mathbb{R}^2)} + \left\| \int_0^t \nabla \cdot \left( -\partial_t \tilde{\varphi}(\mathbf{R}) \tilde{\mathbf{R}} + \tilde{W}(\mathbf{R}) \tilde{\mathbf{R}} \right) (\cdot, s) ds \right\|_{L^2(\mathbb{R}^2)}. \end{aligned}$$

The residual term on the right-hand side is bounded by  $C\varepsilon^{\frac{3}{2}-a}$  due to (42). The second and fourth term on the right-hand side of (77) are estimated by (73) and the third and fifth term by (67). In the first norm on the last line of (77) all terms have been treated in Steps I, II or III except for those of the type  $\varepsilon^a \epsilon_3 \partial_{x_1} r_{\alpha,1} R_j U_{\text{ext},k}$  and  $\epsilon_3 \partial_{x_1} r_{\alpha,1} U_{\text{ext},j} U_{\text{ext},k}$ . Using  $\mathbf{R} \in L^{\infty}(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  and  $\|\mathbf{U}_{\text{ext}}\|_{L^{\infty}(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \leq C\varepsilon$ , we have

$$\begin{aligned} & \|\varepsilon^a \epsilon_3 (\partial_{x_1} r_{\alpha,1} R_j U_{\text{ext},k})(\cdot, t)\|_{L^2(\mathbb{R}^2)} \leq C\varepsilon^{1+a} \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)}, \\ & \|\epsilon_3 (\partial_{x_1} r_{\alpha,1} U_{\text{ext},j} U_{\text{ext},k})(\cdot, t)\|_{L^2(\mathbb{R}^2)} \leq C\varepsilon^2 \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)}. \end{aligned}$$

In the last norm of the right-hand side of (77), the terms which have not been estimated so far are of the type  $\partial_{x_1} r_{\alpha,1} \partial_t U_{\text{ext},j} U_{\text{ext},k}$ ,  $\varepsilon^a \partial_t (U_{\text{ext},j} R_k) \partial_{x_1} r_{\alpha,1}$ , and  $\varepsilon^a U_{\text{ext},j} R_k \partial_t \partial_{x_1} r_{\alpha,1}$  for  $j, k \in \{1, 2, 3\}$ . Using

$\mathbf{R}, \partial_t \mathbf{R} \in L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  and  $\|\partial_t \mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \leq C\varepsilon$ , we obtain

$$\begin{aligned} \int_0^t \|\varepsilon_3 (\partial_{x_1} r_{\alpha,1} \partial_t U_{\text{ext},j} U_{\text{ext},k}) (\cdot, s)\|_{L^2(\mathbb{R}^2)} ds &\leq C\varepsilon^2 \int_0^t \|\partial_{x_1} r_{\alpha,1}(\cdot, s)\|_{L^2(\mathbb{R}^2)} ds, \\ \int_0^t \|\varepsilon^a \varepsilon_3 (\partial_{x_1} r_{\alpha,1} \partial_t (R_k U_{\text{ext},j})) (\cdot, s)\|_{L^2(\mathbb{R}^2)} ds &\leq C\varepsilon^{1+a} \int_0^t \|\partial_{x_1} r_{\alpha,1}(\cdot, s)\|_{L^2(\mathbb{R}^2)} ds, \end{aligned}$$

and, integrating by parts in time,

$$\begin{aligned} &\left\| \int_0^t \varepsilon^a \varepsilon_3 (U_{\text{ext},j} R_k \partial_t \partial_{x_1} r_{\alpha,1}) (\cdot, s) ds \right\|_{L^2(\mathbb{R}^2)} \\ &\leq \left\| \left[ \varepsilon^a (U_{\text{ext},j} R_k \partial_{x_1} r_{\alpha,1}) (\cdot, s) \right]_0^t - \int_0^t \varepsilon^a (\partial_t (U_{\text{ext},j} R_k) \partial_{x_1} r_{\alpha,1}) (\cdot, s) ds \right\|_{L^2(\mathbb{R}^2)} \\ &\leq C\varepsilon^{1+a} \left( \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} + \|\partial_{x_1} r_{\alpha,1}(\cdot, 0)\|_{L^2(\mathbb{R}^2)} + \int_0^t \|\partial_{x_1} r_{\alpha,1}(\cdot, s)\|_{L^2(\mathbb{R}^2)} ds \right) \\ &\leq C\varepsilon^{1+a} \left( \rho_0 + \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} + \int_0^t \|\partial_{x_1} r_{\alpha,1}(\cdot, s)\|_{L^2(\mathbb{R}^2)} ds \right). \end{aligned}$$

Combining the above inequalities, for  $a \in (1, \frac{11}{2})$  and  $0 \leq t \leq T_{\rho_0, \varepsilon_0} \leq T_0 \varepsilon^{-2}$  we infer

$$\begin{aligned} \vartheta \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} &\leq C \left( \rho_0 + \varepsilon^{\frac{1}{2}(\frac{3}{2}-a)} + \varepsilon^{\frac{1}{2}(a-1)} + \varepsilon^2 \|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} + \varepsilon^{\frac{3}{2}-a} \right) \\ &\quad + C\varepsilon^2 \int_0^t \left( \rho_0 + \varepsilon^{\frac{1}{2}(\frac{3}{2}-a)} + \varepsilon^{\frac{1}{2}(a-1)} \right) ds + C\varepsilon^2 \int_0^t \|\partial_{x_1} r_{\alpha,1}(\cdot, s)\|_{L^2(\mathbb{R}^2)} ds. \end{aligned}$$

For  $\varepsilon$  small enough and  $a \in [\frac{5}{4}, \frac{11}{2})$  (so that  $\frac{3}{2} - a \leq a - 1$ ) it follows

$$\|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} \leq C \left( \rho_0 + \varepsilon^{\frac{1}{2}(\frac{3}{2}-a)} + \varepsilon^2 \int_0^t \|\partial_{x_1} r_{\alpha,1}(\cdot, s)\|_{L^2(\mathbb{R}^2)} ds \right).$$

Finally, Gronwall's inequality yields

$$\|\partial_{x_1} r_{\alpha,1}(\cdot, t)\|_{L^2(\mathbb{R}^2)} \leq C \left( \rho_0 + \varepsilon^{\frac{1}{2}(\frac{3}{2}-a)} \right) e^{C\varepsilon^2 t} \leq C \left( \rho_0 + \varepsilon^{\frac{1}{2}(\frac{3}{2}-a)} \right).$$

(ii) We iterate the process from (i) for higher  $\alpha_2 = \beta_2$  and  $\alpha_t = \beta_t$  (keeping  $\beta_1 = 1$ ). For instance, the following sequence of  $\alpha$ 's can be chosen:  $\alpha = (0, 1, 0)^\top, (0, 0, 1)^\top, (0, 2, 0)^\top, (0, 0, 2)^\top, (0, 1, 1)^\top$ . Note that  $|\alpha| = \beta_t + \beta_2 \leq 2$  therefore we can always use integration by parts and Lemma C.1. In the terms with  $\widetilde{W}$  again three derivatives can fall on  $\partial_t U_{\text{ext},k}$ . If  $\partial_{x_1}$  is included, then one can proceed as above by means of (45). Otherwise, one uses (47) and argues as in (71).

**Substep 2:**  $\beta_1 > 1$ . In this last step we have to iterate over  $\beta_1$  and increase it to 3. For  $\beta_1 = 2$  we set  $\beta = (1, 0, 0)^\top + \alpha$  with  $\alpha = (1, \beta_2, \beta_t)^\top$ . The estimates work like in Substep 1(i) since we have  $|\alpha| \leq 2$ . Finally, for  $\beta_1 = 3$  we have  $\beta = (3, 0, 0)^\top = (1, 0, 0)^\top + \alpha$  with  $\alpha = (2, 0, 0)^\top$  and apply Substep 1(i) again. Here, factors  $\partial_{x_1}^3 \partial_t U_{\text{ext},k}$  occur in the terms with  $\widetilde{W}$ , which are treated with (46).

In summary, collecting all the above estimates, one concludes

$$z(t) \leq C \left( \rho_0^2 + \varepsilon^{\frac{3}{2}-a} \right)$$

for every  $t \in J_{\rho_0, \varepsilon_0}$  and  $\varepsilon \in (0, \varepsilon_0)$  if  $a \in [\frac{5}{4}, \frac{7}{2})$  and  $\varepsilon_0$  is small enough.

Next, we keep  $\rho$  fixed, choose  $a \in [\frac{5}{4}, \frac{3}{2})$  and  $\rho_0, \varepsilon_0$  so small that  $C \left( \rho_0^2 + \varepsilon_0^{3/2-a} \right) < \frac{1}{2}\rho^2$  and

$$\varepsilon_0^a \rho + \|\mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times [0, T_0 \varepsilon^{-2}])^3} \leq \varpi$$



where we recall (59) and that  $\|\mathbf{U}_{\text{ext}}\|_{L^\infty(\mathbb{R}^2 \times [0, T_0 \varepsilon^{-2}])^3} \leq C\varepsilon \leq C\varepsilon_0$ . With this choice we have

$$z(t) < \frac{1}{2}\rho^2$$

for every  $t \in J_{\rho_0, \varepsilon_0}$  and  $\varepsilon \in (0, \varepsilon_0)$  if  $a \in [\frac{5}{4}, \frac{3}{2})$ . The definition (61) of  $T_{\rho_0, \varepsilon_0}$  now implies that  $T_{\rho_0, \varepsilon_0} = T_0 \varepsilon^{-2} < t_M$  and that (60) holds with  $t_* = T_0 \varepsilon^{-2}$ .

## 6.2 Final error estimate

To finalize the proof of Theorem 1.1, we first compare  $\mathbf{U}_{\text{ext}}$  from (40) and  $\mathbf{U}_{\text{ans}}$  from (12). Similar as deducing  $\text{Res}(\mathbf{U}_{\text{ext}}) \in \mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3$  from (43), one can show that the same condition yields

$$\begin{aligned} & \|\mathbf{U}_{\text{ext}} - \mathbf{U}_{\text{ans}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \\ &= \left\| \left( -\varepsilon^2 i \partial_{X_2} A \partial_k \mathbf{w}(k_0) - \varepsilon^3 \frac{1}{2} \partial_{X_2}^2 A \partial_k^2 \mathbf{w}(k_0) + \varepsilon^3 |A|^2 A \mathbf{p} \right) F_1 + \varepsilon^3 A^3 \mathbf{h} F_1^3 + \text{c.c.} \right\|_{\mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \\ &\leq C\varepsilon^{\frac{3}{2}}. \end{aligned}$$

Note that we again lose half a power of  $\varepsilon$  since  $A$  depends on  $X_2 = \varepsilon(x_2 - \nu_1 t)$ . Second, we use (62) and the triangle inequality to conclude

$$\begin{aligned} \|\mathbf{U} - \mathbf{U}_{\text{ans}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} &\leq \|\mathbf{U} - \mathbf{U}_{\text{ext}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} + \|\mathbf{U}_{\text{ext}} - \mathbf{U}_{\text{ans}}\|_{\mathcal{G}^3(\mathbb{R}^2 \times J_{\rho_0, \varepsilon_0})^3} \\ &\leq C \left( \varepsilon^a + \varepsilon^{\frac{3}{2}} \right) \leq C_\delta \varepsilon^{\frac{3}{2} - \delta} \end{aligned}$$

for all  $\delta > 0$ . □

## A Numerical method for the eigenvalue problem

To solve (15) numerically we rewrite the problem as a second-order ordinary differential equation

$$\begin{aligned} \partial_{x_1}^2 w_3 &= i \partial_{x_1} \varepsilon_1(x_1) \omega w_2 + i \varepsilon_1(x_1) \omega \partial_{x_1} w_2 \\ &= \frac{\partial_{x_1} \varepsilon_1(x_1)}{\varepsilon_1(x_1)} \partial_{x_1} w_3 - \varepsilon_1(x_1) \omega (\mu_0 \omega w_3 + k w_1) \\ &= \frac{\partial_{x_1} \varepsilon_1(x_1)}{\varepsilon_1(x_1)} \partial_{x_1} w_3 - \varepsilon_1(x_1) \mu_0 \omega^2 w_3 + k^2 w_3 \end{aligned}$$

on  $\mathbb{R} \setminus \{0\}$ . The interface condition  $\llbracket w_2 \rrbracket_{\text{ID}} = 0$  implies that  $\left\llbracket \frac{\partial_{x_1} w_3}{\varepsilon_1} \right\rrbracket_{\text{ID}} = 0$ . Now we have to solve the eigenvalue problem

$$\begin{cases} -\partial_{x_1}^2 w_3(x_1) + \frac{\partial_{x_1} \varepsilon_1(x_1)}{\varepsilon_1(x_1)} \partial_{x_1} w_3(x_1) + k^2 w_3(x_1) = \varepsilon_1(x_1) \mu_0 \omega^2 w_3(x_1), & x_1 \in \mathbb{R} \setminus \{0\}, \\ \llbracket w_3 \rrbracket_{\text{ID}} = \left\llbracket \frac{\partial_{x_1} w_3}{\varepsilon_1} \right\rrbracket_{\text{ID}} = 0. \end{cases} \quad (78)$$

Note that we can use  $w_1 = -\frac{k}{\varepsilon_1 \omega} w_3$  and  $w_2 = -\frac{i}{\varepsilon_1 \omega} \partial_{x_1} w_3$  to calculate the remaining components of  $\mathbf{w}$ . We also see that the interface conditions  $\llbracket \varepsilon_1 w_1 \rrbracket_{\text{ID}} = \llbracket w_2 \rrbracket_{\text{ID}} = 0$  are satisfied if  $w_3$  solves (78).

To simplify the numerics we write  $w_3 = w_{3,r} + w_{3,s}$ , with a smooth function  $w_{3,r}$  and a function  $w_{3,s}$  that has a discontinuous first derivative at  $x_1 = 0$ . For instance, we take

$$w_{3,s}(x_1) = \begin{cases} w_{3,s}^- = \text{const.}, & x_1 \leq 0, \\ w_{3,s}^+(x_1), & x_1 > 0, \end{cases}$$

and choose  $w_{3,s}^+(0) = w_{3,s}^-$  so that  $w_{3,s}$  is continuous. Note that with this choice  $w_3$  satisfies the first interface condition. For the second interface condition we calculate  $\partial_{x_1} w_3$  and get that

$$\begin{aligned} \left[ \frac{\partial_{x_1} w_3}{\epsilon_1 \omega} \right]_{\text{ID}} = 0 &\iff \epsilon_1^-(0) (\partial_{x_1} w_{3,r}(0) + \partial_{x_1} w_{3,s}^+(0)) = \epsilon_1^+(0) \partial_{x_1} w_{3,r}(0) \\ &\iff \partial_{x_1} w_{3,s}^+(0) = \frac{\epsilon_1^+(0) - \epsilon_1^-(0)}{\epsilon_1^-(0)} \partial_{x_1} w_{3,r}(0) =: \tilde{\epsilon} \partial_{x_1} w_{3,r}(0). \end{aligned}$$

We now set

$$w_{3,s}(x_1) = (\mathcal{L}w_{3,r})(x_1) := \begin{cases} -\text{sgn}(\tilde{\epsilon}) \partial_{x_1} w_{3,r}(0), & x_1 < 0, \\ -\text{sgn}(\tilde{\epsilon}) \partial_{x_1} w_{3,r}(0) e^{-|\tilde{\epsilon}|x_1}, & x_1 \geq 0. \end{cases}$$

to satisfy the second interface condition. Thus,  $w_{3,r}$  has to solve

$$\begin{cases} \left( -\partial_{x_1}^2 + \frac{\partial_{x_1} \epsilon_1(x_1)}{\epsilon_1(x_1)} \partial_{x_1} + k^2 \right) ((I + \mathcal{L})w_{3,r}(x_1)) = \epsilon_1(x_1) \mu_0 \omega^2 (I + \mathcal{L})w_{3,r}(x_1), & x_1 \in \mathbb{R} \setminus \{0\}, \\ \llbracket w_{3,r} \rrbracket_{\text{ID}} = \llbracket \partial_{x_1} w_{3,r} \rrbracket_{\text{ID}} = 0. \end{cases} \quad (79)$$

We are interested in  $H^1(\mathbb{R})$ -solutions, therefore we have at least the boundary conditions

$$\lim_{x_0 \rightarrow -\infty} w_{3,r}(x_1) = \text{sgn}(\tilde{\epsilon}) \partial_{x_1} w_{3,r}(0), \quad \lim_{x_1 \rightarrow \infty} w_{3,r}(x_1) = 0.$$

To solve (79) numerically for a fixed  $k \in \mathbb{R}$  we discretize the problem over a finite interval  $[-d, d] \subset \mathbb{R}$  and apply a solver for a generalized eigenvalue problem, e.g. a solver based on a Krylov-Schur algorithm. To be more precise, we used  $d$  ranging from  $10^2$  to  $10^4$  with the step size  $h = 0.01$  in space. We used the second-order difference quotients with zero Dirichlet boundary conditions to discretize the derivatives. The generalized eigenvalue problem was then solved with the Matlab functions ‘‘eigs’’, where we calculated the first 10 eigenvalues closest to  $\nu_0$  with a convergence tolerance of  $10^{-10}$ . We then only selected solutions where the corresponding eigenfunctions were almost zero in a small neighborhood of the boundary of  $[-d, d]$ , i.e., the norm of  $w_3$  on  $[-d, -d + 100h] \cup [d - 100h, d]$  is smaller than  $10^{-6}$ .

## B Residual of order $\epsilon^4$

For  $\text{Res} := \text{Res}(U_{\text{ext}})$  and  $F_1 = e^{i(k_0 x_2 - \nu_0 t)}$  we have

$$\begin{aligned} \text{Res}_3 &= F_1 \epsilon^4 \left( \frac{1}{2} \partial_{X_2}^3 A \partial_k^2 w_1(k_0) - (2|A|^2 \partial_{X_2} A + A^2 \partial_{X_2} \bar{A}) p_1 \right) \\ &\quad + F_1 \epsilon^4 \mu_0 \left( \frac{\nu_1}{2} \partial_{X_2}^3 A \partial_k^2 w_3(k_0) - i \partial_T \partial_{X_2} A \partial_k w_3(k_0) - \nu_1 (2|A|^2 \partial_{X_2} A + A^2 \partial_{X_2} \bar{A}) p_3 \right) \\ &\quad - 3F_1^3 \epsilon^4 (\partial_{X_2} A A^2 h_1 + \mu_0 \nu_1 A^2 \partial_{X_2} A h_3) + \text{c.c.} + \mathcal{O}(\epsilon^5) \end{aligned}$$

and the parts of  $\text{Res}_1$  and  $\text{Res}_2$  that are linear in  $U_{\text{ext}}$  are given by

$$\begin{aligned} \text{Res}_{\text{lin},1} &= F_1 \epsilon^4 \left( \frac{1}{2} \partial_{X_2}^3 A \partial_k^2 w_3(k_0) - (2|A|^2 \partial_{X_2} A + A^2 \partial_{X_2} \bar{A}) p_3 \right) \\ &\quad + F_1 \epsilon^4 \epsilon_1 \left( \frac{\nu_1}{2} \partial_{X_2}^3 A \partial_k^2 w_1(k_0) - i \partial_T \partial_{X_2} A \partial_k w_1(k_0) - \nu_1 (2|A|^2 \partial_{X_2} A + A^2 \partial_{X_2} \bar{A}) p_1 \right) \\ &\quad - 3F_1^3 \epsilon^4 (\epsilon_1 \nu_1 A^2 \partial_{X_2} A h_1 + A^2 \partial_{X_2} A h_3) + \text{c.c.} + \mathcal{O}(\epsilon^5), \\ \text{Res}_{\text{lin},2} &= F_1 \epsilon^4 \epsilon_1 \left( \frac{\nu_1}{2} \partial_{X_2}^3 A \partial_k^2 w_2(k_0) - i \partial_T \partial_{X_2} A \partial_k w_2(k_0) - \nu_1 (2|A|^2 \partial_{X_2} A + A^2 \partial_{X_2} \bar{A}) p_2 \right) \\ &\quad - 3F_1^3 \epsilon^4 \epsilon_1 \nu_1 A^2 \partial_{X_2} A h_2 + \text{c.c.} + \mathcal{O}(\epsilon^5). \end{aligned}$$

For the nonlinear part of  $\text{Res}_1$  we get

$$\begin{aligned} \text{Res}_{\text{nl},1} = & -\varepsilon^4 \epsilon_3 (3\nu_1 F_1^3 A^2 \partial_{X_2} A (m_1^3 + m_1 m_2^2) \\ & + \nu_0 F_1^3 A^2 \partial_{X_2} A (3m_1^2 \partial_k w_1(k_0) + m_2^2 \partial_k w_1(k_0) + 2m_1 m_2 \partial_k w_2(k_0)) \\ & + \nu_1 F_1 A^2 \partial_{X_2} \bar{A} (3|m_1|^2 m_1 + 2m_1 |m_2|^2 + \bar{m}_1 m_2^2) \\ & + \nu_0 F_1 A^2 \partial_{X_2} \bar{A} (3m_1^2 \partial_k \bar{w}_1(k_0) + m_2^2 \partial_k \bar{w}_1(k_0) + 2m_1 m_2 \partial_k \bar{w}_2(k_0)) \\ & + 2\nu_0 F_1 |A|^2 \partial_{X_2} A (3|m_1|^2 \partial_k w_1(k_0) + |m_2|^2 \partial_k w_1(k_0) + \bar{m}_1 m_2 \partial_k w_2(k_0) + m_1 m_2 \partial_k w_2(k_0)) \\ & + 2\nu_1 F_1 |A|^2 \partial_{X_2} A (3|m_1|^2 m_1 + 2m_1 |m_2|^2 + \bar{m}_1 m_2^2)) + \text{c.c.} + \mathcal{O}(\varepsilon^5), \end{aligned}$$

and for  $\text{Res}_{\text{nl},2}$  we simply have to change the indices of the components of  $\mathbf{m}$  and  $\partial_k \mathbf{w}(k_0)$  in  $\text{Res}_{\text{nl},2}$ .

## C Calculus Lemma

### Lemma C.1

Let  $m_1, m_2 \in \mathbb{N}_0$  with  $m_1 \geq m_2$  and  $m_1 \geq 2$  and let  $J \subset \mathbb{R}$  be an interval.

(i) Let  $j \in \{0, \dots, m_1\}$ ,  $f \in \mathcal{H}^{m_1-j}(\mathbb{R}^2)$  and  $g \in \mathcal{H}^j(\mathbb{R}^2)$ . Then  $fg \in L^2(\mathbb{R}^2)$  and

$$\|fg\|_{L^2(\mathbb{R}^2)} \leq C \|f\|_{\mathcal{H}^{m_1-j}(\mathbb{R}^2)} \|g\|_{\mathcal{H}^j(\mathbb{R}^2)}.$$

(ii) Let  $f \in \mathcal{H}^{m_1}(\mathbb{R}^2)$  and  $g \in \mathcal{H}^{m_2}(\mathbb{R}^2)$ . Then  $fg \in \mathcal{H}^{m_2}(\mathbb{R}^2)$  and

$$\|fg\|_{\mathcal{H}^{m_2}(\mathbb{R}^2)} \leq C \|f\|_{\mathcal{H}^{m_1}(\mathbb{R}^2)} \|g\|_{\mathcal{H}^{m_2}(\mathbb{R}^2)}.$$

(iii) Let  $f \in \mathcal{F}_0^{m_1,1}(\mathbb{R}^2)$  and  $g \in \mathcal{H}^{m_2}(\mathbb{R}^2)$ . Then  $fg \in \mathcal{H}^{m_2}(\mathbb{R}^2)$  and

$$\|fg\|_{\mathcal{H}^{m_2}(\mathbb{R}^2)} \leq C \|f\|_{\mathcal{F}_0^{m_1,1}(\mathbb{R}^2)} \|g\|_{\mathcal{H}^{m_2}(\mathbb{R}^2)}.$$

(iv) Let  $f \in \mathcal{F}^{m_1,1}(\mathbb{R}^2 \times J)$  and  $g \in \mathcal{G}^{m_2}(\mathbb{R}^2 \times J)$ . Then  $fg \in \mathcal{G}^{m_2}(\mathbb{R}^2 \times J)$  and

$$\|fg\|_{\mathcal{G}^{m_2}(\mathbb{R}^2 \times J)} \leq C \|f\|_{\mathcal{F}^{m_1,1}(\mathbb{R}^2 \times J)} \|g\|_{\mathcal{G}^{m_2}(\mathbb{R}^2 \times J)}.$$

(v) Let  $f \in \mathcal{F}_0^{m_1,1}(\mathbb{R}^2)$  and  $g \in \mathcal{F}_0^{m_2,1}(\mathbb{R}^2)$ . Then  $fg \in \mathcal{F}_0^{m_2,1}(\mathbb{R}^2)$  and

$$\|fg\|_{\mathcal{F}_0^{m_2,1}(\mathbb{R}^2)} \leq C \|f\|_{\mathcal{F}_0^{m_1,1}(\mathbb{R}^2)} \|g\|_{\mathcal{F}_0^{m_2,1}(\mathbb{R}^2)}.$$

PROOF: The proofs can be done analogously to Lemma 2.22 in [27]. □

## Acknowledgements

Tomas Dohnal and Daniel P. Tietz acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project-ID DO1467/4-1. Roland Schnaubelt acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 258734477 - SFB 1173.

## References

- [1] G. Agrawal. *Nonlinear Fiber Optics*. Elsevier Science, 2013.
- [2] A. Babin and A. Figotin. Nonlinear photonic crystals: IV. Nonlinear Schrödinger equation regime. *Waves Random Complex Media*, 15(2):145–228, 2005.
- [3] A. Ben-Artzi and I.I Gohberg. Dichotomy of systems and invertibility of linear ordinary differential operators. In *Time-Variant Systems and Interpolation*, volume 56 of *Oper. Theory Adv. Appl.*, pages 90–119. Birkhäuser, Basel, 1992.
- [4] M. Brown, T. Dohnal, M. Plum, and I. Wood. Spectrum of the Maxwell Equations for the Flat Interface between Dispersive Media. 2022. in preparation.
- [5] K. Busch, G. Schneider, L. Tkeshelashvili, and H. Uecker. Justification of the nonlinear Schrödinger equation in spatially periodic media. *Z. Angew. Math. Phys.*, 57(6):905–939, 2006.
- [6] W. Coppel. *Dichotomies in Stability Theory*. Springer-Verlag, Berlin, New York, 1978.
- [7] S. Crutcher and A. Osei. Derivation of the Effective Nonlinear Schrödinger Equations for Dark and Power Law Spatial Plasmon-Polariton Solitons Using Nano Self-Focusing. *Progress In Electromagnetics Research B*, 29:83–103, 2011.
- [8] R. Dautray and J. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 1 Physical Origins and Classical Methods*. Springer-Verlag, 1990.
- [9] A. Davoyan, I. Shadrivov, and Y. Kivshar. Self-focusing and spatial plasmon-polariton solitons. *Opt. Express*, 17(24):21732–21737, Nov 2009.
- [10] M. Davydova, D. Dodonov, A. Kalish, V. Belotelov, and A. Zvezdin. Schrödinger plasmon solitons in kerr nonlinear heterostructures with magnetic manipulation. *Opt. Lett.*, 40(23):5439–5442, 2015.
- [11] T. Dohnal, G. Romani, and D. Tietz. A quasilinear transmission problem with application to maxwell equations with a divergence-free D-field. *J. Math. Anal. Appl.*, 511(1):126067, 2022.
- [12] T. Dohnal and D. Rudolf. Nls approximation for wavepackets in periodic cubically nonlinear wave problems in  $\mathbb{R}^d$ . *Appl. Anal.*, 99(10):1685–1723, 2020.
- [13] M. Eller. On symmetric hyperbolic boundary problems with nonhomogeneous conservative boundary conditions. *SIAM J. Math. Anal.*, 44(3):1925–1949, 2012.
- [14] C. Jamois, R. Wehrspohn, L. Andreani, C. Hermann, O. Hess, and U. Gösele. Silicon-based two-dimensional photonic crystal waveguides. *Photonics Nanostructures: Fundam. Appl.*, 1(1):1–13, 2003.
- [15] L. Kalyakin. Long wave asymptotics. integrable equations as asymptotic limits of non-linear systems. *Russian Math. Surveys*, 44(1):3, 1989.
- [16] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.
- [17] M. Kauranen and A. Zayats. Nonlinear plasmonics. *Nature Photonics*, 6(11):737–748, Nov 2012.
- [18] P. Kirrmann, G. Schneider, and A. Mielke. The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. Roy. Soc. Edinburgh A*, 122(1-2):85–91, 1992.
- [19] Y. Kivshar and S. Mingaleev. Nonlinear photonic crystals: waveguides, all-optical switching, and solitons. In *Nonlinear Optics: Materials, Fundamentals and Applications*, page ThD1. Optica Publishing Group, 2002.

- [20] V. Lescaudet and G. Schneider. Diffractive optics with harmonic radiation in 2d nonlinear photonic crystal waveguide. *Z. Angew. Math. Phys.*, 63(3):401–427, Jun 2012.
- [21] G. Li and S. Seshadri. Weakly nonlinear surface polariton. *J. Opt. Soc. Am. B*, 6(6):1125–1137, 1989.
- [22] A. Marini and D. Skryabin. Ginzburg-Landau equation bound to the metal-dielectric interface and transverse nonlinear optics with amplified plasmon polaritons. *Phys. Rev. A*, 81:033850, 2010.
- [23] R. Schnaubelt and M. Spitz. Local wellposedness of quasilinear Maxwell equations with conservative interface conditions. *Commun. Math. Sci.*, accepted for publication, 2022.
- [24] G. Schneider. Validity and limitation of the Newell-Whitehead equation. *Math. Nachr.*, 176(1):249–263, 1995.
- [25] G. Schneider, D. Sunny, and D. Zimmermann. The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions. *J. Dynam. Differential Equations*, 27(3):1077–1099, 2015.
- [26] G. Schneider and H. Uecker. Existence and stability of modulating pulse solutions in Maxwell’s equations describing nonlinear optics. *Z. Angew. Math. Phys.*, 54(4):677–712, 2003.
- [27] M. Spitz. *Local wellposedness of nonlinear Maxwell equations*. PhD thesis, Karlsruhe Institute of Technology, Karlsruhe, 2017.
- [28] T. Tao. *Nonlinear Dispersive Equations*. American Mathematical Society, Providence, RI, 2006.
- [29] Z. Wang, Y. Chong, J. Joannopoulos, and M. Soljačić. Reflection-free one-way edge modes in a gyromagnetic photonic crystal. *Phys. Rev. Lett.*, 100:013905, 2008.