Scaling relations for auxin waves

Bente Hilde Bakker[1] · Timothy E. Faver[2] · Hermen Jan Hupkes[1] ·
Roeland M. H. Merks[3] · Jelle van der Voort[1]

Abstract

We analyze an 'up-the-gradient' model for the formation of transport channels of the phytohormone auxin, through auxin-mediated polarization of the PIN1 auxin transporter. We show that this model admits a family of travelling wave solutions that is parameterized by the height of the auxin-pulse. We uncover scaling relations for the speed and width of these waves and verify these rigorous results with numerical computations. In addition, we provide explicit expressions for the leading-order wave profiles, which allows the influence of the biological parameters in the problem to be readily identified. Our proofs are based on a generalization of the scaling principle developed by Friesecke and Pego to construct pulse solutions to the classic Fermi–Pasta–Ulam–Tsingou model, which describes a one-dimensional chain of coupled nonlinear springs.

✉ Hermen Jan Hupkes
hhupkes@math.leidenuniv.nl

Bente Hilde Bakker
b.h.bakker@math.leidenuniv.nl

Timothy E. Faver
tfaver1@kennesaw.edu

Roeland M. H. Merks
merksrmh@math.leidenuniv.nl

Jelle van der Voort
jelvoort@live.nl

[1] Mathematical Institute, Universiteit Leiden, P.O. Box 9512, 2300 RA Leiden, The Netherlands

[2] Department of Mathematics, Kennesaw State University, 850 Polytechnic Lane, MD #9085, Marietta, GA 30060, USA

[3] Mathematical Institute and Institute of Biology Leiden, Universiteit Leiden, P.O. Box 9512, 2300 RA Leiden, The Netherlands

Mathematics Subject Classification Primary 34A33 · 92C37; Secondary 34K26

1 Introduction

1.1 Polar auxin transport

The phytohormone auxin is a central player in practically all aspects of the development and growth of plants, for example in phyllotaxis, root development and the initiation of lateral roots, the formation of vascular tissues in stems, the patterning of leaf veins, and flower development (Paque and Weijers 2016). The pattern formation principles underlying these developmental mechanisms have been uncovered to a large part through an intensive cross-talk between experimental approaches and mathematical modeling (Shi and Vernoux 2018; Autran et al. 2021; Cieslak et al. 2021). Auxin is transported between cells and between cells and the cell walls both through diffusion and through transport proteins that are localized at the plasma membrane (PM). Some of these transport proteins, mostly notably several members of the PIN-FORMED family including PIN1 (Adamowski and Friml 2015) are distributed in a polarised manner inside the cells. Such polarised localisation of PINs is coordinated in plant tissue, leading to a directed transport of auxin through plant tissues in a mechanism called polar auxin transport (PAT) (Adamowski and Friml 2015). For example, in fully developed seed plants, auxin is synthesized in leaves, then is transported through the central tissues of the stem and the root towards the root tips, where it redirected along the superficial tissues of the root back to towards the stem and recycled towards the internal tissues of the root (Adamowski and Friml 2015).

Despite new details being uncovered incessantly (see e.g. Verna et al. 2019; Hajný et al. 2020), it is still incompletely understood what mechanisms drive the polarization of PINs inside cells and the coordinated polarization among adjacent cells. In a series of classical experiments, Sachs applied artificial auxin to bean plants, and observed that these become the source of new vascular tissue that then joins the existing vasculature; see e.g. Sachs (1975) and the review (Hajný et al. 2022). These initial observations, together with the discovery of PIN1 and subsequently discovered members of the PIN-FORMED protein family suggested that auxin drives the polarization of its own transporters, and hence the direction of its own transport (reviewed in Merks et al. 2007; Hajný et al. 2022). Initial models aimed to explain the formation of transport channels as observed in Sachs' experiments. These models therefore assumed that the rate of auxin flux from cell to cell further polarised auxin transport. This positive feedback led to the self-organised formation of auxin transport channels in a process called auxin canalisation. When it was realised that auxin accumulations mark the formation of new leaves at the shoot apex, an alternative model was proposed, in which cells polarised towards the locally increased concentrations of auxin, thus forming self-organised accumulation of auxin (Reinhardt et al. 2003). Mathematical models of the self-organisation of polar auxin transport therefore follow these two broad categories. 'With-the-gradient' models formalise the canalisation hypothesis and assume that the rate of cell polarisation depends on the auxin *flux* towards the relevant neighbour (Mitchison 1980, 1981; Rolland-Lagan and Prusinkiewicz 2005;

Rolland-Lagan 2008). 'Up-the-gradient' models assume that PIN polarizes in the direction of neighbouring cells at a rate that positively depends on the auxin *concentration* in that neighbour (Jönsson et al. 2006; Smith et al. 2006). Attempts to reconcile these two seemingly contradicting ideas have followed two broad approaches. The first approach proposed that with-the-gradient and up-the-gradient models act at different positions of the plant or at different stages during development. For example Bayer et al. (2009) proposed that the up-the-gradient model act at superficial tissue layers of the shoot apical meristem where it forms auxin accumulation points leading to the initial of new leaves. The deeper tissue layers could follow the with-the-gradient model channeling auxin away from the auxin accumulation point towards the vascular tissues (Bayer et al. 2009). A similar approach was recently taken to explain the leaf venation patterning in combination with auxin convergence at the edge of the leaf primordium (Holloway and Wenzel 2021). The second approach looked for variants of the with-the-gradient or up-the-gradient models that could explain both auxin canalisation and auxin accumulation depending on the parameter settings. In this line of reasoning Walker et al. have proposed a with-the-gradient hypothesis for phyllotaxis (Walke et al. 2013), whereas one of us has proposed an up-the-gradient hypothesis for canalisation (Merks et al. 2007).

More recent analyses of the role of auxin and PINs in the formation of leaf veins (Verna et al. 2019) put the key role of a feedback between auxin signaling and the polar localisation of PINs into question, and therefore the validity of canalisation hypothesis or its alternatives including the traveling-wave hypothesis (Merks et al. 2007) for formation of vascular tissues. In particular, quadruple mutants strongly reducing functionality of all plasma-membrane-localised PINs, i.e., of all PINs that are responsible for PAT in the leaf veins, show relatively mild venation pattern phenotypes. Further knock-out of PIN6 and PIN8, expressed in the leaf veins but not localised in the PM, thus excluding a role of these PINs in PAT, led to further defects in leaf venation patterning (Verna et al. 2019) identical to those due to a chemical block of auxin transport. Nevertheless, in these mutants the polar ordering of the cells in the vasculature stays intact and supernumerary veins are induced by exogenous application of auxin, showing that auxin can induce veins in absence of polar transport. Using further mutations of auxin sensing proteins, it was found that this PAT-independent vein formation requires auxin sensing and the activity of GNOM, a protein regulating the constitutive recycling of PM-localised proteins, including PINs. How the available mathematical models of auxin-regulated patterning in plants will need to be updated or rejected is a topic of ongoing investigation, but what seems clear at this moment is that such models must involve auxin sensing and coordination of cell polarisation possibly through polar transport of other small chemicals besides auxine [e.g., acidification of the cell walls (Fendrych et al. 2016)], facilitated diffusion (Mitchison 1980) or coordination of polarity through other means such as mechanical signaling as studied in mathematical models of phyllotaxis (Julien et al. 2019) and leaf venation patterning (Kneuper et al. 2020).

In this paper we formally analyse an existing up-the-gradient model for establishment of polar auxin transport during leaf venation patterning (Jönsson et al. 2006; Heisler and Jonsson 2006; Merks et al. 2007). Although this model is a strong oversimplification of the experimental state-of-the-art, which in part invalidates it, it includes
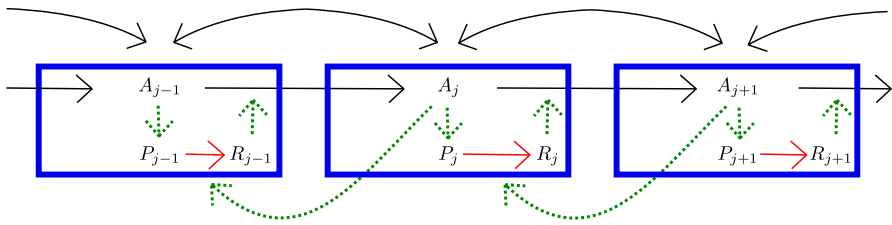
Fig. 1 Schematic representation of the model (1.1). Black arrows represent transport, red arrows describe polarization and the green dashed arrows indication promotion. In particular, the PIN1 polarization rate correlates positively with the neighbouring auxin concentration, making this a model of 'up-the-gradient' type

(1) auxin sensing, (2) polar transport, and (3) constitutive recycling, and thus likely contains key elements of updated, future models while still retaining the simplicity required for mathematical analysis. Thus, despite clear discrepancies of recent experimental insights with both the up-the-gradient and with-the-gradient models, the insights obtained in a formal analysis as well as the mathematical approaches developed in this work will likely apply to future, updated models of auxin-regulated patterning in plants.

1.2 Mathematical motivation

In order to distinguish between the available phenomenological models of auxin-driven pattern formation and the general developmental principles that they represent, mathematical insight into the models' structure and the models' solutions will be crucial. This will help pinpoint key differences between the model structures and may uncover potential structural instabilities in the models upon which evolution may have acted, so as to produce new developmental patterning modules (Benítez et al. 2018). From the mathematical side, almost all previous studies have focused on the types of patterns that can be generated by different models once the transitory dynamics have died out. An important example is the study by Van Berkel and coworkers (van Berkel et al. 2013), where a number of models for polar auxin transport are recast into a common mathematical framework that allows them to be compared. A steady state analysis for a general class of active transport models can be found in Draelants et al. (2015), using advanced tools such as snaking from the field of bifurcation theory. Both periodic and stationary patterns are examined in Allen and Ptashnyk (2020), where the authors consider an extended with-the-gradient model. Haskovec and his coworkers derive local and global existence results together with an appropriate continuum limit for their graph-based diffusion model in Haskovec et al. (2019).

Important qualitative examples of the up-the-gradient model are the formation of regularly spaced auxin maximums that lead to the growth of new leaves, as well as the formation of auxin channels that have been hypothesized to precede the formation of veins. Our goal here is to move beyond the well-studied equilibrium settings above and focus instead on understanding the dynamical behavior that leads to these patterns. In particular, we provide a rigorous framework to study a class of wave solutions that

underpin the dynamical behaviour associated to up-the-gradient models. Ultimately, we hope that this analytic approach will provide an additional lens through which models of PAT can be examined and compared.

1.3 The model

Inspired by Jönsson et al. (2006), Heisler and Jonsson (2006) and Merks et al. (2007), the system we will study is given by

$$
\begin{cases}
\dot{A}_j = T_{\text{act}} \left(R_{j-1} \frac{A_{j-1}}{k_a + A_{j-1}} - R_j \frac{A_j}{k_a + A_j} \right) + T_{\text{diff}}(A_{j+1} - 2A_j + A_{j-1}), \\[2ex]
\dot{P}_j = -k_1 \frac{A_{j+1}}{k_r + A_{j+1}} \left(\frac{P_j}{k_m + P_j} \right) + \alpha A_j, \\[2ex]
\dot{R}_j = k_1 \frac{A_{j+1}}{k_r + A_{j+1}} \left(\frac{P_j}{k_m + P_j} \right),
\end{cases}
\tag{1.1}
$$

posed on the one-dimensional lattice $j \in \mathbb{Z}$; see Fig. 1. The variable $A_j(t)$ denotes the auxin concentration in cell $j \in \mathbb{Z}$, while $P_j(t)$ and $R_j(t)$ represent the unpolarized respectively right-polarized PIN1 in this cell. PIN1 is the PIN-variant that is believed to play a central role during auxin-based pattern formation in the shoot apical meristem and during leaf venation patterning (Reinhardt et al. 2003; Jönsson et al. 2006; Smith et al. 2006; Scarpella et al. 2006; Verna et al. 2019), and we therefore consider PIN1 here. However, note that the general structure of this model would apply to other polarised transporter proteins with similar behavior.

The parameters appearing in the problem are all strictly positive and labelled in the same manner as in Merks et al. (2007).[1] In particular, T_{act} and T_{diff} denote the strengths of the active PIN1-induced rightward auxin transport and its diffusive counterpart, respectively. Unpolarized PIN1 is formed in the presence of auxin at a rate α, while k_1 denotes the polarization rate. Finally, k_a, k_r, and k_m are the Michaelis constants associated to the active transport of auxin and the polarization of PIN1, which depends on the auxin-concentration in the right-hand neighbouring cell. In particular, this model is of 'up-the-gradient' type.

The main difference compared to Merks et al. (2007) is that we are neglecting the presence of left-polarized PIN1 and have set the decay and depolarization rates of PIN1 to zero. Although this step of course imposes a pre-existing polarity on the system, we need to do this for technical reasons that we explain in the sequel. For now we simply point out that we wish to focus our attention on the dynamics of rightward auxin propagation, which takes place on timescales that are much faster than these decay and depolarization processes, and that the results will give novel insight into the full problem.

We will look for solutions of the special type

$$
(A_j, P_j, R_j)(t) = (\phi_A, \phi_P, \phi_R)(j - ct),
\tag{1.2}
$$

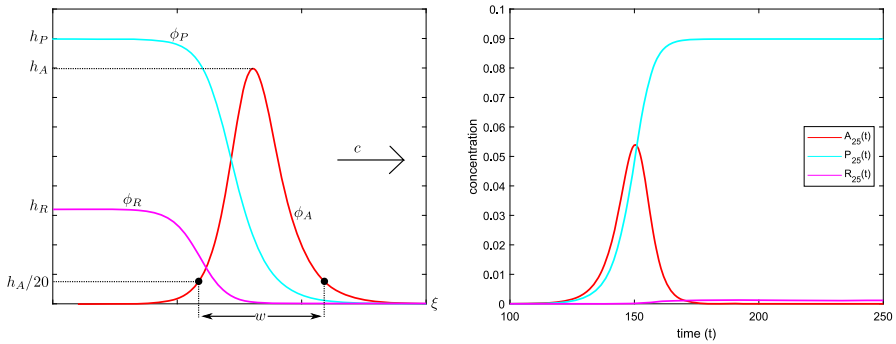[1] For presentation purposes, the parameters L and r appearing in Merks et al. (2007) have been set to unity.

Fig. 2 Left: cartoon of the waveprofiles (ϕ_A, ϕ_P, ϕ_R), illustrating the definition of the width w of the auxin-pulse and the limits (1.3). Right: numerical simulation of an auxin pulse passing through cell 25, leaving a residue of (polarized) PIN1. We used the procedure described in Sect. 1.4, with $A_1(0) = A_\diamond = 0.15$. The remaining parameters were fixed as $T_{\text{act}} = 800$, $T_{\text{diff}} = 0.15$, $k_a = 1$, $k_m = k_r = 100$, $k_1 = 200$ and $\alpha = 0.1$

with $c > 0$, in which we impose the limits

$$\lim_{\xi \to -\infty} \phi_A(\xi) = 0, \qquad \lim_{\xi \to \infty} (\phi_A, \phi_P, \phi_R)(\xi) = 0; \qquad (1.3)$$

see Fig. 2. From a modelling perspective, such solutions represent a pulse of auxin that moves to the right through a one-dimensional row of cells. Ahead of the wave the cells are clear of both polarized and unpolarized PIN, but behind the wavefront a residual amount of PIN is left in the cells, representing the coordinated polarisation of the tissue.

In reality these residues start to depolarize and decay, which can be included by adding linear decay terms to (1.1). This leads to the expanded system

$$\begin{cases} \dot{A}_j = T_{\text{act}} \left(R_{j-1} \frac{A_{j-1}}{k_a + A_{j-1}} - R_j \frac{A_j}{k_a + A_j} \right) + T_{\text{diff}}(A_{j+1} - 2A_j + A_{j-1}), \\[3mm] \dot{P}_j = -k_1 \frac{A_{j+1}}{k_r + A_{j+1}} \left(\frac{P_j}{k_m + P_j} \right) + \alpha A_j + k_2 R_j - \delta P_j, \\[3mm] \dot{R}_j = k_1 \frac{A_{j+1}}{k_r + A_{j+1}} \left(\frac{P_j}{k_m + P_j} \right) - k_2 R_j, \end{cases} \qquad (1.4)$$

in which the positive parameters δ and k_2 represent the decay and depolarization rate of PIN1, respectively. Mathematically, these terms can be included into our framework provided that the parameters δ and k_2 are small compared to the amplitude of the pulses, but we do not pursue this level of generality in the current paper for presentational clarity. Note in any case that in Merks et al. (2007) these parameters were chosen to be orders of magnitude smaller than α and k_1.

Travelling waves have played a fundamental role in the analysis of many spatially discrete systems (Kevrekidis 2011; Mallet-Paret 1999; Chen et al. 2008; Hupkes and Sandstede 2010; Keener 1987). They can be seen as a lossless mechanism to transport
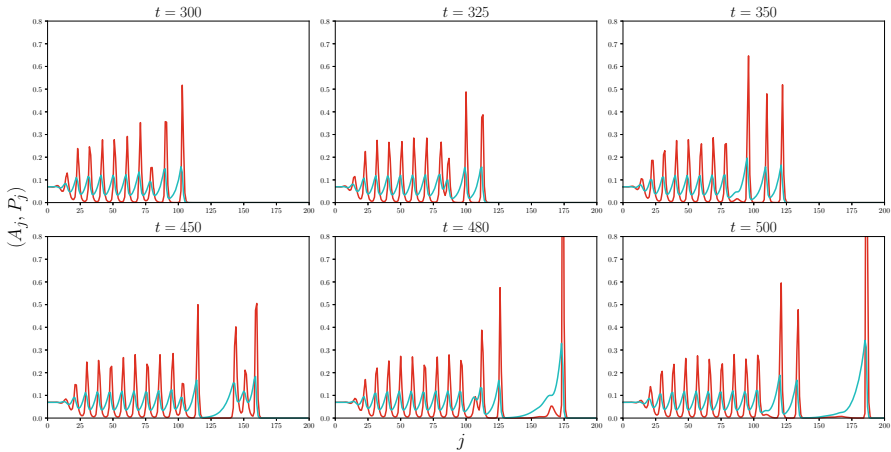
Fig. 3 Six snapshots of a wavetrain simulation for the expanded system (1.4). Higher pulses travel faster than lower pulses, in correspondence with the scaling relations (1.7). These speed differences lead to merge events where even higher pulses are formed, which detach from the bulk. We used the procedure described in Sect. 1.4, taking $A_1(0) = A_\diamond = 0.0$ but adding 0.025 to $\dot{A}_1(t)$ to simulate a constant auxin influx at the left boundary. We picked $\delta = 0.1$ and $k_2 = 0.2$, leaving the remaining parameters from Fig. 2 unchanged. The full simulation can be found in supplementary video S1

matter or energy over arbitrary distances. As such, they are interesting in their own right, but they can also be viewed as building blocks to describe more complicated behaviour of nonlinear systems (Aronson and Weinberger 1975, 1978). In the present case for example, one can construct wavetrain solutions to (1.4) by adding a persistent auxin source; see Fig. 3 and Supplementary Video S1. Initially, these solutions can be seen in an approximate sense as a concatenation of the individual auxin pulses that we consider here (Moser 2021). As a consequence of the amplitude variations, small speed differences occur between these pulses which leads to highly interesting collision processes. Due to this type of versatility, travelling waves play an important role in many applications and have been extensively studied in a variety of settings (Sandstede 2002; Kevrekidis 2011; Hochstrasser et al. 1989; Jones et al. 1991).

1.4 Main results

Our goal will be to obtain quantitative scaling information concerning the speed and shape of these waves. In particular, we will show rigorously that (1.1) admits a family of travelling wave solutions that are parameterized by the amplitude of the auxin-pulse. In addition, we show that the speed and width of these waves scale with this amplitude via a fractional power law. We state our results in full technical detail in Theorem 6.3 below.

More precisely, we provide an explicit triplet of functions $(\phi_A^*, \phi_P^*, \phi_R^*)$ that satisfy the limits (1.3) and construct solutions to (1.1) of the form

$$
\begin{aligned}
\left(A_j, P_j, R_j\right)(t) = {}& \left(\epsilon\phi_A^*, \epsilon^{1/5}\phi_P^*, \epsilon^{2/5}\phi_R^*\right)\left(\epsilon^{2/5}(j - c_*\epsilon^{2/5}t)\right) \\
& + \left(\mathcal{O}(\epsilon^{17/15}), \mathcal{O}(\epsilon^{1/3}), \mathcal{O}(\epsilon^{3/5})\right),
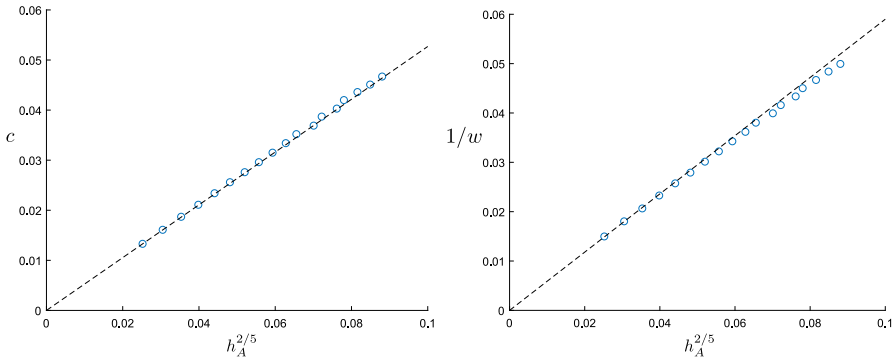\end{aligned}
\tag{1.5}
$$

Fig. 4 Scaling behaviour of the wavespeed c (left) and the auxin width w (right) against the height h_A of the auxin pulse. The dashed lines represent the explicit predictions (1.7). The circles arise from numerical simulations, following the procedure described in Sect. 1.4 with several different values for $A_\diamond$. The other parameters were chosen as in Fig. 2

for a constant c_*, which we state exactly in (1.8). Here the limiting profile ϕ_A^* is scaled in such a way that $\|\phi_A^*\|_{L^\infty} = 1$. Upon introducing the heights[2]

$$(h_A, h_P, h_R) = \left(\|A\|_\infty, \|P\|_\infty, \|R\|_\infty \right) \tag{1.6}$$

associated to the three components of our waves, this choice ensures that the auxin-height h_A is equal to the parameter $\epsilon > 0$ at leading order. In particular, comparing this to (1.2) we uncover the leading order scaling relations

$$c \sim c_* h_A^{2/5}, \qquad w \sim w_* h_A^{-2/5}, \qquad h_P \sim h_P^* h_A^{1/5}, \qquad h_R \sim h_R^* h_A^{2/5} \tag{1.7}$$

for the speed c, width[3] w and heights of the wave. Here the constant w_* denotes the width of the limiting profile ϕ_A^*, while the other constants are given explicitly by

$$c_* = \left(\frac{9\alpha k_1 T_{\mathrm{act}} T_{\mathrm{diff}}^2}{8 k_a k_m k_r} \right)^{1/5},$$

$$h_P^* = \sqrt{6} \left(\frac{9\alpha^6 k_a^4 k_m^4 k_r^4 T_{\mathrm{diff}}^2}{8 k_1^4 T_{\mathrm{act}}^4} \right)^{1/10},$$

$$h_R^* = 3 \left(\frac{9\alpha k_a^4 k_1 T_{\mathrm{diff}}^2}{8 k_r k_m T_{\mathrm{act}}^4} \right)^{1/5}. \tag{1.8}$$

In particular, for a fixed height of the auxin-pulse our results state that the speed and residual PIN1 will increase as the PIN1-production parameter $\alpha > 0$ is increased.

---

[2] Here we use the abbreviation $\|A\|_\infty = \sup_{j,t} |A_j(t)|$ and its analogues for P and R.

[3] We define the width of the auxin pulse as the distance between the two points where the pulse attains 5% of its maximum value.
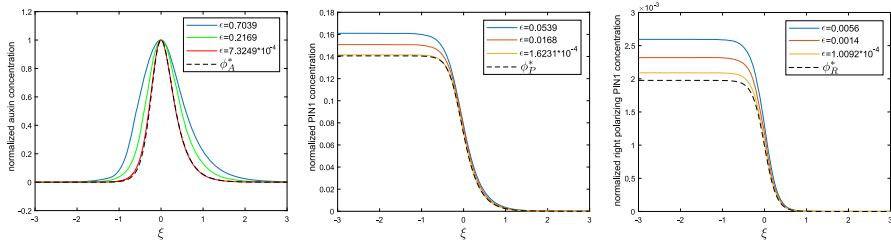
Fig. 5 Convergence of the (scaled) profiles ϕ_A (left), ϕ_P (center) and ϕ_R (right) to their limits $(\phi_A^*, \phi_P^*, \phi_R^*)$. To perform the scalings, we wrote $h_A = \|\phi_A\|_{L^\infty}$, compressed space by a factor of $h_A^{2/5}$ and divided the three profiles by the respective factors $(h_A, h_A^{1/5}, h_A^{2/5})$, in line with the relations (1.7)

Although our proof requires the parameter $\epsilon > 0$ and hence the amplitude of the auxin-pulses to be small, this branch of solutions continues to exist well beyond this asymptotic regime. Indeed, we numerically confirmed the existence (and stability) of these waves by a direct simulation of (1.1) on a row of cells $j \in \{1, \dots 500\}$, initialized with $A_j(0) = P_j(0) = R_j(0) = 0$ for $2 \le j \le 500$, together with $P_1(0) = R_1(0) = 0$ and $A_1(0) = A_\diamond$ for some $A_\diamond > 0$ that we varied between simulations. In order to close the system, we used the Neumann-type condition $A_0(t) = A_1(t)$ on the left-boundary, together with $R_0(t) = 0$ and a sink condition $A_{501}(t) = 0$ on the right. An example of such a simulation can be found in Fig. 2 (right). By varying the initial auxin concentration $A_\diamond$, we were able to generate waves with a range of amplitudes. We subsequently numerically computed the speed and width of these waves, which allowed us to confirm the leading order behaviour (1.7); see Fig. 4. In addition, we verified the convergence to the limiting profiles $(\phi_A^*, \phi_P^*, \phi_R^*)$ by comparing the appropriately rescaled numerical waveprofiles; see Fig. 5.

1.5 Cross-diffusion

From a mathematical perspective, the problem (1.1) is interesting due to its interpretation as a so-called cross-diffusion problem, where the transport coefficient of one component is influenced by one of the other components. Work in this area was stimulated by developments in the modeling of bacterial cell membranes (Shih et al. 2019) and biofilms (Emerenini et al. 2015), where self-organization of biological molecules plays an important role. In the continuum regime, such problems are tough to analyze on account of potential degeneracies in the coefficients. The well-posedness of the underlying problem was analyzed in Sonner et al. (2011), while a numerical method for such problems was developed in Ghasemi et al. (2018).

The key phenomenological assumption behind such models is that particles behave differently when they are isolated compared to when they are part of a cluster. A simplified agent-based approach to capture this mechanism can be found in Johnston et al. (2017), which reduces naturally to a scalar PDE with nonlinear diffusion in the continuum limit. After adding a small regularization term, it is possible to use geometric singular perturbation theory to show that this PDE admits travelling wave

solutions (Li et al. 2021). In this setting, the steepness of the wavefronts provides the necessary scale-separation required for rigorous results.

Our approach in this paper proceeds along entirely different lines, using the amplitude of the auxin pulse as a small continuation parameter to construct a family of travelling wave solutions to (1.1). The key insight is that one can extract an effective limiting system by scaling the width and speed of the wave in an appropriate fashion and sending the amplitude to zero. By means of a fixed-point analysis one can show in a rigorous fashion that solutions to this limiting system can be continued to form a family of solutions to the full system.

1.6 Relation to FPUT pulses

Our technique is a generalization of the approach developed by Friesecke and Pego (1999) to construct small-amplitude travelling pulse solutions to the Fermi–Pasta–Ulam–Tsingou (FPUT) problem (Fermi et al. 1955; Dauxois 2008)

$$\ddot{x}_j = F(x_{j+1} - x_j) - F(x_j - x_{j-1}), \qquad j \in \mathbb{Z}. \tag{1.9}$$

This models an infinite, one-dimensional chain of particles that can only move horizontally and are connected to their nearest neighbours by springs. These springs transmit a force

$$F(r) = r + r^2 \tag{1.10}$$

that hence depends nonlinearly on the relative distance r between neighbouring particles; see Friesecke and Pego (1999), Herrmann and Matthies (2015) and Pankov (2005) for the impact of other choices. The FPUT system is well-established as a fundamental model to study the propagation of disturbances through spatially discrete systems, such as granular media, artificial metamaterials, DNA strands, and electrical transmission lines (Brillouin 1953; Kevrekidis 2011).

Looking for a travelling wave in the relative displacement coordinates, one introduces an Ansatz of the form

$$x_{j+1}(t) - x_j(t) = \phi(j - \sigma t), \tag{1.11}$$

which leads to the scalar functional differential equation of mixed type (MFDE)

$$\sigma^2 \phi''(\xi) = F\big(\phi(\xi + 1)\big) - 2F\big(\phi(\xi)\big) + F\big(\phi(\xi - 1)\big). \tag{1.12}$$

Following the classic papers by Friesecke in combination with Wattis (Friesecke and Wattis 1994) and Pego (Friesecke and Pego 1999, 2002, 2004a, b), we introduce the scaling

$$\phi(\xi) = \epsilon^2 \varphi_\epsilon(\epsilon \xi) \tag{1.13}$$

and write $\sigma = \sigma_\epsilon$, which transforms (1.12) into the MFDE

$$\sigma_\epsilon^2 \epsilon^2 \varphi_\epsilon'' = (S^\epsilon + S^{-\epsilon} - 2)\big[\varphi_\epsilon + \epsilon^2 \varphi_\epsilon^2\big]. \tag{1.14}$$

Here the shift operator S^d acts as

$$(S^d f)(\xi) = f(\xi + d) \tag{1.15}$$

for any $d \in \mathbb{R}$. Since the symbol $S^\epsilon + S^{-\epsilon} - 2$ represents a discrete Laplacian, we can interpret (1.14) as a wave equation with a nonlinear diffusion term. To some extent, this clarifies the link with our original problem (1.1) and the discussion above.

Applying the Fourier transform to (1.14) with k as the frequency variable, we arrive at

$$
\begin{aligned}
- \sigma_\epsilon^2 \epsilon^2 k^2 \widehat{\varphi_\epsilon}(k) &= 2(\cos(\epsilon k) - 1)\big[\widehat{\varphi_\epsilon} + \epsilon^2 \widehat{\varphi_\epsilon^2}\big](k) \\
&= -4 \sin^2(\epsilon k/2)\big[\widehat{\varphi_\epsilon} + \epsilon^2 \widehat{\varphi_\epsilon^2}\big](k).
\end{aligned} \tag{1.16}
$$

Upon introducing the symbol

$$\widetilde{\mathcal{M}}_{\mathrm{FPUT}}^{(\epsilon)}(k) = \frac{4\epsilon^2 \sin^2(\epsilon k/2)}{\sigma_\epsilon^2 \epsilon^2 k^2 - 4\sin^2(\epsilon k/2)}, \tag{1.17}$$

this can be recast into the compact form

$$\widehat{\varphi_\epsilon}(k) = \widetilde{\mathcal{M}}_{\mathrm{FPUT}}^{(\epsilon)}(k)\widehat{\varphi_\epsilon^2}(k). \tag{1.18}$$

Upon choosing the speed

$$\sigma_\epsilon = 1 + \frac{\epsilon^2}{3}, \tag{1.19}$$

we can exploit the expansion $\sin^2(z/2) = \frac{1}{4}z^2 - \frac{1}{48}z^4 + O(z^6)$ to obtain the pointwise limit

$$\widetilde{\mathcal{M}}_{\mathrm{FPUT}}^{(\epsilon)}(k) \to \frac{12}{8 + k^2}, \qquad \epsilon \to 0. \tag{1.20}$$

Using the fact that $(8 + k^2)$ is the Fourier symbol for $8 - \partial_\xi^2$, this suggests that the relevant system for φ_ϵ in the formal $\epsilon \to 0$ limit is given by

$$8\varphi_* - \varphi_*'' = 12\varphi_*^2, \tag{1.21}$$

which has the nontrivial even solution

$$\varphi_*(\xi) = \mathrm{sech}^2(\sqrt{2}\xi). \tag{1.22}$$

By casting the problem in an appropriate functional analytic framework, one can show that this explicit solution φ_* can be continued to yield solutions φ_ϵ to (1.14) for small $\epsilon > 0$. In this fashion, one establishes the existence of a family of pulse solutions (Friesecke and Pego 1999)

$$x_{j+1}(t) - x_j(t) = \epsilon^2 \operatorname{sech}^2\left(\sqrt{2}\epsilon(j - \sigma_\epsilon t)\right) + \mathcal{O}(\epsilon^4). \tag{1.23}$$

Roughly speaking, the main mathematical contribution in this paper is that we show how this analysis can be generalized to the setting of (1.1). The first main obstacle is that this is a multi-component system, which requires us to explicitly reduce the order before a tractable limit can be obtained. The second main obstacle is that the analysis of our Fourier symbol is considerably more delicate, since in our setting the wavespeed c converges to zero instead of one as $\epsilon \to 0$. Indeed, the denominator of $\widetilde{\mathcal{M}}_{\mathrm{FPUT}}^{(\epsilon)}$ above depends only on the product ϵk, while in our case there is a separate dependence on $\epsilon^2 k$. This introduces a quasi-periodicity into the problem that requires our convergence analysis to carefully distinguish between 'small' values of k and several separate regions of 'large' k.

The third main difference is that we cannot use formal spectral arguments to analyze the limiting linear operator, which in our case is related to the Bernoulli equation. Instead, we apply a direct solution technique using variation-of-constants formulas. On the one hand this is much more explicit, but on the other hand the resulting estimates are rather delicate on account of the custom function spaces involved.

1.7 Discussion

Due to the important organizing role that wave solutions often play in complex systems, scaling information such as (1.7) can be used as the starting point to uncover more general dynamical information concerning models such as (1.1) and related models of polar auxin trasnport. As such, we hope that the ideas we present here will provide a robust analytical tool to analyze different types of models as well. The resulting insights and predictions could help to prioritize competing models on the basis of dynamical experimental observations. Indeed, scaling laws appear to play a role in many aspects of biological systems, such as the structural properties of vascular systems (Razavi et al. 2018), the mass dependence of metabolic rates (West and Brown 2004) and the functional constraints imposed by size (Schmidt-Nielsen and Knut 1984).

Although we have included only right-polarizing PIN in our system, we believe that our techniques can be adapted to cover the full case where also left-polarizing PIN is included. However, the computations rapidly become unwieldy and the limiting system is expected to differ qualitatively. For this reason, we have not chosen to pursue this level of generality in the present paper, as it would only obscure the main ideas behind our framework. One of the main generalizations that we intend to pursue in the future is to study the model in two spatial dimensions. This is motivated by recent numerical observations concerning the formation of auxin channels and their associated PIN polarization under the influence of travelling patterns that are localized in both spatial dimensions (Althuis 2021; Merks et al. 2007).

1.8 Notation

We summarize a few aspects of our (mostly standard) notation.

- If $f = f(X)$ is a differentiable function on $\mathbb{R}$, then we sometimes write $f' = \partial_X[f]$.
- If $\mathcal{X}$ and $\mathcal{Y}$ are normed spaces, then we denote the space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ by $\mathbf{B}(\mathcal{X}, \mathcal{Y})$. We put $\mathbf{B}(\mathcal{X}) := \mathbf{B}(\mathcal{X}, \mathcal{X})$.
- We sometimes abbreviate $\mathbb{R}_+ := (0, \infty)$ and $\mathbb{R}_- := (-\infty, 0)$.

2 The travelling wave problem

2.1 Rewriting the original problem (1.1)

We will reduce the problem (1.1) to a system of equations involving only A_j and P_j, and it will be this resulting system on which we make the long wave-scaled travelling wave Ansatz.

2.1.1 Changes of notation

We begin by rewriting (1.1) in a slightly more compressed manner that also exposes more transparently the leading order terms in the nonlinearities. Let $\delta^{\pm}$ be the left and right difference operators that act on sequences (x_j) in $\mathbb{R}$ via

$$\delta^+ x_j := x_{j+1} - x_j \qquad \text{and} \qquad \delta^- x_j := x_j - x_{j-1}.$$

Next, for $k, x \in \mathbb{R}$ with $k + x \neq 0$ we have

$$\frac{x}{k+x} = \frac{x}{k} - \frac{x^2}{k(k+x)}.$$

We put

$$\mathsf{Q}_1(x, y) := \frac{x^2 y}{k_a + x} \tag{2.1}$$

and compress

$$\tau_1 := \frac{T_{\text{act}}}{k_a} \qquad \text{and} \qquad \tau_2 := T_{\text{diff}} \tag{2.2}$$

to see that our equation for A_j now reads

$$\dot{A}_j = \tau_2 \delta^+ \delta^- A_j - \tau_1 \delta^- (R_j A_j) + \tau_1 \delta^- \mathsf{Q}_1(A_j, R_j).$$

Next, we abbreviate

$$\kappa := \frac{k_1}{k_r k_m} \tag{2.3}$$

and put

$$Q_2(x, y) := \kappa \left(\frac{k_r y + k_m x + xy}{(k_r + x)(k_m + y)} \right) xy \tag{2.4}$$

to see that, the equation for P_j is

$$\dot{P}_j = -\kappa A_{j+1} P_j + \alpha A_j + Q_2(A_{j+1}, P_j).$$

The equation for R_j is updated similarly, and so we have rewritten (1.1) as

$$\begin{cases} \dot{A}_j = \tau_2 \delta^+ \delta^- A_j - \tau_1 \delta^- (R_j A_j) + \tau_1 \delta^- Q_1(A_j, R_j), \\[2mm] \dot{P}_j = -\kappa A_{j+1} P_j + \alpha A_j + Q_2(A_{j+1}, P_j), \\[2mm] \dot{R}_j = \kappa A_{j+1} P_j - Q_2(A_{j+1}, P_j). \end{cases} \tag{2.5}$$

We observe that the equation for R_j depends only on A_{j+1} and P_j and therefore can be solved by direct integration. Before we do that, however, we rewrite the new equation for P_j using Duhamel's formula.

2.1.2 Rewriting the P_j equation

We can view the equation for P_j in (2.5) as a first-order linear differential equation forced by $\alpha A_j + Q_2(A_{j+1}, P_j)$, and so we can solve it via the integrating factor method. For $f, g \in L^1$ and $h \in L^\infty$ we introduce the operators

$$E(f)(s, t) := \exp \left(-\kappa \int_s^t f(\xi) \, d\xi \right), \quad s, t \in \mathbb{R}, \tag{2.6}$$

$$P_1(f, g)(t) := \alpha \int_{-\infty}^t E(f)(s, t) g(s) \, ds, \tag{2.7}$$

and

$$P_2(f, h)(t) := \int_{-\infty}^t E(f)(s, t) Q_2(f(s), h(s)) \, ds. \tag{2.8}$$

Recall from (1.3) that we want P_j to vanish at $-\infty$. The unique solution for P_j in (2.5) that does vanish at $-\infty$ must satisfy

$$P_j(t) = P_1(A_{j+1}, A_j)(t) + P_2(A_{j+1}, P_j)(t).$$

2.1.3 Solving the R_j equation

Since, per (1.3), we want R_j to vanish at $-\infty$, and since we are assuming that each A_j vanishes sufficiently fast at both $\pm\infty$ and P_j vanishes at $-\infty$ and remains bounded at $+\infty$, we may solve for R_j by integrating the third equation in (2.5) from $-\infty$ to t. For $f, g \in L^1$ and $h \in L^\infty$, we define more integral operators:

$$\mathsf{R}_1(f, g)(t) := \kappa\tau_1 \int_{-\infty}^{t} f(s)\mathsf{P}_1(f, g)(s)\, ds, \ t \in \mathbb{R}, \tag{2.9}$$

$$\mathsf{R}_2(f, g, h)(t) := \int_{-\infty}^{t} \big(\kappa f(s)\mathsf{P}_2(f, g)(s) - \mathsf{Q}_2(f(s), \mathsf{P}_1(f, g)(s) \\ +\mathsf{P}_2(f, h)(s))\big)\, ds, \tag{2.10}$$

and

$$\mathsf{R}(f, g, h)(t) := \mathsf{R}_1(f, g)(t) + \mathsf{R}_2(f, g, h)(t). \tag{2.11}$$

We have defined P_1 and P_2 just above, respectively, in (2.7) and (2.8) and Q_2 earlier in (2.4). Then the solution to the third equation in (2.5) that vanishes at $-\infty$ is

$$R_j(t) = \mathsf{R}(A_{j+1}, A_j, P_j)(t) = \mathsf{R}_1(A_{j+1}, A_j)(t) + \mathsf{R}_2(A_{j+1}, A_j, P_j)(t). \tag{2.12}$$

2.1.4 The final system for A_j and P_j

We rewrite (part of) the A_j equation once more to incorporate the new expression for R_j. For $f, g \in L^1$ and $h \in L^\infty$ and $t \in \mathbb{R}$ put

$$\mathsf{N}(f, g, h)(t) := \tau_1\mathsf{Q}_1(g(t), \mathsf{R}(f, g, h)(t)) - \tau_1\mathsf{R}_2(f, g, h)(t)g(t), \tag{2.13}$$

where we defined Q_1 in (2.1). Then A_j must satisfy

$$\dot{A}_j = \tau_2\delta^+\delta^- A_j - \delta^-\big(\mathsf{R}_1(A_{j+1}, A_j)A_j\big) + \delta^-\mathsf{N}(A_{j+1}, A_j, P_j),$$

and so our system for A_j and P_j is now

$$\begin{cases} \dot{A}_j = \tau_2\delta^+\delta^- A_j - \delta^-\big(\mathsf{R}_1(A_{j+1}, A_j)A_j\big) + \delta^-\mathsf{N}(A_{j+1}, A_j, P_j), \\ \\ P_j = \mathsf{P}_1(A_{j+1}, A_j) + \mathsf{P}_2(A_{j+1}, P_j). \end{cases} \tag{2.14}$$

That is, using the formula (2.12) for R_j in terms of A_j and P_j, we can solve (2.5) if we can solve (2.14).

We will make two changes of variables on (2.14). First, in Sect. 2.2, we make a travelling wave Ansatz for A_j and P_j. We reformulate (2.14) for the travelling wave profiles as the system (2.28) below. Then, in Sect. 3.1, we introduce our long wave

Table 1 Summary of notational evolution	Symbol	Use
	R	The original problem (2.14)
	$\widetilde{\mathsf{R}}^c$	The travelling wave problem (2.28)
	$\check{\mathsf{R}}^\epsilon$	The preliminary long wave problem (3.14)
	$\mathcal{R}^\nu$	The final long wave problem (3.42)

scaling on these travelling wave profiles. After numerous adjustments, we arrive at the final system (3.14) for the scaled travelling wave profiles, which we solve in Sect. 6. The reader uninterested in these intermediate stages may wish to proceed directly to Proposition 3.5, which discusses the equivalence of the problem (2.14) for A_j and P_j and the ultimate long wave system (3.42). Of course, our notation must keep up with these changes of variables, and we summarize in Table 1 the evolution of a typical operator's typesetting across these different problems.

Remark 2.1 The linearization of (2.14) at 0 yields

$$\dot{A}_j = \tau_2 \delta^+ \delta^- A_j, \qquad P_j = R_j = 0.$$

If we follow the discussion after Friesecke and Pego (1999, Thm. 1.1), as well as Faver and Wright (2018, Rem. 2.2), and look for plane wave solutions $A_j(t) = e^{ikj - i\omega t}$ with $\omega, k \in \mathbb{R}$, we find the dispersion relation

$$- i\omega = 2\tau_2 (\cos(k) - 1). \tag{2.15}$$

The only real solutions are $\omega = 0$ and $k \in 2\pi\mathbb{Z}$. Previously, in Friesecke and Pego (1999) and Faver and Wright (2018) a nontrivial dispersion relation $\omega = \omega(k)$ was found by making the same kind of plane wave Ansatz, and the result 'phase speed' $k \mapsto \omega(k)/k$ had a nonzero maximum c_s, which was called the 'speed of sound.' These articles then proceeded to look for travelling waves with speed slightly above their respective values of c_s; these were 'supersonic' waves. For us, $\omega(k)$ is identically zero, which suggests that the speed of sound for our auxin problem is 0. Our long wave scaling in Sect. 3.1 analytically justifies this intuition.

2.2 The travelling wave Ansatz

We now look for solutions A_j and P_j to (2.14) of the form

$$A_j = \phi_1(j - ct) \qquad \text{and} \qquad P_j = \phi_2(j - ct). \tag{2.16}$$

The profiles ϕ_1 and ϕ_2 are real-valued functions of a single real variable and $c \in \mathbb{R}$. The following manipulations will be justified if we assume $\phi_1 \in H_q^1$ and $\phi_2 \in W^{1,\infty}$; we discuss the exponentially localized Sobolev space H_q^1 in Appendix A.3. Working on an exponentially localized space, as opposed to an algebraically weighted space,

both allows us to capture precisely certain very fast decay properties and permits us to use some technical results on approximating Fourier multipliers. Furthermore, since we want P_j to vanish at $-\infty$ and be asymptotically constant at $+\infty$, per the limits (1.3) and the numerical predictions of Fig. 2, we expect that ϕ_2 should vanish at $+\infty$ and be asymptotically constant at $-\infty$.

We will convert the problem (2.14) for A_j and P_j into a nonlocal system for ϕ_1 and ϕ_2, with c as a parameter. Doing so amounts to little more than changing variables *many* times in the integral operators defined in Sects. 2.1.2 and 2.1.3 and gives us a host of new integral operators that will constitute the problem for ϕ_1 and ϕ_2.

In what follows we assume $f \in L^1$ and $g \in L^\infty$, so that the operators below are defined in the special cases of $f = \phi_1 \in H_q^1$ and $g = \phi_2 \in W^{1,\infty}$. First, for $x, v \in \mathbb{R}$, put

$$\widetilde{\mathsf{E}}^c(f)(v, x) := \exp\left(\frac{\kappa}{c}\int_v^x f(u+1)\, du\right) \tag{2.17}$$

and

$$\widetilde{\mathsf{P}}_1^c(f)(x) := \frac{\alpha}{c}\int_x^\infty \widetilde{\mathsf{E}}^c(f)(v, x)f(v)\, dv. \tag{2.18}$$

Then we use the Ansatz (2.16) and the definition of P_1 in (2.7) to find

$$P_1(A_{j+1}, A_j)(t) = \alpha \int_{-\infty}^t \exp\left(-\kappa\int_s^t \phi_1(j - c\xi + 1)\, d\xi\right)\phi_1(j - cs)\, ds$$
$$= \widetilde{\mathsf{P}}_1^c(\phi_1)(j - ct).$$

Here we have substituted $u = j - c\xi$ in the exponential's integral and then $v = j - cs$ throughout.

Similar substitutions, which we do not discuss, yield the following identities. Put

$$\widetilde{\mathsf{P}}_2^c(f, g)(x) := \frac{1}{c}\int_x^\infty \widetilde{\mathsf{E}}^c(f)(v, x)\mathsf{Q}_2(f(v+1), g(v))\, dv, \tag{2.19}$$

so that with P_2 defined in (2.8) we have

$$P_2(A_{j+1}, P_j)(t) = \widetilde{\mathsf{P}}_2^c(\phi_1, \phi_2)(j - ct).$$

Thus ϕ_2 must satisfy

$$\phi_2 = \widetilde{\mathsf{P}}_1^c(\phi_1) + \widetilde{\mathsf{P}}_2^c(\phi_1, \phi_2), \tag{2.20}$$

which indicates that, as expected, ϕ_2 should vanish at $+\infty$ and be asymptotically constant at $-\infty$.

Now we reformulate the equation for A_j, equivalently, for ϕ_1. Put

$$\widetilde{\mathsf{R}}_1^c(f)(x) := \frac{\kappa \tau_1}{c} \int_x^\infty f(u+1)\widetilde{\mathsf{P}}_1^c(f)(u)\, du, \tag{2.21}$$

so that with R_1 defined in (2.9) we have

$$\mathsf{R}_1(A_{j+1}, A_j)(t) = \widetilde{\mathsf{R}}_1^c(\phi_1)(j - ct).$$

Put

$$\widetilde{\mathsf{R}}_2^c(f, g)(x) := \frac{1}{c} \int_x^\infty \big(\kappa f(u+1)\widetilde{\mathsf{P}}_2^c(f, g)(u)$$
$$- Q_2(f(u+1), g(u)) \big)\, du \tag{2.22}$$

and

$$\widetilde{\mathsf{R}}^c(f, g) := \widetilde{\mathsf{R}}_1^c(f) + \widetilde{\mathsf{R}}_2^c(f, g), \tag{2.23}$$

so that with R_2 defined in (2.10) and R in (2.11) we have

$$\mathsf{R}_2(A_{j+1}, A_j, P_j)(t) = \widetilde{\mathsf{R}}_2^c(\phi_1, \phi_2)(j - ct) \qquad \text{and}$$
$$\mathsf{R}(A_{j+1}, A_j, P_j)(t) = \widetilde{\mathsf{R}}^c(\phi_1, \phi_2)(j - ct).$$

Last, put

$$\widetilde{\mathsf{N}}^c(f, g)(x) := \tau_1 \widetilde{\mathsf{R}}_2^c(f, g)(x) f(x) - \tau_1 Q_1(f(x), \widetilde{\mathsf{R}}^c(f, g)(x)), \tag{2.24}$$

so that with N defined in (2.13) we have

$$\mathsf{N}(A_{j+1}, A_j, P_j)(t) = \widetilde{\mathsf{N}}^c(\phi_1, \phi_2)(j - ct).$$

For a function $f : \mathbb{R} \to \mathbb{R}$ and $d \in \mathbb{R}$, define, as in (1.15), the shift operator S^d by

$$(S^d f)(x) := f(x+d). \tag{2.25}$$

This final piece of notation, along with the Eq. (2.20), allows us to convert the problem (2.14) for A_j and P_j into the following nonlocal system for ϕ_1 and ϕ_2:

$$\begin{cases} -c\phi_1' = \tau_2(S^1 - 2 + S^{-1})\phi_1 + (S^{-1} - 1)\big(\widetilde{\mathsf{R}}_1^c(\phi_1)\phi_1 + \widetilde{\mathsf{N}}^c(\phi_1, \phi_2)\big), \\ \\ \phi_2 = \widetilde{\mathsf{P}}_1^c(\phi_1) + \widetilde{\mathsf{P}}_2^c(\phi_1, \phi_2). \end{cases} \tag{2.26}$$

2.3 The Fourier multiplier structure

We summarize our conventions and definitions for Fourier transforms and Fourier multipliers in Appendix A. If we take the Fourier transform of the equation for ϕ_1 in (2.26), we find

$$\big(ick + 2\tau_2(\cos(k) - 1)\big)\widehat{\phi}_1(k) = (1 - e^{-ik})\mathfrak{F}\big[\widetilde{\mathsf{R}}_1^c(\phi_1)\phi_1 + \widetilde{\mathsf{N}}^c(\phi_1, \phi_2)\big](k).$$

For $k \in \mathbb{R}$, we have $ick + 2\tau_2(\cos(k) - 1) = 0$ if and only if $k = 0$. Consequently, the function

$$\widetilde{\mathsf{M}}_c(k) := \frac{1 - e^{-ik}}{ick + 2\tau_2(\cos(k) - 1)} \tag{2.27}$$

has a removable singularity at 0 and is in fact analytic on $\mathbb{R}$. We therefore define M_c to be the Fourier multiplier with symbol $\widetilde{\mathsf{M}}_c$, i.e., M_c satisfies

$$\widehat{\mathsf{M}_c f}(k) = \widetilde{\mathsf{M}}_c(k)\widehat{f}(k).$$

We discuss some further properties of Fourier multipliers in Appendix A.2. Now the problem (2.26) is equivalent to

$$\begin{cases} \phi_1 = \mathsf{M}_c\big(\widetilde{\mathsf{R}}_1^c(\phi_1)\phi_1 + \widetilde{\mathsf{N}}^c(\phi_1, \phi_2)\big) \\ \phi_2 = \widetilde{\mathsf{P}}_1^c(\phi_1) + \widetilde{\mathsf{P}}_2^c(\phi_1, \phi_2). \end{cases} \tag{2.28}$$

3 The long wave problem

3.1 The long wave scaling

We now make the long wave Ansatz

$$\phi_1(x) = \epsilon\psi_1(\epsilon^\mu x), \qquad \phi_2(x) = \epsilon^\beta\psi_2(\epsilon^\mu x), \qquad \text{and} \qquad c = \epsilon^\gamma c_0. \tag{3.1}$$

We assume, as with ϕ_1 and ϕ_2, that the scaled profiles satisfy $\psi_1 \in H_q^1$ and $\psi_2 \in W^{1,\infty}$. We think of $\epsilon > 0$ as small and keep the exponents $\beta, \gamma, \mu > 0$ arbitrary for now; eventually we will pick

$$\gamma = \mu = \frac{2}{5} \qquad \text{and} \qquad \beta = \frac{1}{5}.$$

The reasoning behind this choice is by no means obvious at this point and will not be for some time; leaving μ, β, and γ arbitrary will allow this choice to appear more naturally (at the cost of temporarily more cumbersome notation).

As we intuited in Remark 2.1, our wave speed is now close to 0, which is the auxin problem's natural 'speed of sound'. The parameter $c_0 \neq 0$ affords us some additional

flexibility in choosing the wave speed. A properly chosen value of c_0 will cause the maximum of the leading-order term of ϕ_1 to be ϵ, which will fulfill our promise in Sect. 1.4 that the auxin-height is, to leading order, ϵ. Friesecke and Pego introduce a similar auxiliary parameter into their ϵ-dependent wave speed, see Friesecke and Pego (1999, Eq. (2.5), (2.13)). This parameter allows them to prove that the dependence of their travelling wave profile on wave speed is sufficiently regular in different function spaces, a result needed for their subsequent stability arguments in Friesecke and Pego (2002, 2004a, b). We did not provide this extra parameter in our version (1.19) of the Friesecke–Pego wave speed, but rather we selected it so that the amplitude of the leading order sech^2-profile term in (1.22) is 1. Similarly, we will not pursue their depth of wave-speed analysis on our profiles' dependence on c_0.

We convert (2.28) to another nonlocal system for ψ_1 and ψ_2, which now depends heavily on the parameter ϵ. As before, this process mostly amounts to changing variables in many integrals. For example, we use the definition of $\widetilde{\mathsf{P}}_1^c$ in (2.18) and the Ansatz (3.1) to find

$$\widetilde{\mathsf{P}}_1^c(\phi_1)(x) = \frac{\alpha}{\epsilon^\gamma c_0} \int_x^\infty \widetilde{\mathsf{E}}^c(\phi_1)(v, x) \epsilon \psi_1(\epsilon^\mu v) \, dv, \tag{3.2}$$

where, using the definition of $\widetilde{\mathsf{E}}^c$ in (2.17), we have

$$\widetilde{\mathsf{E}}^c(\phi_1)(v, x) = \exp\left(\frac{\kappa}{\epsilon^\gamma c_0} \int_v^x \epsilon \psi_1(\epsilon^\mu u + \epsilon^\mu) \, du\right)$$

$$= \exp\left(\frac{\kappa}{c_0} \epsilon^{1-(\gamma+\mu)} \int_{\epsilon^\mu v}^{\epsilon^\mu x} \psi_1(U + \epsilon^\mu) \, dU\right).$$

Here we have substituted $U = \epsilon^\mu u$.

Now for $f \in L^1$ we put

$$\mathcal{E}(f)(V, X) := \exp\left(\frac{\kappa}{c_0} \int_V^X f(U) \, dU\right), \quad V, X \in \mathbb{R}, \tag{3.3}$$

so that (3.2) becomes

$$\widetilde{\mathsf{P}}_1^c(\phi_1)(x) = \frac{\alpha}{c_0} \epsilon^{1-\gamma} \int_x^\infty \mathcal{E}(\epsilon^{1-(\gamma+\mu)} S^{\epsilon^\mu} \psi_1)(\epsilon^\mu v, \epsilon^\mu x) \psi_1(\epsilon^\mu v) \, dv.$$

Here S^{ϵ^μ} is the shift operator defined in (2.25) with $d = \epsilon^\mu$. We substitute again with $V = \epsilon^\mu v$ and define

$$\check{\mathsf{P}}_1^\epsilon(f)(X) := \frac{\alpha}{c_0} \int_X^\infty \mathcal{E}(\epsilon^{1-(\gamma+\mu)} S^{\epsilon^\mu} f)(V, X) f(V) \, dV \tag{3.4}$$

to conclude that

$$\widetilde{\mathsf{P}}_1^c(\phi_1)(x) = \epsilon^{1-(\gamma+\mu)} \check{\mathsf{P}}_1^\epsilon(\psi_1)(\epsilon^\mu x).$$

Similar careful substitutions will allow us to reformulate the integral operators from Sect. 2.2 in terms of the long wave Ansatz. First, however, we define

$$\check{Q}_1^\epsilon(X, Y) := \frac{X^2 Y}{k_a + \epsilon X} \quad \text{and} \quad \check{Q}_2^\epsilon(X, Y) := \kappa \frac{k_r Y + k_m \epsilon^{1-\beta} X + \epsilon XY}{(k_r + \epsilon X)(k_m + \epsilon^\beta Y)} XY.$$

$$(3.5)$$

When $\epsilon \neq 0$, this definition permits the very convenient factorizations

$$Q_1(\epsilon X, \epsilon^{1-(\gamma+\mu)} Y) = \epsilon^{3-(\gamma+\mu)} \check{Q}_1^\epsilon(X, Y) \quad \text{and} \quad Q_2(\epsilon X, \epsilon^\beta Y) = \epsilon^{1+2\beta} \check{Q}_2^\epsilon(X, Y),$$

where Q_1 was defined in (2.1) and Q_2 in (2.4).

Now we work on the travelling wave integral operators. Below we will assume $f \in L^1$ and $g \in L^\infty$. Put

$$\check{P}_2^\epsilon(f, g)(X) := \frac{1}{c_0} \int_X^\infty \mathcal{E}(\epsilon^{1-(\gamma+\mu)} S^{\epsilon^\mu} f)(V, X) \check{Q}_2^\epsilon(f(V + \epsilon^\mu), g(V)) \, dV,$$

$$(3.6)$$

so that with $\widetilde{P}_2^c$ defined in (2.19) we have

$$\widetilde{P}_2^c(\phi_1, \phi_2)(x) = \epsilon^{1-(\gamma+\mu)+2\beta} \check{P}_2^\epsilon(\psi_1, \psi_2)(\epsilon^\mu x).$$

This converts the second equation in (2.28) for ϕ_2 to

$$\epsilon^\beta \psi_2(\epsilon^\mu x) = \epsilon^{1-(\gamma+\mu)} \check{P}_1^\epsilon(\psi_1)(\epsilon^\mu x) + \epsilon^{1-(\gamma+\mu)+2\beta} \check{P}_2^\epsilon(\psi_1, \psi_2)(\epsilon^\mu x).$$

Passing to $X = \epsilon^\mu x$, we find that ψ_2 must satisfy

$$\psi_2(X) = \epsilon^{1-(\gamma+\mu)-\beta} \check{P}_1^\epsilon(\psi_1)(X) + \epsilon^{1-(\gamma+\mu)+\beta} \check{P}_2^\epsilon(\psi_1, \psi_2)(X). \qquad (3.7)$$

Now put

$$\check{R}_1^\epsilon(f)(X) := \frac{\kappa \tau_1}{c_0} \int_X^\infty \check{P}_1^\epsilon(f)(V) f(V + \epsilon^\mu) \, dV, \qquad (3.8)$$

so that with $\widetilde{R}_1^c$ defined in (2.21) we have

$$\widetilde{R}_1^c(\phi_1)(x) = \epsilon^{2(1-(\gamma+\mu))} \check{R}_1^\epsilon(\psi_1)(\epsilon^\mu x).$$

Put

$$\check{R}_2^\epsilon(f, g)(X) := \frac{1}{c_0} \int_X^\infty \left(\epsilon^{1-(\gamma+\mu)} \kappa f(V + \epsilon^\mu) \check{P}_2^\epsilon(f, g)(V) \right.$$
$$\left. -\check{Q}_2^\epsilon(f(V + \epsilon^\mu), g(V)) \right) dV \qquad (3.9)$$

and

$$\check{R}^\epsilon(f,g)(X) := \epsilon^{1-(\gamma+\mu)}\check{R}_1^\epsilon(f)(X) + \epsilon^{2\beta}R_2^\epsilon(f,g)(X), \qquad (3.10)$$

so that with $\widetilde{R}_2^c$ defined in (2.22) and $\widetilde{R}^c$ defined in (2.23) we have

$$\begin{aligned}\widetilde{R}_2^c(\phi_1,\phi_2)(x) &= \epsilon^{1-(\gamma+\mu)+2\beta}\check{R}_2^\epsilon(\psi_1,\psi_2)(\epsilon^\mu x) \qquad \text{and}\\\widetilde{R}^c(\phi_1,\phi_2)(x) &= \epsilon^{1-(\gamma+\mu)}\check{R}^\epsilon(\psi_1,\psi_2)(\epsilon^\mu x).\end{aligned}$$

Finally, put

$$\check{N}^\epsilon(f,g)(X) := \tau_1\check{R}_2^\epsilon(f,g)(X)f(X) - \epsilon^{1-2\beta}\tau_1\check{Q}_1^\epsilon(f(X),\check{R}^\epsilon(f,g)(X)), (3.11)$$

so that with $\widetilde{N}^c$ defined in (2.24) we have

$$\widetilde{N}^c(\phi_1,\phi_2)(x) = \epsilon^{2-(\gamma+\mu)+2\beta}\check{N}^\epsilon(\psi_1,\psi_2)(\epsilon^\mu x).$$

The definition of scaled Fourier multipliers from (A.3) tells us that, for $\epsilon > 0$, $\mathsf{M}_{\epsilon^\gamma c_0}^{(\epsilon^\mu)}$ is the Fourier multiplier satisfying

$$\widehat{\mathsf{M}_{\epsilon^\gamma c_0}^{(\epsilon^\mu)}f}(k) = \widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\mu k)\widehat{f}(k),$$

where $\widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}$ is defined by taking $c = \epsilon^\gamma c_0$ in (2.27). This converts the first equation in (2.28) for ϕ_1 to

$$\epsilon\psi_1(\epsilon^\mu x) = \mathsf{M}_{\epsilon^\gamma c_0}^{(\epsilon^\mu)}[\epsilon^{2(1-(\gamma+\mu))}\check{R}_1^\epsilon(\psi_1)\epsilon\psi_1 + \epsilon^{2-(\gamma+\mu)+2\beta}\check{N}^\epsilon(\psi_1,\psi_2)](\epsilon^\mu x).$$

We factor this to reveal

$$\psi_1(X) = \epsilon^{2(1-(\gamma+\mu))}\mathsf{M}_{\epsilon^\gamma c_0}^{(\epsilon^\mu)}[\check{R}_1^\epsilon(\psi_1)\psi_1 + \epsilon^{-1+\gamma+\mu+2\beta}\mathsf{N}^\epsilon(\psi_1,\psi_2)](X). \quad (3.12)$$

We abbreviate

$$\check{\mathsf{M}}_\epsilon := \epsilon^{2(1-(\gamma+\mu))}\mathsf{M}_{\epsilon^\gamma c_0}^{(\epsilon^\mu)} \qquad (3.13)$$

to conclude from (3.12) and the prior Eq. (3.7) for ψ_2 that the long wave profiles must satisfy

$$\begin{cases} \psi_1 = \check{\mathsf{M}}_\epsilon[\check{R}_1^\epsilon(\psi_1)\psi_1 + \epsilon^{-1+\gamma+\mu+2\beta}\check{N}^\epsilon(\psi_1,\psi_2)] \\[2mm] \psi_2 = \epsilon^{1-(\gamma+\mu+\beta)}\check{P}_1^\epsilon(\psi_1) + \epsilon^{1-(\gamma+\mu)+\beta}\check{P}_2^\epsilon(\psi_1,\psi_2). \end{cases} \qquad (3.14)$$

We have been tacitly assuming that all of the exponents on powers of ϵ above are nonnegative so that the various ϵ-dependent operators and prefactors are actually defined at $\epsilon = 0$. In particular, this demands

$$1 - 2\beta \geq 0, \qquad -1 + \gamma + \mu + 2\beta \geq 0, \qquad \text{and} \qquad 1 - (\gamma + \mu + \beta) \geq 0. \tag{3.15}$$

3.2 The formal long wave limit and exponent selection

Our intention is now to take the limit $\epsilon \to 0$ in the Eq. (3.14) for ψ_1 and ψ_2. Doing so in a way that the limit is both meaningful (i.e., defined and nontrivial) and reflective of what the numerics predict at $\epsilon = 0$ will teach us what the exponents μ, γ, and β should be, beyond the requirements of (3.15).

3.2.1 The formal limit on $\check{M}_\epsilon$ and the selection of the exponents γ and μ

We want to assign a 'natural' definition to $\check{M}_0$, where $\check{M}_\epsilon$ was defined, for $\epsilon > 0$, in (3.13). However, we relied above on having $\epsilon > 0$ to invoke the scaled Fourier multiplier identity (A.3) that gave us $\check{M}_\epsilon$, and naively setting $\epsilon = 0$ in that identity is meaningless. Additionally, we should be careful that the prefactor $\epsilon^{2(1-(\gamma+\mu))}$ in (3.13) does not lead us to define $\check{M}_0 = 0$; otherwise, we would have $\psi_1 = 0$ when $\epsilon = 0$, and that is not what the numerics in Fig. 2 predict.

A natural starting point, then, is to study $\check{M}_\epsilon$ in the limit $\epsilon \to 0^+$, and this amounts to considering the limit of its symbol, whose definition we extract from the definition of $\check{M}_\epsilon$ in (3.13) and the definition of the scaled Fourier multiplier in (A.3). Thus, for each $k \in \mathbb{R}$, we want the limit

$$\lim_{\epsilon \to 0^+} \epsilon^{2(1-(\gamma+\mu))} \widetilde{M}_{\epsilon^\gamma c_0}(\epsilon^\mu k) \tag{3.16}$$

to exist without being identically zero. The function $\widetilde{M}_{\epsilon^\gamma c_0}$ was defined in (2.27).

To calculate this limit, we first state the Taylor expansions

$$1 - e^{-iz} = iz + iz^2 N_1(z) \qquad \text{and} \qquad \cos(z) - 1 = -\frac{z^2}{2} + \frac{iz^4 N_2(z)}{2\tau_2} \tag{3.17}$$

for $z \in \mathbb{C}$. The functions N_1 and N_2 are analytic and uniformly bounded on strips in the sense that

$$C_q := \sup_{x \in \mathbb{R}} |N_1(x \pm iq)| + |N_2(x \pm iq)| < \infty \tag{3.18}$$

for any $q > 0$. The choice of constants on N_1 and N_2 will permit some useful cancellations later. Then

$$\widetilde{M}_c(k) = \frac{ik + ik^2 N_1(k)}{ick - \tau_2 k^2 + ik^4 N_2(k)} = \frac{1 + k N_1(k)}{c + i\tau_2 k + k^3 N_2(k)},$$

and so

$$\epsilon^{2(1-(\gamma+\mu))}\widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\mu k) = \epsilon^{2(1-(\gamma+\mu))}\frac{1+\epsilon^\mu k N_1(\epsilon^\mu k)}{\epsilon^\gamma c_0 + i\tau_2\epsilon^\mu k + \epsilon^{3\mu}k^3 N_2(\epsilon^\mu k)}. \quad (3.19)$$

At this point it does not make sense to set $\epsilon = 0$, as then the denominator would be identically zero. So, we would like to factor some power of ϵ out of the denominator. Since the first term in the denominator has a factor of ϵ^γ and the second a factor of ϵ^μ, we assume $\gamma = \mu$ and remove the power of ϵ from both the first and the second terms. We discuss the choice of $\gamma = \mu$ further in Remark 3.2.

Then

$$\epsilon^{2(1-(\gamma+\mu))}\widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\mu k) = \epsilon^{2(1-2\gamma)}\widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\gamma k)$$
$$= \epsilon^{2(1-2\gamma)-\gamma}\frac{1+\epsilon^\gamma k N_1(\epsilon^\gamma k)}{c_0 + i\tau_2 k + \epsilon^{2\gamma}k^3 N_2(\epsilon^\gamma k)}. \quad (3.20)$$

Pointwise in k we have

$$\lim_{\epsilon\to 0^+}\frac{1+\epsilon^\gamma k N_1(\epsilon^\gamma k)}{c_0 + i\tau_2 k + \epsilon^{2\gamma}k^3 N_2(\epsilon^\gamma k)} = \frac{1}{c_0 + i\tau_2 k},$$

and so we want

$$2(1-2\gamma)-\gamma = 0$$

so that the prefactor of $\epsilon^{2(1-2\gamma)-\gamma}$ in (3.20) does not induce a trivial or undefined limit. Thus we take

$$\gamma = \mu = \frac{2}{5}.$$

Certainly doing so does not contradict any of the inequalities in (3.15), provided that β is chosen appropriately. Moreover, the power of 2/5 agrees with the height-speed-width relations suggested in Fig. 4. And so

$$\lim_{\epsilon\to 0^+}\epsilon^{2(1-(\gamma+\mu))}\widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\mu k) = \lim_{\epsilon\to 0^+}\epsilon^{4/5}\widetilde{\mathsf{M}}_{\epsilon^{2/5}c_0}(\epsilon^{2/5}k) = \frac{1}{c_0 + \tau_2 ik}.$$

Put

$$\widetilde{\mathcal{M}}^{(0)}(z) := \frac{1}{c_0 + i\tau_2 z}, \quad (3.21)$$

so $\widetilde{\mathcal{M}}^{(0)}$ is analytic on any strip $\{z \in \mathbb{C} \mid |\mathrm{Im}(z)| < q\}$ for $q \in (0, \tau_2/c_0)$. Let $\mathcal{M}^{(0)}$ be the Fourier multiplier with symbol $\widetilde{\mathcal{M}}^{(0)}$.

Lemma A.2 then gives the following properties of $\mathcal{M}^{(0)}$; the identities (3.22) are direct calculations with the Fourier transform.

Lemma 3.1 *Fix* $q \in (0, \tau_2/c_0)$. *Then* $\mathcal{M}^{(0)} \in \mathbf{B}(H_q^r, H_q^{r+1})$ *for all* r. *More generally, if* $f \in H^1$ *and* $g \in L^2$, *then*

$$\mathcal{M}^{(0)}(c_0 + \tau_2 \partial_X) f = f \quad and \quad (c_0 + \tau_2 \partial_X) \mathcal{M}^{(0)} g = g. \tag{3.22}$$

Because of the identities (3.22), we write $\mathcal{M}^{(0)} = (c_0 + \tau_2 \partial_X)^{-1}$. The formal analysis above then leads us to expect

$$\lim_{\epsilon \to 0^+} \check{\mathsf{M}}_\epsilon = \mathcal{M}^{(0)} = (c_0 + \tau_2 \partial_X)^{-1}. \tag{3.23}$$

However, we have not yet proved this rigorously by any means.

Remark 3.2 Here is why we take $\gamma = \mu$ when factoring the power of ϵ out of the denominator in (3.19). First, taking $\gamma > \mu$ produces

$$\epsilon^{2(1-(\gamma+\mu))} \widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\mu k) = \epsilon^{2(1-(\gamma+\mu))-\mu} \frac{1 + \epsilon^\mu k N_1(\epsilon^\mu k)}{\epsilon^{\gamma-\mu} c_0 + i \tau_2 k + \epsilon^{2\mu} k^3 N_2(\epsilon^\mu k)}$$

instead of (3.20). If $2(1 - (\gamma + \mu)) - \mu > 0$, then the right side above is identically zero at $\epsilon = 0$, and so we demand $2(1 - (\gamma + \mu)) - \mu = 0$; there are many pairs of γ and μ that work here. But then

$$\lim_{\epsilon \to 0^+} \frac{1 + \epsilon^\mu k N_1(\epsilon^\mu k)}{\epsilon^{\gamma-\mu} c_0 + i \tau_2 k + \epsilon^{2\mu} k^3 N_2(\epsilon^\mu k)} = \frac{1}{i \tau_2 k}.$$

This suggests that instead of (3.23), we have

$$\lim_{\epsilon \to 0^+} \check{\mathsf{M}}_\epsilon = (\tau_2 \partial_X)^{-1}.$$

However, this is meaningless: differentiation is not invertible from H_q^r to H_q^{r+1}.

Taking $\gamma < \mu$ also does not work. In that case, instead of (3.20) we would have found

$$\epsilon^{2(1-(\gamma+\mu))} \widetilde{\mathsf{M}}_{\epsilon^\gamma c_0}(\epsilon^\mu k) = \epsilon^{2(1-(\gamma+\mu))-\gamma} \frac{1 + \epsilon^\mu k N_1(\epsilon^\mu k)}{c_0 + i \tau_2 \epsilon^{\mu-\gamma} k + \epsilon^{3\mu-\gamma} k^3 N_2(\epsilon^\mu k)}.$$

Since $\gamma < \mu$ we find

$$\lim_{\epsilon \to 0^+} \frac{1 + \epsilon^\mu k N_1(\epsilon^\mu k)}{c_0 + i \tau_2 \epsilon^{\mu-\gamma} k + \epsilon^{3\mu-\gamma} k^3 N_2(\epsilon^\mu k)} = \frac{1}{c_0}.$$

We would then want $2 - 3\gamma - 2\mu = 0$ to prevent a nontrivial limit.

Choosing γ and μ appropriately, we conclude that at $\epsilon = 0$ the equation for ψ_1 from (3.14) formally reduces to

$$\psi_1 = \frac{1}{c_0} \check{\mathsf{R}}_1^0(\psi_1) \psi_1.$$

Numerically we expect $\psi_1(X) > 0$ for all X when $\epsilon = 0$, and so, using the definition of $\check{\mathsf{R}}_1^0$ from (3.8), we have

$$c_0 = \check{\mathsf{R}}_1^0(\psi_1)(X) = \frac{\alpha\kappa\tau_1}{c_0^2} \int_X^\infty \left(\int_V^\infty \psi_1(W)\,dW \right) \psi_1(V)\,dV.$$

Differentiating, we find

$$\left(\int_X^\infty \psi_1(W)\,dW \right) \psi_1(X) = 0.$$

But since $\psi_1(W) > 0$ for all W, we cancel the integral factor to find $\psi_1(X) = 0$, a contradiction to our numerical predictions.

3.2.2 The formal leading order equation for ψ_1

At $\epsilon = 0$ the equation for ψ_1 in (3.14) becomes (again, formally)

$$\psi_1 = \mathcal{M}^{(0)}\big(\check{\mathsf{R}}^0(\psi_1)\psi_1\big) = (c_0 + \tau_2\partial_X)^{-1}\big(\check{\mathsf{R}}^0(\psi_1)\psi_1\big).$$

This is equivalent to

$$c_0\psi_1 + \tau_2\psi_1' = \check{\mathsf{R}}^0(\psi_1)\psi_1. \tag{3.24}$$

We will rewrite this equation so that each term is a perfect derivative.

The definition of $\check{\mathsf{R}}_1^\epsilon$ in (3.8), valid for all ϵ, gives

$$\check{\mathsf{R}}^0(\psi_1)(X) = \frac{\alpha\kappa\tau_1}{c_0^2} \int_X^\infty \left(\int_V^\infty \psi_1(W)\,dW \right) \psi_1(V)\,dV. \tag{3.25}$$

Write

$$\Psi_1(X) := \int_X^\infty \psi_1(W)\,dW,$$

so that $\Psi_1' = -\psi_1$. The double integral from (3.25) is

$$\int_X^\infty \left(\int_V^\infty \psi_1(W)\,dW \right) \psi_1(V)\,dV = -\int_X^\infty \Psi_1(V)\Psi_1'(V)\,dV$$

$$= -\int_X^\infty \partial_V \left[\frac{\Psi_1^2(V)}{2} \right] dV$$

$$= \frac{\Psi_1(X)^2}{2}.$$

Here we are using the requirement that $\psi_1 \in H_q^1$, which implies $\Psi_1(X) \to 0$ as $X \to \infty$. Thus

$$\overset{\circ}{\mathsf{R}}{}^0(\psi_1)\psi_1 = -\left(\frac{\alpha\kappa\tau_1}{2c_0^2}\right)\Psi_1^2\Psi_1' = -\left(\frac{\alpha\kappa\tau_1}{6c_0^2}\right)\partial_X[\Psi_1^3].$$

Then (3.24) is equivalent to

$$\tau_2\Psi_1'' + c_0\Psi_1' - \left(\frac{\alpha\kappa\tau_1}{6c_0^2}\right)\partial_X[\Psi_1^3] = 0.$$

We integrate both sides from 0 to ∞ and use the aforementioned fact that Ψ_1 and its derivatives are required to vanish at ∞ to find

$$\tau_2\Psi_1' + c_0\Psi_1 - \left(\frac{\alpha\kappa\tau_1}{6c_0^2}\right)\Psi_1^3 = 0. \tag{3.26}$$

This is a Bernoulli equation, and it has the solution

$$\Psi_1(X) = \Sigma(X) := \left(\frac{6c_0^3}{\alpha\kappa\tau_1 + 6c_0^2\exp\left(2c_0X/\tau_2 + \theta\right)}\right)^{1/2}. \tag{3.27}$$

Here $\theta \in \mathbb{R}$ is an arbitrary phase shift. It follows that putting

$$\psi_1(X) = \sigma(X) := -\Sigma'(X) = \frac{(6c_0^3)^{3/2}\exp\left(2c_0X/\tau_2 + \theta\right)}{\tau_2\left[\alpha\kappa\tau_1 + 6c_0^2\exp\left(2c_0X/\tau_2 + \theta\right)\right]^{3/2}} \tag{3.28}$$

solves (3.24).

Remark 3.3 We view the free parameter θ in (3.28) as an artifact of the translation invariance of our original problem (1.1). Friesecke and Pego (1999) do not incorporate a phase shift like θ into their leading order sech^2-type KdV solution, since their broader existence result relies on working in spaces of even functions, and phase shifts destroy evenness. We will not need such symmetry in our subsequent arguments (nor could we achieve it, since no translation of σ is even or odd), and so we will leave θ as an arbitrary free parameter and not specify its value. Instead, we will restrain the extra degree of freedom of translation invariance by imposing a certain integral condition, which we make precise in (5.2).

3.2.3 The formal leading order equation for ψ_2 and the selection of the exponent β

From our choice of $\gamma = \mu = 2/5$ and the inequalities in (3.15), we need, at the very least,

$$\frac{1}{10} \le \beta \le \frac{1}{5}.$$

If the strict inequality $\beta < 1/5$ holds, then at $\epsilon = 0$ the equation for ψ_2 in (3.14) reduces to the trivial result $\psi_2 = 0$. This is not at all what we expect numerically from Fig. 2; rather, we anticipate that ψ_2 will asymptote to some nonzero constant at ∞.

However, if we instead take β so that

$$0 = 1 - (\gamma + \mu + \beta) = \frac{1}{5} - \beta,$$

which is to say,

$$\beta = \frac{1}{5},$$

then the equation for ψ_2 in (3.14) at $\epsilon = 0$ becomes

$$\psi_2 = \check{\mathsf{P}}_1^0(\psi_1).$$

Putting

$$\psi_2(X) = \zeta(X) := \check{\mathsf{P}}_1^0(\sigma)(X) = \frac{\alpha}{c_0} \int_X^\infty \sigma(V)\, dV \tag{3.29}$$

therefore solves the leading order equation for ψ_2. We really have

$$\zeta(X) = \frac{\alpha}{c_0} \Sigma(X) = \frac{\alpha}{c_0} \left(\frac{6c_0^3}{\alpha\kappa\tau_1 + 6c_0^2 e^{2c_0 X/\tau_2 + \theta}} \right)^{1/2},$$

where Σ was defined in (3.27).

3.3 The final long wave system

With the choices of exponents $\gamma = \mu = 2/5$ and $\beta = 1/5$, it becomes convenient to introduce the new small parameter

$$\nu := \epsilon^{2/5} \tag{3.30}$$

into the problem (3.14) and then recast that problem more cleanly in terms of ν. First, the long wave Ansatz (3.1) becomes

$$\phi_1(x) = \nu^{5/2}\psi_1(\nu x), \qquad \phi_2(x) = \nu^{1/2}\psi_2(\nu x), \qquad \text{and} \qquad c = \nu c_0. \tag{3.31}$$

Proceeding very much as in Sect. 3.1, we then define

$$\mathcal{Q}_1^\nu(X, Y) := \frac{X^2 Y}{k_a(k_a + \nu^{5/2}X)} \qquad \text{and}$$

$$\mathcal{Q}_2^\nu(X, Y) := \kappa \frac{k_r Y + k_m \nu^2 X + \nu^{5/2} XY}{(k_r + \nu^{5/2}X)(k_m + \nu^{1/2}Y)} XY \tag{3.32}$$

for $X, Y \in \mathbb{R}$, while for $f \in L^1$ and $g \in L^\infty$, we put

$$\mathcal{P}_1^\nu(f)(X) := \frac{\alpha}{c_0} \int_X^\infty \mathcal{E}(\nu^{1/2} S^\nu f)(V, X) f(V) \, dV, \qquad (3.33)$$

where $\mathcal{E}$ was defined in (3.3), and

$$\mathcal{P}_2^\nu(f, g)(X) := \frac{1}{c_0} \int_X^\infty \mathcal{E}(\nu^{1/2} S^\nu f)(V, X) \mathcal{Q}_2^\nu(f(V + \nu), g(V)) \, dV, \quad (3.34)$$

$$\mathcal{R}_1^\nu(f)(X) := \frac{\kappa \tau_1}{c_0} \int_X^\infty \mathcal{P}_1^\nu(f)(V) f(V + \nu) \, dV, \qquad (3.35)$$

$$\mathcal{R}_2^\nu(f, g)(X) := \frac{1}{c_0} \int_X^\infty \left(\nu^{1/2} \kappa f(V + \nu) \mathcal{P}_2^\nu(f, g)(V) \right.$$
$$\left. - \mathcal{Q}_2^\nu(f(V + \nu), g(V)) \right) dV, \qquad (3.36)$$

$$\mathcal{R}^\nu(f, g)(X) := \mathcal{R}_1^\nu(f)(X) + \nu^{1/2} \mathcal{R}_2^\nu(f, g)(X), \qquad (3.37)$$

and

$$\mathcal{N}^\nu(f, g)(X) := \tau_1 \mathcal{R}_2^\nu(f, g)(X) f(X) - \nu^{3/2} \tau_1 \mathcal{Q}_1^\nu(f(X), \mathcal{R}^\nu(f, g)(X)). \qquad (3.38)$$

Remark 3.4 The operators $\mathcal{P}_1^\nu$ and $\mathcal{R}_1^\nu$ map L^1 into L^∞, while $\mathcal{P}_2^\nu$ and $\mathcal{R}_2^\nu$ map $L^1 \times L^\infty$ into L^∞, and $\mathcal{N}^\nu$ maps $L^1 \times L^\infty$ into L^1. More precisely, we could replace L^1 with H_q^1 and L^∞ with $W^{1,\infty}$ and the preceding statement would still be true; see the estimates in Appendix B.1.

The operator $\mathcal{R}_1^0$ has the especially simple form

$$\mathcal{R}_1^0(f)(X) = \left(\frac{\alpha \kappa \tau_1}{6 c_0^2} \right) \int_X^\infty \left(\int_V^\infty f(W) \, dW \right) f(V) \, dV \qquad (3.39)$$

and therefore is differentiable from L^1 to L^∞.

Last, for $\nu > 0$, let $\mathcal{M}^{(\nu)}$ be the Fourier multiplier with symbol

$$\widetilde{\mathcal{M}}^{(\nu)}(z) := \nu \frac{1 - e^{-i\nu z}}{i c_0 \nu^2 z + 2\tau_2(\cos(\nu z) - 1)}. \qquad (3.40)$$

When $\nu = 0$ we have already defined $\mathcal{M}^{(0)}$ as the Fourier multiplier whose symbol $\widetilde{\mathcal{M}}^{(0)}$ is given in (3.21). We will show momentarily in Sect. 4 how $\mathcal{M}^{(0)}$ is a good approximation of $\mathcal{M}^{(\nu)}$ in the operator norm topology, which we formally anticipated in the limit (3.23).

We now summarize the work of this section and the preceding one. In Sect. 2 we reformulated our original problem (1.1) into the more concise structure (2.14) and then made a travelling wave ansatz on this latter system. That led to the travelling wave

problem (2.28). In this section, we introduced the long wave Ansatz (3.1) into this travelling wave problem and studied several formal expansions and limits to deduce the 'correct' choice of exponents and scalings. With the operators defined above, we can summarize our long wave problem in the following form.

Proposition 3.5 *Suppose*

$$\begin{cases} A_j(t) = \nu^{5/2}\psi_1(\nu(j - \nu c_0 t)), \\ P_j(t) = \nu^{1/2}\psi_2(\nu(j - \nu c_0 t)) \end{cases} \tag{3.41}$$

for some $\psi_1 \in H_q^1$ and $\psi_2 \in L^\infty$, where $c_0, \nu > 0$ and $q \in (0, c_0/\tau_2)$. Then A_j and P_j satisfy (2.14) if and only if ψ_1 and ψ_2 satisfy

$$\begin{cases} \psi_1 = \mathcal{M}^{(\nu)}\big(\mathcal{R}_1^\nu(\psi_1)\psi_1 + \nu^{1/2}\mathcal{N}^\nu(\psi_1, \psi_2)\big), \\ \psi_2 = \mathcal{P}_1^\nu(\psi_1) + \nu\mathcal{P}_2^\nu(\psi_1, \psi_2). \end{cases} \tag{3.42}$$

Moreover, taking

$$\psi_1 = \sigma \qquad and \qquad \psi_2 = \zeta = \mathcal{P}_1^0(\sigma),$$

where σ is defined in (3.28) and ζ is given explicitly in (3.29), solves (3.42) when $\nu = 0$.

We will analyze the system (3.42) with a quantitative contraction mapping argument that tracks its dependence on ν. Specifically, we will look for solutions ψ_1 and ψ_2 to (3.42) that are close to σ and ζ, respectively, when ν is close to 0. We provide the details of this argument in Sect. 6.

Before doing so, we need to understand the behavior of two key operators, on whose good behavior the successful contraction argument will hinge. Our first task, as we mentioned above, is to study how $\mathcal{M}^{(0)}$ approximates $\mathcal{M}^{(\nu)}$, and we do this in Sect. 4. Next, since we are looking for solutions (ψ_1, ψ_2) to (3.42) that are close to (σ, ζ), it is natural to study the linearization of (3.42) at (σ, ζ) for $\nu = 0$. It turns out that we will only need the linearization of the first equation, which is the operator

$$\mathcal{T}\psi := \psi - \mathcal{M}^{(0)}\big[\mathcal{R}_1^0(\sigma)\psi + (D\mathcal{R}_1^0(\sigma)\psi)\sigma\big]. \tag{3.43}$$

We will show in Sect. 5 that $\mathcal{T}$ is surjective from H_q^1 to H_q^1 with a one-dimensional kernel. Restricting $\mathcal{T}$ off this kernel will yield an extremely useful bijectivity result.

4 Analysis of the Fourier multiplier $\mathcal{M}^{(\nu)}$

We show that the Fourier multiplier $\mathcal{M}^{(\nu)}$, whose symbol was defined in (3.40), converges to the multiplier $\mathcal{M}^{(0)}$, whose symbol was defined in (3.21). More precisely, we prove the following estimate.

Proposition 4.1 *Fix $q \in (0, c_0/\tau_2)$. There exist $\nu_\mathcal{M}$, $C_\mathcal{M} > 0$ such that if $0 < \nu < \nu_\mathcal{M}$, then*

$$\|\mathcal{M}^{(\nu)} - \mathcal{M}^{(0)}\|_{\mathbf{B}(H_q^1)} \leq C_\mathcal{M} \nu^{1/3}.$$

The proof of this estimate depends on the following lemma, whose proof we give in the remainder of this section.

Lemma 4.2 *Let $q \in (0, c_0/\tau_2)$. There exist $C_\mathcal{M}$, $\nu_\mathcal{M} > 0$ such that if $0 < \nu < \nu_\mathcal{M}$, then the map $\widetilde{\mathcal{M}}^{(\nu)}$ defined in (3.40) is analytic on the strip $\overline{\mathcal{U}}_q :=\{z \in \mathbb{C} \mid |\mathrm{Im}(z)| \leq q\}$ and satisfies*

$$\sup_{k \in \mathbb{R}} \left| \widetilde{\mathcal{M}}^{(\nu)}(k \pm iq) - \widetilde{\mathcal{M}}^{(0)}(k \pm iq) \right| < C_\mathcal{M} \nu^{1/3}, \tag{4.1}$$

where $\widetilde{\mathcal{M}}^{(0)}$ was defined in (3.21).

This lemma allows us to invoke Beale's result in Lemma A.2 (with $r = 1$ and $s = 0$) to prove Proposition 4.1. Beale's result depends very much on our working in exponentially weighted Sobolev spaces, and this is another of our reasons for preferring these spaces to algebraically weighted ones.

We will estimate the difference $\left| \widetilde{\mathcal{M}}^{(\nu)}(z) - \widetilde{\mathcal{M}}^{(0)}(z) \right|$ over two regimes, one in which $z = k \pm iq$ is 'close' to 0, and the other in which z is 'far from' 0. Part of these estimates will involve bounding the denominator of $\widetilde{\mathcal{M}}^{(\nu)}$ away from zero; this will ensure the analyticity of $\widetilde{\mathcal{M}}^{(\nu)}$, since it is the quotient of two analytic functions.

To quantify these regimes, we introduce two positive constants p and m; we say that z is 'close' to 0 if $|z| \leq \nu^{-p}$ and 'far from' 0 if $|z| > \nu^{-p}$. The constant m will later control how close the real part of νz is to an integer multiple of 2π, a bound that will be very useful in certain estimates to come. All constants C in the work below are allowed to depend on m, p, and q, but they are always independent of ν and z.

Our estimates will depend on the parameters p and m; once we have all the estimates together, we will choose useful values for p and m. We feel that this approach allows the otherwise nonobvious final values for p and m to emerge very naturally. This strategy of splitting the estimates over regions close to and far from 0 is modeled on the proofs of Faver and Wright (2018, Lem. A.13) and Stefanov and Wright (2020, Lem. 3) and the strategy in Johnson and Wright (2020, App. A.3). Friesecke and Pego Friesecke and Pego (1999, Sec. 3) give a rather different proof of symbol convergence that relies on more knowledge of the poles of $\widetilde{\mathcal{M}}^{(\nu)}$ than we care to discover.

4.1 Estimates for z 'close to' 0

In this regime we fix $|z| \leq \nu^{-p}$. We recall the Taylor expansions

$$1 - e^{-iz} = iz + iz^2 N_1(z) \qquad \text{and} \qquad \cos(z) - 1 = -\frac{z^2}{2} + \frac{iz^4 N_2(z)}{2\tau_2}$$

from (3.17), as well as the estimate

$$C_q := \sup_{x \in \mathbb{R}} |N_1(x \pm iq)| + |N_2(x \pm iq)| < \infty.$$

Now we can write

$$\widetilde{\mathcal{M}}^{(\nu)}(z) = \frac{1 + \nu z N_1(\nu z)}{c_0 + \tau_2 i z + \nu^2 z^3 N_2(\nu z)}.$$

With this expression we find the following equality

$$\widetilde{\mathcal{M}}^{(\nu)}(z) - \widetilde{\mathcal{M}}^0(z) = I_\nu(z) + II_\nu(z),$$

where

$$I_\nu(z) := \frac{c_0 \nu z N_1(\nu z) - \nu^2 z^3 N_2(\nu z)}{(c_0 + \tau_2 i z + \nu^2 z^3 N_2(\nu z))(c_0 + i \tau_2 z)} \tag{4.2}$$

and

$$II_\nu(z) := \frac{i \tau_2 \nu z^2 N_1(\nu z)}{(c_0 + \tau_2 i z + \nu^2 z^3 N_2(\nu z))(c_0 + i \tau_2 z)}. \tag{4.3}$$

We work on the denominators. We use the reverse triangle inequality to find

$$\left| (c_0 + \tau_2 i z - 2\tau_2 i \nu^2 z^3 N_2(\nu z))(c_0 + \tau_2 i z) \right| \geq \left| c_0 - \tau_2 q - 2\tau_2 \nu^2 |z|^3 |N_2(\nu z)| \right|$$
$$|c_0 - \tau_2 q|.$$

As $q \in (0, c_0/\tau_2)$, we have $|c_0 - \tau_2 q| > 0$. Also, since $|z| \leq \nu^{-p}$, we have $\nu^2 |z|^3 \leq \nu^{2-3p}$. If we take

$$0 < p < \frac{2}{3} \tag{4.4}$$

and assume $\nu \in (0, \nu_1)$, where

$$\nu_1 := \min \left\{ 1, \left(\frac{|c_0 - \tau_2 q|}{4 C_q \tau_2} \right)^{1/(2-3p)} \right\}, \tag{4.5}$$

then

$$\left| c_0 - \tau_2 q - \nu^2 |z|^3 |N_2(\nu z)| \right| \geq \frac{|c_0 - \tau_2 q|}{2}. \tag{4.6}$$

In particular,

$$\left| c_0 - \tau_2 q - 2\tau_2 \nu^2 |z|^3 |N_2(\nu z)| \right| |c_0 - \tau_2 q| \geq \frac{|c_0 - \tau_2 q|^2}{2}, \tag{4.7}$$

and this inequality guarantees that $\widetilde{\mathcal{M}}^{(\nu)}$ is defined (and analytic) for $|z| \leq \nu^{-p}$ and $|\mathrm{Im}(z)| < q$. Then we use (4.7) to estimate $I_\nu(z)$ from (4.2) as

$$|I_\nu(z)| \leq C\nu^{1-p} + C\nu^{2-3p}.$$

Next, we use (4.6) to estimate $II_\nu(z)$ from (4.3) as

$$|II_\nu(z)| \leq C\nu^{1-p} \frac{|z|}{|c_0 + i\tau_2 z|}.$$

Setting $z = x \pm iq$ we note

$$\frac{|z|^2}{|c_0 + i\tau_2 z|^2} = \frac{x^2 + q^2}{(c_0 \pm \tau_2 q)^2 + \tau_2^2 x^2} \leq \frac{x^2 + q^2}{(c_0 - \tau_2 q)^2 + \tau_2^2 x^2}.$$

We know

$$D := \sup_{x \in \mathbb{R}} \frac{x^2 + q^2}{(c_0 - \tau_2 q)^2 + c_0^2 x^2} < \infty,$$

and thus

$$|II_\nu(z)| \leq C\nu^{1-p}.$$

We conclude

$$\left| \widetilde{\mathcal{M}}^{(\nu)}(z) - \widetilde{\mathcal{M}}^{(0)}(z) \right| \leq |I_\nu(z)| + |II_\nu(z)| \leq C\left(\nu^{1-p} + \nu^{2-3p}\right). \qquad (4.8)$$

As we required $p \in (0, 2/3)$, the final estimate contains only positive powers of ν. Since we will always consider $0 < \nu < \nu_1$ in the future, the definition of ν_1 in (4.5) ensures $0 < \nu < 1$ in the following regimes.

4.2 Estimates for z 'far from' 0

In this regime we assume $|z| > \nu^p$. Take

$$\nu_2 < \min \left\{ \nu_1, \left(\frac{\tau_2}{2c_0} \right)^{1/p} \right\}, \qquad (4.9)$$

with ν_1 defined in (4.5), so that if $0 < \nu < \nu_2$, then $|z| > c_0/\tau_2$. With the reverse triangle inequality we find

$$|\widetilde{\mathcal{M}}^{(0)}(z)| \leq \frac{1}{\big||c_0| - |\tau_2 z|\big|} < \frac{1}{\tau_2 \nu^{-p} - c_0} < \frac{2}{\tau_2} \nu^p. \qquad (4.10)$$

Consequently, it suffices in this regime to show that $\widetilde{\mathcal{M}}^{(\nu)}$ is bounded by a multiple of some power of ν. It will be convenient now to rewrite $\widetilde{\mathcal{M}}^{(\nu)}$ as

$$\widetilde{\mathcal{M}}^{(\nu)}(z) = \frac{\nu \widetilde{\mathcal{M}}_1^{(\nu)}(z)}{\widetilde{\mathcal{M}}_2^{(\nu)}(z)},$$

where

$$\widetilde{\mathcal{M}}_1^{(\nu)}(z) := 1 - e^{-i\nu z} \quad \text{and} \quad \widetilde{\mathcal{M}}_2^{(\nu)}(z) := i c_0 \nu^2 z + 2\tau_2 (\cos(\nu z) - 1). \tag{4.11}$$

The analyticity of $\widetilde{\mathcal{M}}^{(\nu)}$ for $|z| > \nu^p$ will follow if we bound $\widetilde{\mathcal{M}}_2^{(\nu)}$ away from zero here.

The presence of the factor $\cos(\nu z) - 1$ in the denominator of $\widetilde{\mathcal{M}}^{(\nu)}$ suggests that the behavior of this function may be different when $\mathrm{Re}(\nu z)$ is 'close' to an integer multiple of 2π and when it is not. For this reason, we expand $z = x \pm iq$ and let $n \in \mathbb{Z}$ be the unique integer such that $|\nu x - 2\pi n| \leq \pi$. We consider three cases on the behavior of νx and n.

4.2.1 Estimates for Re(νz) 'close to' a nonzero integer multiple of 2π

In this regime we assume $|\nu x - 2\pi n| \leq \nu^m$ with $n \neq 0$.

We first rewrite the numerator as

$$\widetilde{\mathcal{M}}_1^{(\nu)}(z) = 1 - e^{-i(\nu x - 2\pi n)} + e^{-i\nu x}(1 - e^{\pm \nu q}).$$

Since the map $y \mapsto e^{-iy}$ is uniformly Lipschitz on $\mathbb{R}$ we have

$$|1 - e^{-i(\nu x - 2\pi n)}| \leq |\nu x - 2\pi n| \leq \nu^m.$$

Since the map $y \mapsto e^{-y}$ is locally Lipschitz on $\mathbb{R}$ we have, if we take $0 < \nu < \nu_3$ with

$$\nu_3 := \min\left\{\nu_2, \frac{1}{q}\right\} \tag{4.12}$$

and ν_2 defined in (4.9), the estimate

$$|1 - e^{\pm \nu q}| \leq \nu q.$$

Then

$$\left|\widetilde{\mathcal{M}}_1^{(\nu)}(z)\right| \leq \nu^m + \nu q \leq C(\nu^m + \nu). \tag{4.13}$$

We remark that we did not need $n \neq 0$ here, although we will momentarily.

We now turn to the denominator, $\mathcal{M}_2^{(\nu)}(z)$. Using the identity

$$\cos(a + bi) = \cos(a)\cosh(b) - i\sin(a)\sinh(b) \tag{4.14}$$

for $a, b \in \mathbb{R}$ we find

$$\text{Im}\left(\widetilde{\mathcal{M}}_1^{(\nu)}(z)\right) = c_0\nu^2 x - 2\tau_2 \sin(\nu x)\sinh(\nu q).$$

We estimate

$$\left|\text{Im}\left(\widetilde{\mathcal{M}}_1^{(\nu)}(z)\right)\right| \geq C\left(\nu|n| - \nu|\nu x - 2n\pi| - |\sin(\nu x - 2n\pi)||\sinh(\nu q)|\right)$$

We control the three terms on the right as follows. First, $|n| \geq 1$. Next, we are in the regime $|\nu x - 2n\pi| \leq \nu^m$. Finally, we have

$$|\sin(\nu x - 2n\pi)| \leq |\nu x - 2n\pi| \leq \nu^m \quad \text{and} \quad |\sinh(\nu q)| \leq 2|\nu q|,$$

since $|\nu q| \leq 1$. We thus find

$$\left|\text{Im}\left(\widetilde{\mathcal{M}}_2^{(\nu)}(z)\right)\right| \geq C\left(\nu - \nu^{m+1}\right).$$

Now that we have the numerator and the denominator bounded, we can conclude

$$\left|\widetilde{\mathcal{M}}^{(\nu)}(z)\right| \leq C\nu\frac{\nu^m + \nu}{\nu - \nu^{m+1}} = C\frac{\nu^m + \nu}{1 - \nu^m} \leq C\nu^m. \tag{4.15}$$

Here we need to assume

$$0 < m < 1. \tag{4.16}$$

4.2.2 Estimates for Re(νz) 'close to' 0

In this regime we assume $|\nu x| \leq \nu^m$; in particular, we are taking $n = 0$. We will need the following bound on the cosine, which is a consequence of an elementary argument with Taylor's theorem.

Lemma 4.3 *Let $Q \geq 0$. There exist $C_{1,Q}, C_{2,Q} > 0$ such that if $Z \in \mathbb{C}$ with $|Z| \leq C_{1,Q}$ and $|\text{Im}(Z)| \leq Q$, then*

$$|\cos(Z) - 1| \geq C_{2,Q}|Z|^2.$$

In particular, if $Q = 0$, then $C_{1,0} > \pi$.

We use the reverse triangle inequality on $\widetilde{\mathcal{M}}_2^{(\nu)}$ from (4.11) to find

$$\left|\widetilde{\mathcal{M}}_2^{(\nu)}(z)\right| \geq 2\tau_2|\cos(\nu z) - 1| - c_0\nu^2 q - c_0\nu^2|x|. \tag{4.17}$$

Take $0 < v < v_{\mathcal{M}}$, where

$$v_{\mathcal{M}} < \min\left\{ v_3, \left(\frac{1}{q}\right)^{1/(1-m)}, \left(\frac{C_{1,q}}{2}\right)^{1/m} \right\}, \tag{4.18}$$

with v_3 defined in (4.12), to find

$$|vz| \leq v^m + vz \leq 2v^m < C_{1,q}.$$

Lemma 4.3 then guarantees

$$|\cos(vz) - 1| \geq C|vz|^2 \geq Cv^{2-2p}.$$

Finally, since $|x| \leq v^{m-1}$ in this regime we use the bound (4.17) to conclude

$$\left|\widetilde{\mathcal{M}}_2^{(v)}(z)\right| \geq C\left(v^{2-2p} - v^2 - v^{m+1}\right).$$

We remark that the derivation of the estimate (4.13) only assumed $|vx - 2\pi n| \leq v^m$ and did not rely on having $n \neq 0$. So it is still valid here, and we conclude

$$\left|\widetilde{\mathcal{M}}^{(v)}(z)\right| \leq Cv\frac{v^m + v}{v^{2-2p} - v^2 - v^{m+1}} = C\frac{v^{m+2p-1} + v^{2p}}{1 - v^{2p} - v^{m+2p-1}}$$
$$\leq C\left(v^{m+2p-1} + v^{2p}\right). \tag{4.19}$$

Here we are assuming

$$0 \leq 1 - 2p < \min\{1, m\}. \tag{4.20}$$

4.2.3 Estimates for Re(vz) 'far from' a nonzero integer multiple of 2π

In this regime we assume $|vx - 2\pi n| > v^m$. We do not perform separate work on $n = 0$ and $n \neq 0$.

Via (4.14) we find

$$\mathrm{Re}\left(\mathcal{M}_2^{(v)}(z)\right) = -c_0 v^2 q + 2\tau_2(\cos(vx - 2n\pi) - 1) + 2\tau_2\cos(vx)(\cosh(vq) - 1).$$

We estimate

$$\left|\mathrm{Re}\left(\mathcal{M}_2^{(v)}(z)\right)\right| \geq C\left(|\cos(vx - 2n\pi) - 1| - |\cos(vx)||\cosh(vq) - 1| - v^2\right).$$

Now we use Lemma 4.3 with $Q = 0$ to bound

$$|\cos(vx - 2n\pi) - 1| \geq C|vz - 2n\pi|^2 \geq Cv^{2m}.$$

Also, a routine Lipschitz estimate on the hyperbolic cosine gives

$$|\cos(\nu x)||\cosh(\nu q) - 1| \leq C\nu^2$$

since $|\nu q| \leq 1$. We thus find

$$\left|\operatorname{Re}\left(\mathcal{M}_2^{(\nu)}(z)\right)\right| \geq C(\nu^{2m} - \nu^2).$$

As we are assuming $0 < m < 1$ from (4.16), this is a positive lower bound.

Finally, we bound the numerator $\widetilde{\mathcal{M}}_1^\nu(z)$ crudely as $\left|\widetilde{\mathcal{M}}_1^{(\nu)}(z)\right| \leq C$ for all $z \in \mathbb{C}$ with $|\operatorname{Im}(z)| = q$. This follows from the boundedness of $Z \mapsto e^{iZ}$ on strips. We conclude

$$\left|\widetilde{\mathcal{M}}^{(\nu)}(z)\right| < C\frac{\nu}{\nu^{2m} - \nu^2} \leq C\nu^{1-2m}. \tag{4.21}$$

This is a positive bound if we now require

$$0 < m < \frac{1}{2}. \tag{4.22}$$

4.3 Overall estimates

Suppose $0 < \nu < \nu_{\mathcal{M}}$, where $\nu_{\mathcal{M}}$ was specified in (4.18). We conclude from (4.8) that

$$\sup_{\substack{|\operatorname{Im}(z)|=q \\ |z| \leq \nu^{-p}}} \left|\widetilde{\mathcal{M}}^{(\nu)}(z) - \widetilde{\mathcal{M}}^{(0)}(z)\right| \leq C\left(\nu^{1-p} + \nu^{2-3p}\right) \tag{4.23}$$

and, by combining (4.10), (4.15), (4.19), and (4.21), that

$$\sup_{\substack{|\operatorname{Im}(z)|=q \\ |z| > \nu^{-p}}} \left|\widetilde{\mathcal{M}}^{(\nu)}(z) - \widetilde{\mathcal{M}}^{(0)}(z)\right| \leq C\nu^p + C\max\left\{\nu^m, \nu^{m+2p-1} + \nu^{2p}, \nu^{1-2m}\right\}. \tag{4.24}$$

Additionally, we need, per (4.4), and (4.20), and (4.22), the exponents p and m to satisfy

$$0 < p < \frac{2}{3}, \quad 0 < m < \frac{1}{2}, \quad \text{and} \quad 0 < 1 - 2p < \min\{1, m\}. \tag{4.25}$$

There are many possible choices of p and m that will satisfy (4.25). Purely for convenience, we elect to take $m = 1/3$ and then $p = 1/2$. We combine (4.23) and (4.24) to conclude the estimate (4.1).

5 Analysis of the linearization $\mathcal{T}$

The operator

$$\mathcal{T}\psi = \psi - \mathcal{M}^{(0)}\big[\mathcal{R}_1^0(\sigma)\psi + (D\mathcal{R}_1^0(\sigma)\psi)\sigma\big] \tag{5.1}$$

is the linearization of the first equation in our long wave-scaled travelling wave problem (3.42) at $(\psi_1, \psi_2) = (\sigma, \zeta)$ and $\nu = 0$. We recall that the symbol of the Fourier multiplier $\mathcal{M}^{(0)}$ was defined in (3.21) and the operator $\mathcal{R}_1^0$ was defined in (3.39).

We will work with $\mathcal{T}$ defined on the following subspace of H_q^1:

$$H_{q,0}^1 := \left\{ f \in H_q^1 \mid \int_0^\infty f(W)\, dW = 0 \right\}. \tag{5.2}$$

We norm $H_{q,0}^1$ with the H_q^1-norm. In Lemma 5.2 below we show that the kernel of $\mathcal{T}$ in H_q^1 is spanned by σ'. Restricting $\mathcal{T}$ to $H_{q,0}^1$ removes this kernel and guarantees injectivity. We will then prove that $\mathcal{T}$ is surjective onto H_q^1 and so conclude the following result.

Proposition 5.1 *For $q \in (0, c_0/\tau_2)$, the operator $\mathcal{T}: H_{q,0}^1 \to H_q^1$ is invertible with bounded inverse.*

The linearization at the limiting localized solution appears as a key operator in numerous FPUT problems, including (Friesecke and Pego 1999; Faver and Wright 2018; Hoffman and Wright 2017), and the invertibility of this operator is a property essential to the development of the right fixed point formula for the given problem. Our treatment of the invertibility of $\mathcal{T}$ is rather different from the analogous inversions in those papers, as the problem $\mathcal{T}f = g$ is really a linearized Bernoulli equation in disguise, rather than the linearized KdV travelling wave profile equation. In particular, solving $\mathcal{T}f = g$ amounts to studying a first-order linear problem, which we can solve explicitly with an integrating factor. In doing so, we avoid the more abstract spectral theory that controls the second-order KdV linearizations (see, e.g., Friesecke and Pego (1999, Lem. 4.2)).

It will be convenient to abbreviate

$$q_* := \frac{c_0}{\tau_2}, \tag{5.3}$$

and in the following we always assume $0 < q < q_*$.

5.1 The proof of Proposition 5.1

We will reformulate the equation $\mathcal{T}f = g$ in a very convenient manner, which we summarize in Lemma 5.2 below. This lemma will be the key to deducing the injectivity and surjectivity of $\mathcal{T}$.

First, for $h \in H_q^1$, we introduce the operator

$$(\mathcal{A}f)(X) := \int_X^\infty h(W) \, dW, \tag{5.4}$$

so that

$$h = -\partial_X[\mathcal{A}h]. \tag{5.5}$$

Then $h \in H_{q,0}^1$ if and only if $h \in H_q^1$ and $(\mathcal{A}f)(0) = 0$. We will show that the equation $\mathcal{T}f = g$ for $f, g \in H_q^1$ is equivalent to a statement about $\mathcal{A}f$ and $\mathcal{A}g$, which we give precisely in Lemma 5.2.

The following steps are quite similar to the derivation of the Bernoulli solution σ in Sect. 3.2.2. From the definition of $\mathcal{T}$ in (5.1), we have $\mathcal{T}f = g$ if and only if

$$(c_0 + \tau_2 \partial_X) f - \mathcal{R}_1^0(\sigma) f + \left(D\mathcal{R}_1^0(\sigma) f \right)\sigma = g. \tag{5.6}$$

From the definition of $\mathcal{R}_1^0$ in (3.39), we find

$$\left(D\mathcal{R}^0(\sigma)f\right)(X) = \int_X^\infty \left(\int_W^\infty \sigma(V) \, dV \right) f(W) \, dW$$
$$+ \int_X^\infty \left(\int_W^\infty f(V) \, dV \right) \sigma(W) \, dW. \tag{5.7}$$

Since $g, \sigma \in H_q^1$, and since we seek $f \in H_q^1$, we abbreviate

$$F := \mathcal{A}f, \quad G := \mathcal{A}g, \quad \text{and} \quad \Sigma := \mathcal{A}\sigma. \tag{5.8}$$

Then (5.6) is equivalent to

$$\tau_2 F''(X) - c_0 F'(X) - \left(\frac{\alpha\kappa\tau_1}{c_0^2} \right) F'(X) \int_X^\infty \Sigma(W)\Sigma'(W) \, dW$$
$$- \left(\frac{\alpha\kappa\tau_1}{c_0^2} \right) \Sigma'(X) \int_X^\infty \left(\Sigma(W)F'(W) + F(W)\Sigma'(W) \right) dW$$
$$= -c_0 G'(X) - \tau_2 G''(X). \tag{5.9}$$

Although it may not be apparent at first glance, every term in this equation is a perfect derivative. First, since Σ and F must vanish at $+\infty$, we have

$$\int_X^\infty \Sigma(W)\Sigma'(W) \, dW = -\frac{\Sigma(X)^2}{2}$$

and

$$\int_X^\infty \left(\Sigma(W) F'(W) + F(W) \Sigma'(W) \right) dW = -\Sigma(X) F(X).$$

Hence (5.9) really is

$$\tau_2 F'' - c_0 F' + \left(\frac{\alpha \kappa \tau_1}{c_0^2} \right) \left(\frac{F' \Sigma^2}{2} + \Sigma' \Sigma F \right) = -c_0 G' - \tau_2 G'', \qquad (5.10)$$

where

$$\frac{F' \Sigma^2}{2} + \Sigma' \Sigma F = \frac{1}{2} \partial_X [\Sigma^2 F].$$

So, we deduce that F and G must satisfy

$$- \tau_2 F'' - c_0 F' + \left(\frac{\alpha \kappa \tau_1}{c_0^2} \right) \partial_X \left[\frac{\Sigma^2 F}{2} \right] = -c_0 G' - \tau_2 G''. \qquad (5.11)$$

Since both F and G must vanish at $+\infty$, we may integrate (5.11) to find

$$\underbrace{\tau_2 F' + c_0 F - \left(\frac{\alpha \kappa \tau_1}{c_0^2} \right) \frac{\Sigma^2 F}{2}}_{\mathcal{L}F} = c_0 G + \tau_2 G'. \qquad (5.12)$$

The operator $\mathcal{L}$ defined above is the linearization of the Bernoulli equation (3.26) at its solution Σ, and so

$$0 = \mathcal{L}\Sigma' = \mathcal{L}(-\sigma). \qquad (5.13)$$

The operator $\mathcal{L}$, or, more precisely, $\tau_2^{-1}\mathcal{L}$, is also a first-order linear differential operator, and so we can solve (5.12) with an integrating factor. Namely, let P satisfy

$$P' = \frac{c_0}{\tau_2} - \left(\frac{\alpha \kappa \tau_1}{2c_0^2 \tau_2} \right) \Sigma^2. \qquad (5.14)$$

Then F solves (5.12) if and only if

$$F(X) = F(0) e^{P(0)-P(X)} + \sigma(X) \int_0^X e^{P(W)} \left(q_* G(W) + G'(W) \right) dW. \qquad (5.15)$$

In particular, any solution H to $\mathcal{L}H = 0$ must be a scalar multiple of $e^{-P(\cdot)}$, and so, by (5.13), σ is also a scalar multiple of $e^{-P(\cdot)}$. Consequently, we can rewrite (5.15)

as

$$F(X) = \frac{F(0)}{\sigma(0)}\sigma(X) + \sigma(X)\int_0^X \frac{q_* G(W) + G'(W)}{\sigma(W)}\,dW. \qquad (5.16)$$

Conversely, if F satisfies (5.16), then we may undo all of the work above to see that $f := -F'$ solves $\mathcal{T}f = g$. Using the identities (5.8), we can recast this result in terms of the original functions f and g.

Lemma 5.2 *For $g \in H_q^1$, define*

$$(\mathcal{H}g)(X) := q_*(\mathcal{A}g)(X) - g(X) \qquad (5.17)$$

and

$$(\mathcal{K}g)(X) := \sigma(X)\int_0^X \frac{(\mathcal{H}g)(W)}{\sigma(W)}\,dW. \qquad (5.18)$$

Then $f \in H_q^1$ satisfies $\mathcal{T}f = g$ if and only if

$$(\mathcal{A}f)(X) = \frac{(\mathcal{A}f)(0)}{\sigma(0)}\sigma(X) + (\mathcal{K}g)(X). \qquad (5.19)$$

In particular, a function $f \in H_{q,0}^1$ satisfies $\mathcal{T}f = g$ if and only if

$$(\mathcal{A}f)(X) = (\mathcal{K}g)(X). \qquad (5.20)$$

The identity (5.20) allows to prove the bijectivity of $\mathcal{T}\colon H_{q,0}^1 \to H_q^1$. The proof of injectivity is very easy. If $\mathcal{T}f = 0$ for some $f \in H_{q,0}^1$, then (5.20) implies $\mathcal{A}f = 0$, and so the identity (5.5) gives $f = 0$. Observe that if we were working on all of H_q^1, then (5.19) tells us that $\mathcal{T}$ would have a one-dimensional kernel in H_q^1 spanned by σ'. But since $(\mathcal{A}\sigma')(0) = -\sigma(0) \neq 0$, we have $\sigma' \notin H_{q,0}^1$.

Toward surjectivity, suppose $\mathcal{T}f = g$ for some $f \in H_{q,0}^1$ and $g \in H_q^1$. Then (5.5) and (5.20) imply

$$f = -\partial_X[\mathcal{K}g] =: \mathcal{S}g. \qquad (5.21)$$

That is, we expect $\mathcal{T}^{-1} = \mathcal{S}$. Now we make this rigorous.

Lemma 5.3 *The operator $\mathcal{S}$, defined in (5.21), is a bounded linear operator from H_q^1 to $H_{q,0}^1$ that satisfies $\mathcal{T}\mathcal{S}g = g$ for all $g \in H_q^1$.*

Proof Let $g \in H_q^1$. In part (i) of Lemma 5.4 below we show that $\sigma' = -\rho\sigma$ for a certain function $\rho \in L^\infty$. Then the definition of $\mathcal{K}$ in (5.18) gives

$$\mathcal{S}g = \rho(\mathcal{K}g) - \mathcal{H}g, \qquad (5.22)$$

and the definition of $\mathcal{H}$ in (5.17) shows

$$\partial_X[\mathcal{S}g] = \rho'(\mathcal{K}g) - \rho(\mathcal{S}g) + q_*g + g'. \tag{5.23}$$

We claim there is a constant C, independent of g, such that

$$\|\rho'(\mathcal{K}g)\|_{L_q^2} \leq C\|g\|_{H_q^1} \tag{5.24}$$

and

$$\|\mathcal{S}g\|_{L_q^2} \leq C\|g\|_{H_q^1} \tag{5.25}$$

Since $\rho \in L^\infty$, the identities (5.22) and (5.23) show that $\mathcal{S}$ is a bounded operator on H_q^1. We prove the estimate (5.24) in Sect. 5.2.1 below and the estimate (5.25) in Sect. 5.2.2.

To show both that $\mathcal{S}g \in H_{q,0}^1$ and that taking $f = \mathcal{S}g$ satisfies (5.20), we first use the definition of $\mathcal{A}$ in (5.4) to compute

$$(\mathcal{A}\mathcal{S}g)(X) = (\mathcal{K}g)(X) - \lim_{B\to\infty} (\mathcal{K}g)(B). \tag{5.26}$$

We claim that

$$\lim_{B\to\infty} (\mathcal{K}g)(B) = 0. \tag{5.27}$$

Indeed, since $\mathcal{S}g \in H_q^1$, we know that $\mathcal{S}g$ vanishes at infinity; so does $\mathcal{H}g$ by the definition of $\mathcal{H}$ in (5.17). However, $\rho(X) \to q_* \neq 0$ as $X \to \infty$ by part (iii) of Lemma 5.4 below. The first equality in (5.22) then forces the limit (5.27) to be true.

Thus

$$(\mathcal{A}\mathcal{S}g)(X) = (\mathcal{K}g)(X). \tag{5.28}$$

In particular,

$$(\mathcal{A}\mathcal{S}g)(0) = (\mathcal{K}g)(0) = 0$$

by the definition of $\mathcal{K}$ in (5.18). Consequently, $\mathcal{S}g \in H_{q,0}^1$, and so (5.28) shows that $f = \mathcal{S}g$ satisfies (5.20). This implies $\mathcal{T}(\mathcal{S}g) = g$. $\square$

5.2 Auxiliary results for the proof of Lemma 5.3

We first study some properties of σ and its derivative.

Lemma 5.4 (i) *There exists $\rho \in L^\infty$ such that $\sigma' = -\rho\sigma$.*

(ii) *There exist* ς_1^+, ς_2^+, $\varrho^+ \in L^\infty(\mathbb{R}_+)$, ς_1^-, ς_2^-, $\varrho^- \in L^\infty(\mathbb{R}_-)$, *and* C_1, $C_2 \in \mathbb{R}$ *such that*

$$\frac{1}{\sigma(X)} = \begin{cases} C_1 e^{q_* X} + e^{-q_* X} \varsigma_1^+(X), & X > 0 \\ C_2 e^{-2q_* X} + \varsigma_1^-(X), & X < 0, \end{cases} \tag{5.29}$$

$$\sigma(X) = \begin{cases} C_1^{-1} e^{-q_* X} + e^{-3q_* X} \varsigma_2^+(X), & X > 0 \\ C_2^{-1} e^{2q_* X} + e^{4q_* X} \varsigma_2^-(X), & X < 0, \end{cases} \tag{5.30}$$

and

$$\rho(X) = \begin{cases} q_* + e^{-q_* X} \varrho^+(X), & X > 0 \\ -2q_* + e^{q_* X} \varrho^-(X), & X < 0. \end{cases} \tag{5.31}$$

(iii) *There is* $C_3 > 0$ *such that*

$$|\rho'(X)| \leq C_3 e^{-q_* |X|} \tag{5.32}$$

for all $X \in \mathbb{R}$.

Proof (i) Recall that $\sigma = -\Sigma'$, where Σ satisfies the Bernoulli equation (3.26). That is,

$$\sigma = -\Sigma' = \frac{c_0}{\tau_2} \Sigma - \left(\frac{\alpha \kappa \tau_1}{6 c_0^2 \tau_2} \right) \Sigma^3.$$

Then

$$\sigma' = \frac{c_0}{\tau_2} \Sigma' - \left(\frac{\alpha \kappa \tau_1}{2 c_0^2 \tau_2} \right) \Sigma^2 \Sigma' = \left[\left(\frac{\alpha \kappa \tau_1}{2 c_0^2 \tau_2} \right) \Sigma^2 - \frac{c_0}{\tau_2} \right] \sigma.$$

Put

$$\rho = \frac{c_0}{\tau_2} - \left(\frac{\alpha \kappa \tau_1}{2 c_0^2 \tau_2} \right) \Sigma^2. \tag{5.33}$$

By the definition of Σ in (3.27), we have $\rho \in L^\infty$. Note, incidentally, that ρ must be a scalar multiple of P from (5.14).

(ii) The expansions (5.30) and (5.29) follow directly from the formula for σ in (3.28). The expansion (5.31) follows from the definition of ρ in (5.33) and the definition of Σ in (3.27), which gives

$$\lim_{X \to \infty} \Sigma(X)^2 = 0 \quad \text{and} \quad \lim_{X \to -\infty} \Sigma(X)^2 = \frac{3 c_0}{\tau_2} = 3 q_*.$$

(iii) This is a direct consequence of (5.33).

<div align="right">□</div>

We will also need estimates on $\mathcal{A}$ and $\mathcal{H}$.

Lemma 5.5 *There is $C > 0$ such that*

$$\|\mathcal{A}g\|_{L^\infty} + \|\mathcal{H}g\|_{L^\infty} \leq C\|g\|_{H_q^1} \tag{5.34}$$

for all $g \in H_q^1$.

Proof We use the definition of $\mathcal{A}$ in (5.4) to bound

$$|(\mathcal{A}g)(X)| \leq \|f\|_{L^1} \leq C\|f\|_{H_q^1}$$

by the embedding of H_q^1 into L^1, which we discuss in Appendix A.3. The estimate for $\mathcal{H}$ then follows from the triangle inequality. □

In the following we again recall that $0 < q < q_*$.

5.2.1 The proof of the estimate (5.24)

We first use the definition of $\mathcal{K}$ in (5.18) and the estimates (5.32) on ρ' and (5.34) on $\mathcal{H}g$ to bound

$$|\rho'(X)(\mathcal{K}g)(X)| \leq C\|g\|_{H_q^1}e^{-q_*|X|}\sigma(X)\int_0^X \frac{dW}{\sigma(W)}.$$

If $X > 0$, we use the estimates (5.30) on $\sigma(X)$ and (5.29) on $1/\sigma(W)$ to bound

$$\sigma(X)\int_0^X \frac{dW}{\sigma(W)} \leq Ce^{-q_*X}\int_0^X e^{q_*W}\, dW \leq C.$$

If $X < 0$, we use the negative versions of these estimates to bound

$$\sigma(X)\left|\int_0^X \frac{dW}{\sigma(W)}\right| \leq Ce^{2q_*X}\int_X^0 e^{-2q_*W} \leq C.$$

We conclude

$$|\rho'(X)(\mathcal{K}g)(X)| \leq C\|g\|_{H_q^1}e^{-q_*|X|}$$

for all X. Since $0 < q < q_*$, this gives $\|\rho'(\mathcal{K}g)\|_{L_q^2} \leq C\|g\|_{H_q^1}$.

5.2.2 The proof of the estimate (5.25)

It suffices to find $C > 0$ such that for all $g \in H_q^1$, we have

$$\|\mathcal{S}g\|_{L_q^2(\mathbb{R}_+)} + \|\mathcal{S}g\|_{L_q^2(\mathbb{R}_-)} \leq C\|g\|_{H_q^1}.$$

We will rewrite $(\mathcal{S}g)(X)$ in different ways for $X > 0$ and $X < 0$ to exploit the different decay rates of σ at $+\infty$ and $-\infty$.

First suppose $X > 0$. The formula (5.22) for $\mathcal{S}g$, the formula (5.18) for $\mathcal{K}$ and the expansions in Lemma 5.4 allow us to write

$$(\mathcal{S}g)(X) = \left(C_1^{-1}q_* e^{-q_* X} + e^{-2q_* X}\varsigma_3^+(X)\right)$$
$$\int_0^X \left(C_1 e^{q_* W} + e^{-q_* W}\varsigma_1^+(W)\right)(\mathcal{H}g)(W)\, dW - (\mathcal{H}g)(X),$$

where $\varsigma_3^+ \in L^\infty(\mathbb{R}_+)$. We expand this to give

$$(\mathcal{S}g)(X) = \sum_{j=1}^4 (\mathcal{S}_j^+ g)(X),$$

where

$$(\mathcal{S}_1^+ g)(X) := q_* e^{-q_* X} \int_0^X e^{q_* W}(\mathcal{H}g)(W)\, dW - (\mathcal{H}g)(X)$$

$$(\mathcal{S}_2^+ g)(X) := C_1^{-1} q_* e^{-q_* X} \int_0^X e^{-q_* W}\varsigma_1^+(W)(\mathcal{H}g)(W)\, dW$$

$$(\mathcal{S}_3^+ g)(X) := C_1 e^{-2q_* X}\varsigma_3^+(X) \int_0^X e^{q_* W}(\mathcal{H}g)(W)\, dW$$

$$(\mathcal{S}_4^+ g)(X) := e^{-2q_* X}\varsigma_3^+(X) \int_0^X e^{-q_* W}\varsigma_1^+(W)(\mathcal{H}g)(W)\, dW.$$

The estimate (5.34) from Lemma 5.5 allows us to bound the last three terms with

$$|(\mathcal{S}_2^+ g)(X)| + |(\mathcal{S}_4^+ g)(X)| \leq C e^{-q_* X}\|\mathcal{H}g\|_{L^\infty} \int_0^X e^{-q_* W} \leq C e^{-q_* X}\|g\|_{H_q^1} \tag{5.35}$$

and

$$|(\mathcal{S}_3^+ g)(X)| \leq C e^{-2q_* X}\|\mathcal{H}g\|_{L^\infty} \int_0^X e^{q_* W} \leq C e^{-q_* X}\|g\|_{H_q^1}. \tag{5.36}$$

To control $\mathcal{S}_1^+ g$, we first integrate by parts:

$$\int_0^X e^{q_* W}(\mathcal{H}g)(W)\, dW = \frac{e^{q_* X}(\mathcal{H}g)(X) - (\mathcal{H}g)(0)}{q_*} - \frac{1}{q_*}\int_0^X e^{q_* W}(\mathcal{H}g)'(W)\, dW.$$

The definition of $\mathcal{H}$ in (5.17) gives

$$\int_0^X e^{q_* W}(\mathcal{H}g)'(W)\, dW$$
$$= -\int_0^X e^{q_* W}(q_* g(W) + g'(W))\, dW = -\int_0^X \partial_W[e^{q_* W}g(W)]\, dW$$
$$= g(0) - e^{q_* X}g(X). \tag{5.37}$$

It follows that

$$(\mathcal{S}_1^+ g)(X) = -e^{-q_* X}(\mathcal{H}g)(0) - e^{-q_* X}g(0) + g(X). \tag{5.38}$$

We apply Lemma 5.5 to the factor $(\mathcal{H}g)(0)$ and use the Sobolev embedding to estimate $g(0)$, so that

$$|(\mathcal{S}_1^+ g)(X)| \leq Ce^{-q_* X}\|g\|_{H_q^1} + |g(X)|.$$

Since $0 < q < q_*$, the estimates (5.35) and (5.36) and the identity (5.38) give

$$e^{qX}|(\mathcal{S}g)(X)| \leq Ce^{(q - q_*)X}\|g\|_{H_q^1} + e^{qX}|g(X)|$$

for $X > 0$, from which the bound

$$\|\mathcal{S}g\|_{L_q^2(\mathbb{R}_+)} \leq C\|g\|_{H_q^1}$$

follows.

Now suppose $X < 0$. Using the expansions in Lemma 5.4 valid for $X < 0$, we rewrite

$$(\mathcal{S}g)(X) = \left(2q_* C_2^{-1}e^{2q_* X} + e^{3q_* X}\varsigma_3^-(X)\right)\int_X^0 \left(C_2 e^{-2q_* W} + \varsigma_1^-(W)\right)(\mathcal{H}g)(W)$$
$$- (\mathcal{H}g)(X),$$

where $\varsigma_3^- \in L^\infty(\mathbb{R}_-)$. We expand this as

$$(\mathcal{S}g)(X) = \sum_{j=1}^4 (\mathcal{S}_j^- g)(X),$$

where

$$(\mathcal{S}_1^- g)(X) := 2q_* e^{2q_* X} \int_X^0 e^{-2q_* W} (\mathcal{H}g)(W) \, dW - (\mathcal{H}g)(X)$$

$$(\mathcal{S}_2^- g)(X) := 2q_* C_2^{-1} e^{2q_* X} \int_X^0 \varsigma_1^-(W)(\mathcal{H}g)(W) \, dW$$

$$(\mathcal{S}_3^- g)(X) := C_2 e^{3q_* X} \varsigma_3^-(X) \int_X^0 e^{-2q_* W} (\mathcal{H}g)(W) \, dW$$

$$(\mathcal{S}_4^- g)(X) := e^{3q_* X} \varsigma_3^-(X) \int_X^0 \varsigma_1^-(W)(\mathcal{H}g)(W) \, dW.$$

We crudely estimate the last three terms as

$$|(\mathcal{S}_2^- g)(X)| + |(\mathcal{S}_3^- g)(X)| + |(\mathcal{S}_4^- g)(X)| \leq C e^{2q_* X} |X| \|\mathcal{H}g\|_{L^\infty}$$
$$\leq C e^{q_* X} \|g\|_{H_q^1}. \tag{5.39}$$

For the first term, we integrate by parts to find

$$\int_X^0 e^{-2q_* W} (\mathcal{H}g)(W) \, dW = \frac{e^{-2q_* X} (\mathcal{H}g)(X) - (\mathcal{H}g)(0)}{2r}$$
$$- \frac{1}{2q_*} \int_X^0 e^{-2q_* W} (q_* g(W) + g'(W)) \, dW.$$

The difference compared to (5.37) in our treatment of $\mathcal{S}_1^+ g$ is that we no longer have a perfect derivative as the integrand on the right; this is an consequence of the different asymptotic behavior of ρ and σ at $-\infty$ compared to $+\infty$, as specified in Lemma 5.4. Thus

$$(\mathcal{S}_1^- g)(X) = -(\mathcal{H}g)(0) e^{2q_* X} + \mathcal{I}[g](X), \tag{5.40}$$

where

$$\mathcal{I}[g](X) := -e^{2q_* X} \int_X^0 e^{-2q_* W} \left(r g(W) + g'(W) \right) \, dW.$$

To control this integral term, we will use the following lemma, whose proof we defer to Sect. 5.2.3.

Lemma 5.6 *There exists $C > 0$ such that*

$$\int_{-\infty}^0 e^{2X} \left| \int_X^0 e^{-W} h(W) \, dW \right|^2 \, dX \leq C \|h\|_{L^2} \tag{5.41}$$

for all $h \in L^2$.

Since $q_* g + g' \in L_q^2$, we can write

$$q_* g(X) + g'(X) = e^{-q|X|} h(X)$$

for some $h \in L^2$. Then

$$
\int_{-\infty}^{0} e^{-2qX} |\mathcal{I}[g](X)|^2 \, dX = \int_{-\infty}^{0} e^{2(2q_*-q)X} \left| \int_{X}^{0} e^{-(2q_*-q)W} h(W) \, dW \right|^2 \, dX
$$

$$
= \frac{1}{(2q_* - q)^2} \int_{-\infty}^{0} e^{2U} \left| \int_{U}^{0} e^{-V} h\left(\frac{V}{2q_* - q} \right) \, dV \right|^2 \, dU.
$$

Applying Lemma 5.6, we obtain

$$
\|\mathcal{I}[g]\|_{L_q^2(\mathbb{R}_-)} \le C \left\| h\left(\frac{\cdot}{2q_* - q} \right) \right\|_{L^2} \le C \| e^{-q|\cdot|} (e^{q|\cdot|} h) \|_{L^2} \le C \| g \|_{H_q^1}.
$$

All together, we use the estimates (5.39) and the identity (5.40) to bound

$$
|(\mathcal{S}g)(X)| \le C e^{q_* X} \| g \|_{H_q^1} + |\mathcal{I}[g](X)|,
$$

from which we obtain

$$
\|\mathcal{S}g\|_{L_q^2(\mathbb{R}_-)} \le C \| g \|_{H_q^1}.
$$

5.2.3 The proof of Lemma 5.6

Put

$$
\mathcal{W} := \left\{ (X, W, Y) \in \mathbb{R}^3 \mid -\infty < X \le 0, \ X \le W \le 0, \ X \le Y \le 0 \right\},
$$

so that, after using the triangle inequality, the integral in (5.41) is bounded by

$$
\mathcal{J} := \int_{-\infty}^{0} e^{2X} \left(\int_{X}^{0} e^{-W} |h(W)| \, dW \right)^2 \, dX
$$

$$
= \iiint_{\mathcal{W}} e^{2X} e^{-W} e^{-Y} |h(W) h(Y)| \, dY \, dW \, dX.
$$

Next, put

$$
\mathcal{W}_1 := \left\{ (X, W, Y) \in \mathbb{R}^3 \mid -\infty < X \le W, \ W \le Y \le 0, \ -\infty < W \le 0 \right\}
$$

and

$$
\mathcal{W}_2 := \left\{ (X, W, Y) \in \mathbb{R}^3 \mid -\infty < X \le Y, \ Y \le W \le 0, \ -\infty < Y \le 0 \right\},
$$

so $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2$ and $\mathcal{W}_1 \cap \mathcal{W}_2$ has measure zero. Then

$$\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2,$$

where

$$\mathcal{J}_1 := \iiint_{\mathcal{W}_1} e^{2X} e^{-W} e^{-Y} |h(W)h(Y)| \, dY \, dW \, dX$$

and

$$\mathcal{J}_2 := \iiint_{\mathcal{W}_2} e^{2X} e^{-W} e^{-Y} |h(W)h(Y)| \, dY \, dW \, dX.$$

Since the integrands are symmetric in W and Y, it suffices to show

$$\mathcal{J}_1 \le C \int_{-\infty}^{0} |h(X)|^2 \, dX.$$

Change variables to obtain

$$\mathcal{J}_1 = \int_{-\infty}^{0} \int_{W}^{0} \left(\int_{-\infty}^{W} e^{2X} \, dX \right) e^{-W} e^{-Y} |h(W)h(Y)| \, dY \, dW$$
$$= \frac{1}{2} \int_{-\infty}^{0} \int_{W}^{0} e^{W} e^{-Y} |h(W)h(Y)| \, dY \, dW.$$

Now we estimate

$$4|\mathcal{J}_1| \le \mathcal{J}_{12} + \mathcal{J}_{13}, \tag{5.42}$$

where

$$\mathcal{J}_{12} := \int_{-\infty}^{0} \int_{W}^{0} e^{W} e^{-Y} |h(W)|^2 \, dY \, dW \qquad \text{and}$$
$$\mathcal{J}_{13} := \int_{-\infty}^{0} \int_{W}^{0} e^{W} e^{-Y} |h(Y)|^2 \, dY \, dW.$$

We first evaluate

$$\mathcal{J}_{12} = \int_{-\infty}^{0} \left(\int_{W}^{0} e^{-Y} \, dY \right) e^{W} |h(W)|^2 \, dW = \int_{-\infty}^{0} (1 - e^{W}) |h(W)|^2 \, dW.$$

Since $W \le 0$ we have $|1 - e^{W}| \le 2$, and so

$$\mathcal{J}_{12} \le 2 \int_{-\infty}^{0} |h(W)|^2 \, dW \le C \|h\|_{L^2}^2. \tag{5.43}$$

Next, we change variables in $\mathcal{J}_{13}$ to find

$$\mathcal{J}_{13} = \int_{-\infty}^{0} \left(\int_{-\infty}^{Y} e^{W} \, dW \right) e^{-Y} |h(Y)|^2 \, dY = \int_{-\infty}^{0} |h(Y)|^2 \, dY \leq \|h\|_{L^2}^2.$$

(5.44)

Combining the decomposition (5.42) and the estimates (5.43) and (5.44) gives

$$|\mathcal{J}_1| \leq C \|h\|_{L^2}^2,$$

as desired.

6 Analysis of the long wave problem

6.1 The perturbation Ansatz for the long wave problem (3.42)

Throughout this section we keep $q \in (0, c_0/\tau_2)$ fixed. We make the perturbation Ansatz

$$\psi_1 = \sigma + \eta_1 \qquad \text{and} \qquad \psi_2 = \zeta + \eta_2 \tag{6.1}$$

for the long wave problem (3.42). Here $\eta_1 \in H_{q,0}^1$, which was defined in (5.2), and $\eta_2 \in W^{1,\infty}$. We abbreviate

$$\boldsymbol{\eta} = (\eta_1, \eta_2) \in \mathcal{X} := H_{q,0}^1 \times W^{1,\infty},$$

where $\mathcal{X}$ has the norm

$$\|\boldsymbol{\eta}\|_{\mathcal{X}} := \|\eta_1\|_{H_q^1} + \|\eta_2\|_{W^{1,\infty}}.$$

The Ansatz (6.1) solves the system (3.42) if and only if η_1 and η_2 solve

$$\begin{cases} \mathcal{T}\eta_1 = \sum_{k=1}^{5} \mathcal{V}_{1k}^{\nu}(\boldsymbol{\eta}), \\ \eta_2 = \sum_{k=1}^{3} \mathcal{V}_{2k}^{\nu}(\boldsymbol{\eta}), \end{cases} \tag{6.2}$$

where $\mathcal{T}$ was defined in (3.43) and the $\mathcal{V}$-operators are given by

$$\begin{aligned}
\mathcal{V}_{11}^{\nu}(\boldsymbol{\eta}) &:= \left(\mathcal{M}^{(\nu)} - \mathcal{M}^{(0)} \right) \left[\mathcal{R}_1^{\nu}(\sigma + \eta_1)(\sigma + \eta_1) \right], \\
\mathcal{V}_{12}^{\nu}(\boldsymbol{\eta}) &:= \mathcal{M}^{(0)} \left[\left(\mathcal{R}_1^{\nu}(\sigma + \eta_1) - \mathcal{R}_1^{0}(\sigma + \eta_1) \right)(\sigma + \eta_1) \right], \\
\mathcal{V}_{13}^{\nu}(\boldsymbol{\eta}) &:= \mathcal{M}^{(0)} \left[\left(\mathcal{R}_1^{0}(\sigma + \eta_1) - \mathcal{R}_1^{0}(\sigma) - D\mathcal{R}_1^{0}(\sigma)\eta_1 \right)\sigma \right], \\
\mathcal{V}_{14}^{\nu}(\boldsymbol{\eta}) &:= \mathcal{M}^{(0)} \left[\left(\mathcal{R}_1^{0}(\sigma + \eta_1) - \mathcal{R}_1^{0}(\sigma) \right)\eta_1 \right], \\
\mathcal{V}_{15}^{\nu}(\boldsymbol{\eta}) &:= \nu^{1/2} \mathcal{M}^{(0)} \mathcal{N}^{\nu}(\sigma + \eta_1, \zeta + \eta_2)
\end{aligned} \tag{6.3}$$

and

$$
\begin{aligned}
\mathcal{V}_{21}^{v}(\boldsymbol{\eta}) &:= \mathcal{P}_{1}^{v}(\sigma + \eta_1) - \mathcal{P}_{1}^{0}(\sigma + \eta_1), \\
\mathcal{V}_{22}^{v}(\boldsymbol{\eta}) &:= \mathcal{P}_{1}^{0}(\sigma + \eta_1) - \zeta, \\
\mathcal{V}_{23}^{v}(\boldsymbol{\eta}) &:= v\mathcal{P}_{2}^{v}(\sigma + \eta_1, \zeta + \eta_2).
\end{aligned}
\tag{6.4}
$$

We recall that the symbol of $\mathcal{M}^{(v)}$ was defined in (3.40) and the symbol of $\mathcal{M}^{(0)}$ in (3.21). The operator $\mathcal{R}_{1}^{v}$ was defined in (3.35), the operator $\mathcal{N}^{v}$ in (3.38), the operator $\mathcal{P}_{1}^{v}$ in (3.33) and the operator $\mathcal{P}_{2}^{v}$ in (3.34).

Due to Proposition 5.1, the first equation in (6.2) is equivalent to

$$
\eta_1 = \mathcal{T}^{-1} \sum_{k=1}^{5} \mathcal{V}_{1k}^{v}(\boldsymbol{\eta}) =: \mathfrak{N}_{1}^{v}(\boldsymbol{\eta}).
\tag{6.5}
$$

Subsequently, η_1 and η_2 solve (6.2) if and only if

$$
\eta_2 = \mathcal{V}_{21}^{v}(\boldsymbol{\eta}) + \mathcal{V}_{22}^{v}(\mathfrak{N}_{1}^{v}(\boldsymbol{\eta})) + \mathcal{V}_{23}^{v}(\boldsymbol{\eta}) =: \mathfrak{N}_{2}^{v}(\boldsymbol{\eta}).
\tag{6.6}
$$

We have replaced η_1 with its fixed point expression (6.5) in $\mathcal{V}_{22}^{v}$ for the sake of better estimates later; see Appendix B.2.7 for a more precise discussion. Finally, set

$$
\mathfrak{N}^{v}(\boldsymbol{\eta}) := (\mathfrak{N}_{1}^{v}(\boldsymbol{\eta}), \mathfrak{N}_{2}^{v}(\boldsymbol{\eta})),
\tag{6.7}
$$

so $\mathfrak{N}^{v}$ maps $\mathcal{X}$ to $\mathcal{X}$. More precisely, this follows from the mapping estimates in Appendix B.3. We conclude that the problem (6.2) is equivalent to the fixed point problem

$$
\boldsymbol{\eta} = \mathfrak{N}^{v}(\boldsymbol{\eta}),
\tag{6.8}
$$

which we now solve.

6.2 The solution of the fixed point problem (6.8)

For $r > 0$, we define the ball

$$
\mathfrak{B}(r) := \{\boldsymbol{\eta} \in \mathcal{X} \mid \|\boldsymbol{\eta}\|_{\mathcal{X}} \leq r\}.
$$

We prove the following estimates in Appendix B; their verifications are routine, but detailed, so we do not present them here.

Proposition 6.1 *There exist $C_{\star}$, $v_{\star} > 0$ such that if $0 < v < v_{\star}$ then the following hold.*

(i) *If $\boldsymbol{\eta} \in \mathfrak{B}(C_{\star}v^{1/3})$, then $\mathfrak{N}^{v}(\boldsymbol{\eta}) \in \mathfrak{B}(C_{\star}v^{1/3})$.*

(ii) *If η, $\grave{\eta} \in \mathfrak{B}(C_\star \nu^{1/3})$, then*

$$\|\mathfrak{N}^\nu(\eta) - \mathfrak{N}^\nu(\grave{\eta})\|_{\mathcal{X}} \le \frac{1}{2}\|\eta - \grave{\eta}\|_{\mathcal{X}}.$$

Proposition 6.1 then guarantees that $\mathfrak{N}^\nu$ is a contraction on $\mathfrak{B}(C_\star \nu^{1/3})$ for each $0 < \nu < \nu_\star$, and so Banach's fixed point theorem gives the following solution to (6.2).

Proposition 6.2 *Let $C_\star$, $\nu_\star > 0$ be as in Proposition 6.1. For each $0 < \nu < \nu_\star$, there exists a unique $\eta^\nu \in \mathfrak{B}(C_\star \nu^{1/3})$ such that $\eta^\nu = \mathfrak{N}^\nu(\eta^\nu)$.*

The existence of the perturbation terms η^ν enables us to conclude our main results, which are paraphrased nontechnically in (1.5).

Theorem 6.3 *Let α, κ, τ_1, $\tau_2 > 0$, $q \in (0, c_0/\tau_2)$, and $\theta \in \mathbb{R}$. Define the leading-order profile terms*

$$\phi_A^*(X) := \left(\frac{6\sqrt{6}c_0^{9/2}}{\tau_2}\right)\left(\frac{\exp(2c_0 X/\tau_2 + \theta)}{\left[\alpha\kappa\tau_1 + 6c_0^2 \exp(2c_0 X/\tau_2 + \theta)\right]^{3/2}}\right),$$

$$\phi_P^*(X) := \left((6c_0)^{1/2}\alpha\right)\left(\frac{1}{\left[\alpha\kappa\tau_1 + 6c_0^2 \exp\left(2c_0 X/\tau_2 + \theta\right)/\tau_2)\right]^{1/2}}\right),$$

and

$$\phi_R^*(X) := (3\alpha\kappa c_0)\left(\frac{1}{\alpha\kappa\tau_1 + 6c_0^2 \exp(2c_0 X/\tau_2 + \theta)/\tau_2)}\right).$$

There exists $\epsilon_\star > 0$ such that for each $0 < \epsilon < \epsilon_\star$, there are $\phi_A^\epsilon \in H_q^1 \cap \mathcal{C}^\infty$ and ϕ_P^ϵ, $\phi_R^\epsilon \in W^{1,\infty} \cap \mathcal{C}^\infty$ with the following properties.

(i) *Let*

$$A_j(t) = \epsilon\phi_A^*(\epsilon^{2/5}(j - \epsilon^{2/5}c_0 t)) + \epsilon^{17/15}\phi_A^\epsilon(\epsilon^{2/5}(j - \epsilon^{2/5}c_0 t)),$$
$$P_j(t) = \epsilon^{1/5}\phi_P^*(\epsilon^{2/5}(j - \epsilon^{2/5}c_0 t)) + \epsilon^{1/3}\phi_P^\epsilon(\epsilon^{2/5}(j - \epsilon^{2/5}c_0 t)),$$

and

$$R_j(t) = \epsilon^{2/5}\phi_R^*(\epsilon^{2/5}(j - \epsilon^{2/5}c_0 t)) + \epsilon^{3/5}\phi_R^\epsilon(\epsilon^{2/5}(j - \epsilon^{2/5}c_0 t)).$$

Then the triple (A_j, P_j, R_j) solves (1.1).

(ii) *The remainder terms ϕ_A^ϵ, ϕ_P^ϵ, and ϕ_R^ϵ satisfy*

$$\sup_{0 < \epsilon < \epsilon_\star} \|\phi_A^\epsilon\|_{H_q^1} + \|\phi_P^\epsilon\|_{W^{1,\infty}} + \|\phi_R^\epsilon\|_{W^{1,\infty}} < \infty.$$

(iii) *The functions ϕ_P^ϵ and ϕ_R^ϵ vanish exponentially fast at $+\infty$ and are asymptotically constant at $-\infty$ in the following sense: there exist ℓ_P^ϵ, $\ell_R^\epsilon \in \mathbb{R}$ such that*

$$\sup_{0<\epsilon<\epsilon_\star} \left(|\ell_P^\epsilon| + \sup_{X\geq 0} e^{qX}|\phi_P^\epsilon(X)| + \sup_{X\leq 0} e^{-qX}|\phi_P^\epsilon(X) - \ell_P^\epsilon| \right) < \infty$$

and

$$\sup_{0<\epsilon<\epsilon_\star} \left(|\ell_R^\epsilon| + \sup_{X\geq 0} e^{qX}|\phi_R^\epsilon(X)| + \sup_{X\leq 0} e^{-qX}|\phi_R^\epsilon(X) - \ell_R^\epsilon| \right) < \infty.$$

Proof Write the solution of (6.8), which exists due to Proposition 6.2, as $\boldsymbol{\eta}^\nu = (\eta_1^\nu, \eta_2^\nu)$. Define

$$\psi_1^\nu := \sigma + \eta_1^\nu \qquad \text{and} \qquad \psi_2^\nu := \zeta + \eta_2^\nu. \tag{6.9}$$

By the discussion at the start of Sect. 6.1, the pair (ψ_1^ν, ψ_2^ν) then solves the system (3.42).

Now take

$$A_j(t) = \nu^{5/2}\psi_1^\nu(\nu(j - \nu c_0 t)) \qquad \text{and} \qquad P_j(t) = \nu^{1/2}\psi_2^\nu(\nu(j - \nu c_0 t)).$$

This is the scaled travelling wave ansatz from (3.41), and so Proposition 3.5 guarantees that A_j and P_j thus defined solve the simplified system (2.14). Let R_j be given by (2.12). Then the discussion in Sect. 2.1 shows that A_j, P_j and R_j together solve our original problem (1.1).

Next, we use the identity $\epsilon = \nu^{5/2}$ from (3.30) to reintroduce the original long wave parameter ϵ into the solutions. The expansions in (i) and the estimates in (ii) above then follow from (6.9) and the estimates in Proposition 6.2. The exact formulas for the leading order terms follow from the definitions of σ in (3.28) and ζ in (3.29). The asymptotics of part (iii) follow from the definition of ζ, the fixed point property $\eta_2^\nu = \mathfrak{N}_2^\nu(\boldsymbol{\eta}^\nu)$ and the definition of $\mathfrak{N}_2^\nu$ in (6.6), and the definition of R_j from (2.12).

Finally, to obtain the smoothness of the solutions, we first note that $\sigma, \zeta \in \mathcal{C}^\infty$. Next, crude estimates on the symbol of the Fourier multiplier $\mathcal{M}^{(\nu)}$, which we omit, allow us to invoke Lemma A.2 to conclude $\mathcal{M}^{(\nu)} \in \mathbf{B}(H_q^r, H_q^{r+1})$ for each $\nu \geq 0$ and $r \geq 0$; we make no claim about uniform estimates in ν here. This smoothing property of $\mathcal{M}^{(\nu)}$, as well as the smoothing properties of the integral operators that compose $\mathfrak{N}_2^\nu$, show that if $\boldsymbol{\eta} \in \mathcal{C}^r \times \mathcal{C}^2$, then $\mathfrak{N}^\nu(\boldsymbol{\eta}) \in \mathcal{C}^{r+1} \times \mathcal{C}^{r+1}$. By bootstrapping, we obtain $\boldsymbol{\eta}^\nu \in \mathcal{C}^\infty \times \mathcal{C}^\infty$. $\qquad \square$

In order to achieve the normalization $\|\phi_A^*\|_{L^\infty} = 1$, as discussed in Sect. 1.4, we need to use the explicit choice

$$c_0 = \left(\frac{9\alpha\kappa\tau_1\tau_2^2}{8} \right)^{1/5} =: c_*, \tag{6.10}$$

as employed in (1.5). From (6.10), we obtain

$$\|\phi_P^*\|_{L^\infty} = \left(\frac{6\alpha}{\kappa\tau_1}\right)^{1/2}\left(\frac{9\alpha\kappa\tau_1\tau_2^2}{8}\right)^{1/10} \qquad \text{and} \qquad \|\phi_R^*\|_{L^\infty} = \frac{3}{\tau_1}\left(\frac{9\alpha\kappa\tau_1\tau_2^2}{8}\right)^{1/5}.$$

(6.11)

Substituting the abbreviations (2.2) and (2.3) into the quantities in (6.10) and (6.11) then leads to the identities (1.8). This completes the rigorous derivation of the main results that we discussed more informally in Sect. 1.4.

Data availability statement The datasets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Appendix A. Fourier analysis

A.1. The Fourier transform

We use the following conventions for Fourier transforms. If $f \in L^1$, then its Fourier transform is

$$\mathfrak{F}[f](k) = \widehat{f}(k) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx,$$

and its inverse Fourier transform is

$$\mathfrak{F}^{-1}[f](x) = \check{f}(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(k)e^{ikx}\,dk.$$

A.2. Fourier multipliers on Sobolev spaces

For integers $r \geq 0$, we denote by $H^r = H^r(\mathbb{R})$ the usual Sobolev space of all r-times weakly differentiable functions whose weak derivatives are square-integrable.

Our fundamental operator on Sobolev spaces is the Fourier multiplier. The following result above is standard; see, e.g., Faver (2018, Lem. D.2.1).

Lemma A.1 *Let $\widetilde{\mathcal{M}} \colon \mathbb{R} \to \mathbb{C}$ be measurable and suppose*

$$N_{\widetilde{\mathcal{M}}}(r, s) := \sup_{k \in \mathbb{R}} \frac{|\widetilde{\mathcal{M}}(k)|}{(1 + k^2)^{(r-s)/2}} < \infty.$$

*Then the **Fourier multiplier** $\mathcal{M}$ **with symbol** $\widetilde{\mathcal{M}}$ defined by*

$$\mathcal{M}f := \mathfrak{F}^{-1}\big[\widetilde{\mathcal{M}}\widehat{f}\big], \tag{A.1}$$

i.e., by $\widehat{\mathcal{M}f}(k) = \widetilde{\mathcal{M}}(k)\widehat{f}(k)$, is a bounded operator from H^r to H^s, and

$$\|\mathcal{M}\|_{\mathbf{B}(H^r, H^s)} = N_{\widetilde{\mathcal{M}}}(r, s). \tag{A.2}$$

We also need a convenient expression for 'scaled' Fourier multipliers. If f is a function on $\mathbb{R}$ and $\nu \in \mathbb{R} \setminus \{0\}$, let $f(\nu \cdot)$ be the 'scaled' map $x \mapsto f(\nu x)$. Now let $\mathcal{M}$ be the Fourier multiplier with symbol $\widetilde{\mathcal{M}}$ and define $\widetilde{\mathcal{M}}^{(\nu)}(k) := \widetilde{\mathcal{M}}(\nu k)$. Let $\mathcal{M}^{(\nu)}$ be the Fourier multiplier with symbol $\widetilde{\mathcal{M}}^{(\nu)}$. Then standard scaling properties of the Fourier transform imply that

$$\mathcal{M}[f(\nu \cdot)](x) = (\mathcal{M}^{(\nu)}f)(\nu x). \tag{A.3}$$

A.3. Fourier multipliers on weighted Sobolev spaces

We frequently work with weighted Sobolev spaces. For $q \in \mathbb{R}$, let

$$H_q^r := \left\{ f \in H_q^r \mid e^{q|\cdot|}f \in H^r \right\}.$$

We norm this space by

$$\|f\|_{H_q^r} := \|e^{q|\cdot|}f\|_{H^r},$$

and, see Faver (2018, App. C), this norm is equivalent to

$$f \mapsto \sum_{j=0}^{r} \|e^{q|\cdot|}\partial_x^j[f]\|_{L^2}.$$

We put $L_q^2 := H_q^0$. The Cauchy–Schwarz inequality guarantees that L_q^2 embeds into L^1:

$$\|f\|_{L^1} = \|e^{-q|\cdot|}(e^{q|\cdot|}f)\|_{L^1} \le \|e^{-|q\cdot|}\|_{L^2}\|e^{q|\cdot|}f\|_{L^2} \le C_q\|f\|_{L_q^2}.$$

Finally, if $I \subseteq \mathbb{R}$ is an interval, we sometimes denote by $L_q^2(I)$ the set of all measurable functions $f: I \to \mathbb{C}$ such that

$$\int_I e^{2q|X|}|f(X)|^2\, dX < \infty.$$

Since $H_q^r \subseteq H^r$, any Fourier multiplier defined on H^r is defined on H_q^r. A variation on a result of Beale (1980, Lem. 5.1) gives sufficient conditions for a Fourier multiplier on H_q^r to map into another weighted space H_q^s.

Lemma A.2 (Beale) *Fix $q > 0$ and let*

$$\overline{\mathcal{U}}_q := \{z \in \mathbb{C} \mid |\operatorname{Im}(z)| \le q\}.$$

Let $\widetilde{\mathcal{M}}$ be analytic on $\overline{\mathcal{U}}_q$. Suppose there exist $s \ge 0$ and C, $r_0 > 0$ such that if $z \in \overline{\mathcal{U}}_q$ and $r_0 < |z|$, then

$$|\widetilde{\mathcal{M}}(z)| \le \frac{C}{|\operatorname{Re}(z)|^{(s-r)}}.$$

Then the Fourier multiplier $\mathcal{M}$ with symbol $\widetilde{\mathcal{M}}$, defined by (A.1), is a bounded operator from H_q^r to H_q^s and

$$\|\mathcal{M}\|_{\mathbf{B}(H_q^r, H_q^s)} \le \sup_{k \in \mathbb{R}} |(1 + k^2)^{(s-r)/2}\widetilde{\mathcal{M}}(k \pm iq)|.$$

Appendix B. The Proof of Proposition 6.1

Our proof depends on the following lemma, which we prove in the subsequent parts of this appendix.

Lemma B.1 *Let $\nu_{\mathcal{M}} > 0$ be as in Proposition 4.1. There exist $C_{\mathfrak{N}}$, $\rho_{\mathfrak{N}} > 0$ such that if $0 < \nu < \nu_{\mathcal{M}}$ and $\|\eta_1\|_{H_q^1}$, $\|\mathring{\eta}_1\|_{H_q^1}$, $\|\eta_2\|_{H_q^1}$, $\|\mathring{\eta}_2\|_{H_q^1} \le \rho_{\mathfrak{N}}$, then the following hold.*

(i) $\|\mathfrak{N}^\nu(\eta)\|_{\mathcal{X}} \le C_{\mathfrak{N}}(\nu^{1/3} + \|\eta\|_{\mathcal{X}}^2).$
(ii) $\|\mathfrak{N}^\nu(\eta) - \mathfrak{N}^\nu(\mathring{\eta})\|_{\mathcal{X}} \le C_{\mathfrak{N}}(\nu^{1/3} + \|\eta\|_{\mathcal{X}} + \|\mathring{\eta}\|_{\mathcal{X}})\|\eta - \mathring{\eta}\|_{\mathcal{X}}.$

Assuming this lemma to be true, we define

$$C_\star := C_{\mathfrak{N}} \quad \text{and} \quad \nu_\star := \frac{1}{2}\min\left\{\nu_{\mathcal{M}}, 1, \frac{1}{(1 + C_\star^2)^6}, \frac{1}{64 C_\star^6(1 + 2C_\star)^6}, \rho_{\mathfrak{N}}\right\}.$$

Take $0 < \nu < \nu_\star$ and $\eta, \dot{\eta} \in \mathfrak{B}(C_\star \nu)$. Then by part (i) of Lemma B.1 we have

$$\|\mathfrak{N}^\nu(\eta)\|_{\mathcal{X}} \le C_{\mathfrak{N}}(\nu^{1/3} + \|\eta\|_{\mathcal{X}}^2) \le C_\star [C_\star(1 + C_\star^2)\nu^{1/6}]\nu^{1/3} \le C_\star \nu^{1/3}.$$

This proves part (i) of Proposition 6.1. Next, part (ii) of that lemma gives

$$\|\mathfrak{N}^\nu(\eta) - \mathfrak{N}^\nu(\dot{\eta})\|_{\mathcal{X}} \le C_{\mathfrak{N}}(\nu^{1/3} + 2C_{\mathfrak{N}}\nu^{1/6})\|\eta - \dot{\eta}\|_{\mathcal{X}}$$
$$\le C_\star(1 + 2C_\star)\nu^{1/6}\|\eta - \dot{\eta}\|_{\mathcal{X}} \le \frac{1}{2}\|\eta - \dot{\eta}\|_{\mathcal{X}}.$$

This proves part (ii) of Proposition 6.1.

In the remainder of this appendix, we first give some essential auxiliary estimates in Appendix B.1. Then we prove the Lipschitz estimates of part (ii) of Proposition 6.1 in Appendix B.2. Finally, using in part these Lipschitz estimates, we prove in Appendix B.3 the mapping estimates from part (i) of the proposition.

B.1. Auxiliary estimates

Throughout this appendix we will frequently obtain estimates in terms of the L^1- or L^∞-norms of a function $f \in H_q^1$. Afterwards we can use the embedding of L_q^2 into L^1 and the corresponding inequalities

$$\|f\|_{L^1} \le C\|f\|_{L_q^2} \le C\|f\|_{H_q^1}$$

for $f \in H_q^1$, as well as the Sobolev embedding, to turn these L^1- and L^∞-estimates into H_q^1 estimates. For brevity, we will omit those details.

We will also use the operator $\mathcal{A}$ defined in (5.4); we recall

$$(\mathcal{A}f)(X) = \int_X^\infty f(W)\,dW$$

for $f \in H_q^1$, so that

$$f = -\partial_X[\mathcal{A}f] \quad \text{and} \quad \|\mathcal{A}f\|_{L^\infty} \le C\|f\|_{L^1}.$$

Last, we will use the shift operator S^ν, which satisfies

$$(S^\nu f)(X) = f(X + \nu)$$

for any function f defined on $\mathbb{R}$ and any $\nu \in \mathbb{R}$.

The operator $\mathcal{A}$ and the definitions of our fundamental operators $\mathcal{P}_1^\nu$ in (3.33) and $\mathcal{R}_1^\nu$ in (3.35) allow us to rewrite

$$\mathcal{P}_1^\nu(f)(X) = \frac{\alpha}{c_0}\mathcal{A}[\mathcal{E}(\nu^{1/2}S^\nu f)(\cdot, X)f](X)$$

and

$$\mathcal{R}_1^\nu(f)(X) = \frac{\alpha}{c_0^2 \tau_1} \int_X^\infty \mathcal{A}[\mathcal{E}(\nu^{1/2} S^\nu f)(\cdot, V) f](V)(S^\nu f)(V)\, dV.$$

Now we begin our estimates on $\mathcal{E}$, $\mathcal{P}_1^\nu$, and $\mathcal{R}_1^\nu$ in earnest.

Lemma B.2 *There exists $C > 0$ such that if $\|f\|_{L^1}$, $\|\mathring{f}\|_{L^1} \leq 1$, then the following hold.*

(i) $|\mathcal{E}(f)(V, X) - \mathcal{E}(\mathring{f})(V, X)| \leq C\|f - \mathring{f}\|_{L^1}$ *for all V, $X \in \mathbb{R}$.*
(ii) $|\mathcal{E}(f)(V, X)| \leq C$ *for all V, $X \in \mathbb{R}$.*

Proof   (i) Since

$$\left| \frac{\kappa}{c_0} \int_V^X f(U)\, dU \right| \leq \frac{\kappa}{c_0} \|f\|_{L^1}$$

for all V, $X \in \mathbb{R}$, a local Lipschitz estimate on the exponential yields $C > 0$ such that if $\|f\|_{L^1}, \|\mathring{f}\|_{L^1} \leq 1$, then

$$|\mathcal{E}(f)(V, X) - \mathcal{E}(\mathring{f})(V, X)| \leq C \left| \int_V^X f(U)\, dU - \int_V^X \mathring{f}(U)\, dU \right| \leq C\|f - \mathring{f}\|_{L^1}.$$

(ii) Since $\mathcal{E}(0) = 0$, this follows from part (i) by taking $\mathring{f} = 0$.

$\square$

The following lemma guarantees that $\mathcal{P}_1^\nu$ maps H_q^1 to $W^{1,\infty}$, among other results.

Lemma B.3 *There exists $C > 0$ such that if $0 \leq \nu < 1$ and f, $\mathring{f} \in H_q^1$ with $\|f\|_{H_q^1}$, $\|\mathring{f}\|_{H_q^1} \leq 1$, then the following hold.*

(i) $\|\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\mathring{f})\|_{L^\infty} \leq C\big(\nu^{1/2}\|f\|_{H_q^1} + 1\big)\|f - \mathring{f}\|_{H_q^1}$.
(ii) $\|\mathcal{P}_1^\nu(f)\|_{L^\infty} \leq C\|f\|_{H_q^1}$.
(iii) $\|\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f)\|_{L^\infty} \leq C\nu^{1/2}\|f\|_{H_q^1}^2$.
(iv) $\|\partial_X[\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f)]\|_{L^\infty} \leq C\nu^{1/2}\|f\|_{H_q^1}^2$.
(v) $\|\big(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f)\big) - \big(\mathcal{P}_1^\nu(\mathring{f}) - \mathcal{P}_1^0(\mathring{f})\big)\|_{L^\infty} \leq C\nu^{1/2}\big(\|f\|_{H_q^1} + \|\mathring{f}\|_{H_q^1}\big)\|f - \mathring{f}\|_{H_q^1}$.
(vi) $\|\partial_X[\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f)]\big) - \big(\partial_X[\mathcal{P}_1^\nu(\mathring{f}) - \mathcal{P}_1^0(\mathring{f})]\|_{L^\infty} \leq C\nu^{1/2}\big(\|f\|_{H_q^1} + \|\mathring{f}\|_{H_q^1}\big)\|f - \mathring{f}\|_{H_q^1}$.

Proof As we mentioned earlier, in most cases we will conclude bounds in terms of L^∞- or L^1-norms, which then immediately yield the H_q^1-bounds stated above.

(i) We have

$$\mathcal{P}_1^\nu(f)(X) - \mathcal{P}_1^\nu(\mathring{f})(X) = \mathcal{I}_1^\nu(f, \mathring{f})(X) + \mathcal{I}_2^\nu(f, \mathring{f})(X),$$

where

$$\mathcal{I}_1^\nu(f, \mathring{f})(X) := \frac{\alpha}{c_0} \int_X^\infty \left(\mathcal{E}(\nu^{1/2} S^\nu f)(V, X) - \mathcal{E}(\nu^{1/2} S^\nu \mathring{f})(V, X) \right) f(V) \, dV$$

and

$$\mathcal{I}_2^\nu(f, \mathring{f})(X) := \frac{\alpha}{c_0} \int_X^\infty \mathcal{E}(\nu^{1/2} S^\nu \mathring{f})(V, X) \left(f(V) - \mathring{f}(V) \right) dV.$$

We use part (i) of Lemma B.2 to bound

$$\left| \mathcal{E}(\nu^{1/2} S^\nu f)(V, X) - \mathcal{E}(\nu^{1/2} S^\nu \mathring{f})(V, X) \right| \leq C\nu^{1/2} \| S^\nu f - S^\nu \mathring{f} \|_{L^1}$$
$$= C\nu^{1/2} \| f - \mathring{f} \|_{L^1} \tag{B.1}$$

for all $V, X \in \mathbb{R}$. Thus

$$|\mathcal{I}_1^\nu(f, \mathring{f})(X)| \leq C\nu^{1/2} \| f - \mathring{f} \|_{L^1} \int_X^\infty |f(V)| \, dV \leq C\nu^{1/2} \| f - \mathring{f} \|_{L^1} \| f \|_{L^1}$$

for all $X \in \mathbb{R}$.
Next, we use part (ii) of Lemma B.2 to bound

$$|\mathcal{I}_2^\nu(f, \mathring{f})(X)| \leq C \int_X^\infty |f(V) - \mathring{f}(V)| \, dV \leq C \| f - \mathring{f} \|_{L^1}.$$

(ii) Since $\mathcal{P}^\nu(0) = 0$, this follows from part (i) by taking $\mathring{f} = 0$.
(iii) We have

$$\mathcal{P}_1^\nu(f)(X) - \mathcal{P}_1^0(f)(X) = \frac{\alpha}{c_0} \int_X^\infty \left(\mathcal{E}(\nu^{1/2} S^\nu f)(V, X) - 1 \right) f(V) \, dV. \tag{B.2}$$

Since

$$\mathcal{E}(\nu^{1/2} S^\nu f)(V, X) - 1 = \mathcal{E}(\nu^{1/2} S^\nu f)(V, X) - \mathcal{E}(0)(V, X),$$

we may use part (i) of Lemma B.2 to bound

$$|\mathcal{E}(\nu^{1/2} S^\nu f)(V, X) - 1| \leq C\nu^{1/2} \| S^\nu f \|_{L^1} = C\nu^{1/2} \| f \|_{L^1}. \tag{B.3}$$

Thus

$$|\mathcal{P}_1^{\nu}(f)(X) - \mathcal{P}_1^0(f)(X)| \leq C\nu^{1/2}\|f\|_{L^1}\int_X^{\infty}|f(V)|\,dV \leq C\nu^{1/2}\|f\|_{L^1}^2.$$

(iv) We first differentiate under the integral and use the condition $\mathcal{E}(g)(X, X) = 1$, apparent from the definition of $\mathcal{E}$ in (3.3) and valid for all integrable g and $X \in \mathbb{R}$, to calculate

$$\partial_X[\mathcal{P}_1^{\nu}(f)](X) = -\frac{\alpha}{c_0}f(X) + \left(\frac{\alpha}{c_0}\right)^2\nu^{1/2}$$
$$\int_X^{\infty}\mathcal{E}(\nu^{1/2}S^{\nu}f)(V, X)f(V + \nu)f(V)\,dV.$$

Then

$$\partial_X[\mathcal{P}_1^{\nu}(f) - \mathcal{P}_1^0(f)](X) = \left(\frac{\alpha}{c_0}\right)^2\nu^{1/2}$$
$$\int_X^{\infty}\mathcal{E}(\nu^{1/2}S^{\nu}f)(V, X)f(V + \nu)f(V)\,dV. \quad \text{(B.4)}$$

Part (ii) of Lemma B.2 then guarantees

$$\|\partial_X[\mathcal{P}_1^{\nu}(f) - \mathcal{P}_1^0(f)]\|_{L^{\infty}} \leq C\nu^{1/2}\|f\|_{L^{\infty}}\|f\|_{L^1}.$$

(v) We use (B.2) to write

$$\left(\mathcal{P}_1^{\nu}(f) - \mathcal{P}_1^0(f)\right) - \left(\mathcal{P}_1^{\nu}(\mathring{f}) - \mathcal{P}_1^0(\mathring{f})\right) = \mathcal{I}_3^{\nu}(f, \mathring{f}) + \mathcal{I}_4^{\nu}(f, \mathring{f}),$$

where

$$\mathcal{I}_3^{\nu}(f, \mathring{f})(X) := \frac{\alpha}{c_0}\int_X^{\infty}\left(\mathcal{E}(\nu^{1/2}S^{\nu}f)(V, X) - \mathcal{E}(\nu^{1/2}S^{\nu}\mathring{f})(V, X)\right)f(V)\,dV$$

and

$$\mathcal{I}_4^{\nu}(f, \mathring{f})(X) := \frac{\alpha}{c_0}\int_X^{\infty}\left(\mathcal{E}(\nu^{1/2}S^{\nu}\mathring{f})(V, X) - 1\right)\left(f(V) - \mathring{f}(V)\right)dV.$$

We use (B.1) to estimate

$$|\mathcal{I}_3^{\nu}(f, \mathring{f})(X)| \leq C\nu^{1/2}\|f - \mathring{f}\|_{L^1}\int_X^{\infty}|f(V)|\,dV \leq C\nu^{1/2}\|f\|_{L^1}\|f - \mathring{f}\|_{L^1}.$$

We use (B.3) to estimate

$$|\mathcal{I}_3^{\nu}(f, \mathring{f})(X)| \leq C\nu^{1/2}\|\mathring{f}\|_{L^1}\int_X^{\infty}|f(V) - \mathring{f}(V)|\,dV$$

$$\leq C \nu^{1/2} \| \grave{f} \|_{L^1} \| f - \grave{f} \|_{L^1}.$$

(vi) Using (B.4), we have

$$\partial_X [\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f)](X) - \partial_X [\mathcal{P}_1^\nu(\grave{f}) - \mathcal{P}_1^0(\grave{f})](X)$$
$$= \left(\frac{\alpha}{c_0} \right)^2 \nu^{1/2} \int_X^\infty \left[\mathcal{E}(\nu^{1/2} S^\nu f)(V, X) f(V + \nu) f(V) \right.$$
$$\left. - \mathcal{E}(\nu^{1/2} S^\nu \grave{f})(V, X) \grave{f}(V + \nu) \grave{f}(V) \right] dV.$$

The estimate follows in a manner analogous to the proof of part (v) above, so we omit the details.

$\square$

The next lemma guarantees that $\mathcal{R}_1^\nu$ maps H_q^1 to $W^{1,\infty}$.

Lemma B.4 *There exists $C > 0$ such that if $0 \leq \nu < 1$ and $\| f \|_{H_q^1}, \| \grave{f} \|_{H_q^1} \leq 1$, then the following hold.*

(i) $\| \mathcal{R}^\nu(f) \|_{L^\infty} \leq C \| f \|_{H_q^1}^2$.

(ii) $\| \partial_X [\mathcal{R}^\nu(f)] \|_{L^\infty} \leq C \| f \|_{H_q^1}^2$.

(iii) $\| \mathcal{R}^\nu(f) - \mathcal{R}^\nu(\grave{f}) \|_{L^\infty} \leq C \left(\nu^{1/2} + \| f \|_{H_q^1} + \| \grave{f} \|_{H_q^1} \right) \| f - \grave{f} \|_{H_q^1}$.

(iv) $\| \partial_X [\mathcal{R}^\nu(f) - \mathcal{R}^\nu(\grave{f})] \|_{L^\infty} \leq C \left(\nu^{1/2} + \| f \|_{H_q^1} + \| \grave{f} \|_{H_q^1} \right) \| f - \grave{f} \|_{H_q^1}$

(v) $\| \mathcal{R}^\nu(f) - \mathcal{R}_1^0(f) \|_{L^\infty} \leq C \nu^{1/2} \| f \|_{H_q^1}^2$.

(vi) $\| \left(\mathcal{R}^\nu(f) - \mathcal{R}_1^0(f) \right) - \left(\mathcal{R}^\nu(\grave{f}) - \mathcal{R}_1^0(\grave{f}) \right) \|_{L^\infty} \leq C \left(\nu^{1/2} + \| f \|_{H_q^1} + \| \grave{f} \|_{H_q^1} \right) \| f - \grave{f} \|_{H_q^1}$.

(vii) $\| \mathcal{R}^\nu(f) - \mathcal{R}_1^0(f) \|_{L^\infty} \leq C \nu^{1/2} \| f \|_{H_q^1}^2$.

Proof Throughout we will use the inequality

$$\| \mathcal{R}^\nu(f) \|_{L^\infty} \leq \| \mathcal{P}_1^\nu(f)(S^\nu f) \|_{L^1}.$$

As before, we stop when we have bounds in terms of L^1- or L^∞-norms.

(i) We use part (ii) of Lemma B.3 to bound

$$\| \mathcal{R}^\nu(f) \|_{L^\infty} = C \| \mathcal{P}_1^\nu(f)(S^\nu f) \|_{L^1} \leq C \| \mathcal{P}_1^\nu(f) \|_{L^\infty} \| S^\nu f \|_{L^1} \leq C \| f \|_{L^1}^2.$$

(ii) We have

$$\partial_X [\mathcal{R}^\nu(f)] = -\frac{\alpha}{c_0^2 \tau_1} \mathcal{P}_1^\nu(f)(S^\nu f),$$

thus

$$\|\partial_X[\mathcal{R}^\nu(f)]\|_{L^\infty} \le C\|\mathcal{P}_1^\nu(f)(S^\nu f)\|_{L^\infty} \le C\|\mathcal{P}_1^\nu(f)\|_{L^\infty}\|f\|_{H_q^1} \le C\|f\|_{H_q^1}^2$$

by the Sobolev embedding and part (ii) of Lemma B.3.

(iii) We use parts (i) and (ii) of Lemma B.3 to bound

$$\|\mathcal{R}^\nu(f) - \mathcal{R}^\nu(\grave{f})\|_{L^\infty} \le C\|(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(f))f\|_{L^1} + C\|\mathcal{P}_1^\nu(\grave{f})(S^\nu(f - \grave{f}))\|_{L^1}$$
$$\le C\|\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\grave{f})\|_{L^\infty}\|f\|_{L^1} + C\|\mathcal{P}_1^\nu(\grave{f})\|_{L^\infty}\|S^\nu(f - \grave{f})\|_{L^1}$$
$$\le C(\nu^{1/2}\|f\|_{L^1} + 1)\|f\|_{L^1}\|f - \grave{f}\|_{L^1} + C\|\grave{f}\|_{L^1}\|f - \grave{f}\|_{L^1}.$$

(iv) We have

$$\partial_X[\mathcal{R}^\nu(f) - \mathcal{R}^\nu(\grave{f})] = \frac{\alpha}{c_0^2\tau_1}\mathcal{P}_1^\nu(\grave{f})(S^\nu\grave{f}) - \frac{\alpha}{c_0^2\tau_1}\mathcal{P}_1^\nu(f)(S^\nu f),$$

thus

$$\|\partial_X[\mathcal{R}^\nu(f) - \mathcal{R}^\nu(\grave{f})]\|_{L^\infty} \le C\|(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\grave{f}))\grave{f}\|_{L^\infty}$$
$$+ C\|\mathcal{P}_1^\nu(\grave{f})(S^\nu(f - \grave{f}))\|_{L^\infty}.$$

We use part (i) of Lemma B.3 and the Sobolev embedding to estimate

$$\|(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\grave{f}))\grave{f}\|_{L^\infty} \le \|\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\grave{f})\|_{L^\infty}\|f\|_{H_q^1}$$
$$\le C(\nu^{1/2}\|f\|_{L^1} + 1)\|f\|_{H_q^1}\|f - \grave{f}\|_{L^1}$$

and part (ii) of Lemma B.3 and the Sobolev embedding to estimate

$$\|\mathcal{P}_1^\nu(\grave{f})(S^\nu(f - \grave{f}))\|_{L^\infty} \le C\|\mathcal{P}_1^\nu(\grave{f})\|_{L^\infty}\|f - \grave{f}\|_{L^\infty} \le C\|f\|_{L^1}\|f - \grave{f}\|_{H_q^1}.$$

(v) We first estimate

$$\|\mathcal{R}^\nu(f) - \mathcal{R}_1^0(f)\|_{L^\infty} \le C\|\mathcal{P}_1^\nu(f)(S^\nu f) - \mathcal{P}_1^0(f)f\|_{L^1}$$
$$\le C\|(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f))(S^\nu f)\|_{L^1}$$
$$+ C\|\mathcal{P}_1^0(f)(S^\nu f - f)\|_{L^1}.$$

Then part (iii) of Lemma B.3 gives

$$\|(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f))(S^\nu f)\|_{L^1} \le \|\mathcal{P}_1^\nu(f) - \mathcal{P}_1^0(f)\|_{L^\infty}\|S^\nu f\|_{L^1} \le C\nu^{1/2}\|f\|_{L^1}^3.$$

Next, part (ii) of Lemma B.3 implies

$$\|\mathcal{P}_1^0(f)(S^\nu f - f)\|_{L^1} \le \|\mathcal{P}_1^0(f)\|_{L^\infty}\|S^\nu f - f\|_{L^1} \le C\|f\|_{L^1}\|(S^\nu - 1)f\|_{L^1}.$$

Since $f \in H_q^1$, we have

$$\|(S^\nu - 1)f\|_{L^1} \le C_q \|(S^\nu - 1)f\|_{L_q^2}.$$

It follows from Faver and Wright (2018, Lem. A.11) that

$$\|(S^\nu - 1)f\|_{L_q^2} \le C\nu \|f\|_{H_q^1}.$$

(vi) We estimate

$$
\begin{aligned}
\|\big(\mathcal{R}^\nu(f) - \mathcal{R}_1^0(f)\big) &- \big(\mathcal{R}^\nu(\mathring{f}) - \mathcal{R}_1^0(\mathring{f})\big)\|_{L^\infty} \le C\|\big(\mathcal{P}_1^\nu(f)(S^\nu f) \\
&- \mathcal{P}_1^0(f)f\big) - \big(\mathcal{P}^\nu(\mathring{f})(S^\nu \mathring{f}) - \mathcal{P}_1^0(\mathring{f})\mathring{f}\big)\|_{L^1} \\
&\le C\|\big(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\mathring{f})\big)(S^\nu f)\|_{L^1} + C\|\mathcal{P}_1^\nu(\mathring{f})\big(S^\nu(f - \mathring{f})\big)\|_{L^1} \\
&+ C\|\big(P_1^0(f) - \mathcal{P}_1^0(\mathring{f})\big)\mathring{f}\|_{L^1} + C\|\mathcal{P}_1^0(f)(f - \mathring{f})\|_{L^1}.
\end{aligned}
$$

We use part (ii) of Lemma B.3 to bound

$$
\begin{aligned}
\|\mathcal{P}_1^\nu(\mathring{f})\big(S^\nu(f - \mathring{f})\big)\|_{L^1} + \|\mathcal{P}_1^0(f)(f - \mathring{f})\|_{L^1} &\le \|\mathcal{P}_1^\nu(\mathring{f}\|_{L^\infty}\|S^\nu(f - \mathring{f})\|_{L^1} \\
&+ \|\mathcal{P}_1^0(f)\|_{L^\infty}\|f - \mathring{f}\|_{L^1} \\
&\le C\|f\|_{L^1}\|f - \mathring{f}\|_{L^1}.
\end{aligned}
$$

We use part (i) of Lemma B.3 to bound

$$
\begin{aligned}
\|\big(\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\mathring{f})\big)(S^\nu f)\|_{L^1} &+ \|\big(P_1^0(f) - \mathcal{P}_1^0(\mathring{f})\big)\mathring{f}\|_{L^1} \\
&\le \|\mathcal{P}_1^\nu(f) - \mathcal{P}_1^\nu(\mathring{f})\|_{L^\infty}\|S^\nu f\|_{L^1} + \|\mathcal{P}_1^0(f) - \mathcal{P}_1^0(\mathring{f})\|_{L^\infty}\|\mathring{f}\|_{L^1} \\
&\le C\big(\nu^{1/2}\|f\|_{L^1} + 1\big)\|f\|_{L^1}\|f - \mathring{f}\|_{L^1} + C\|\mathring{f}\|_{L^1}\|f - \mathring{f}\|_{L^1}.
\end{aligned}
$$

(vii) We use part (vi) with $\mathring{f} = 0$.

$\square$

Finally, we present estimates on the operators $\mathcal{N}^\nu$ defined in (3.38) and $\mathcal{P}_2^\nu$ from (3.34).

Lemma B.5 *There exist $C, \rho_0 > 0$ such that if $0 \le \nu < 1$, then the following hold.*

(i) *If $f, \mathring{f} \in H_q^1$ and $g, \mathring{g} \in W^{1,\infty}$ with $\|f\|_{H_q^1} + \|g\|_{W^{1,\infty}} \le \rho_0$ and $\|\mathring{f}\|_{H_q^1} + \|\mathring{g}\|_{W^{1,\infty}} \le \rho_0$, then*

$$
\begin{aligned}
\|\mathcal{N}^\nu(f, g) - \mathcal{N}^\nu(\mathring{f}, \mathring{g})\|_{L_q^2} &+ \|\mathcal{P}_2^\nu(f, g) - \mathcal{P}_2^\nu(\mathring{f}, \mathring{g})\|_{H_q^1} \\
&\le C\big(\nu^{1/2} + \|f\|_{H_q^1} + \|\mathring{f}\|_{H_q^1} + \|g\|_{W^{1,\infty}} + \|\mathring{g}\|_{W^{1,\infty}}\big) \\
&\quad \big(\|f - \mathring{f}\|_{H_q^1} + \|g - \mathring{g}\|_{W^{1,\infty}}\big).
\end{aligned}
$$

(ii) *If $f \in H_q^1$ and $g \in W^{1,\infty}$ with $\|f\|_{H_q^1} + \|g\|_{W^{1,\infty}} \leq \rho_0$, then*

$$\|\mathcal{N}^\nu(f, g)\|_{L_q^2} + \|\mathcal{P}_2^\nu(f, g)\|_{W^{1,\infty}} \leq C.$$

Proof Part (ii) follows from part (i) since $\mathcal{N}^\nu(0, 0) = \mathcal{P}_2^\nu(0, 0) = 0$. The proof of the Lipschitz estimates in part (i) follows exactly the strategies deployed above, and we would learn almost nothing new from seeing its argument, so we omit that. The one difference here is that $\mathcal{N}^\nu$ and $\mathcal{P}_2^\nu$ incorporate the maps $\mathcal{Q}_1^\nu$ and $\mathcal{Q}_2^\nu$, which were defined in (3.32) and which are really rational functions from $\mathbb{R}^2$ to $\mathbb{R}$. A glance at the formulas for $\mathcal{Q}_1^\nu$ and $\mathcal{Q}_2^\nu$ provides $\rho_\mathcal{Q} > 0$ such that if $0 < \nu < 1$, then $\mathcal{Q}_1^\nu$ and $\mathcal{Q}_2^\nu$ are defined and smooth on the ball $\{(X, Y) \in \mathbb{R}^2 \mid |X| + |Y| \leq \rho_\mathcal{Q}\}$. By taking $\|f\|_{H_q^1} + \|g\|_{W^{1,\infty}} \leq \rho_0$ for some small $\rho_0 > 0$, we can guarantee that the compositions involving $\mathcal{Q}_1^\nu$ and $\mathcal{Q}_2^\nu$ with f, g, and other operators acting on f and g are all defined and satisfy tame Lipschitz estimates. $\square$

B.2. Lipschitz estimates

We first prove the Lipschitz estimates undergirding part (ii) of Lemma B.1, which we then use to prove the mapping estimates in part (i). From (6.7), we have $\mathfrak{N}^\nu = (\mathfrak{N}_1^\nu, \mathfrak{N}_2^\nu)$, where $\mathfrak{N}_1^\nu$ was defined in (6.5) and $\mathfrak{N}_2^\nu$ in (6.6). Using these definitions and the boundedness of the operator $\mathcal{T}^{-1}$ from Proposition 5.1, we can prove part (ii) of Lemma B.1 if we show

$$\sum_{k=1}^{5} \left(\|\mathcal{V}_{1k}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{1k}^\nu(\check{\boldsymbol{\eta}})\|_{H_q^1} \right) + \left(\|\mathcal{V}_{21}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{21}^\nu(\check{\boldsymbol{\eta}})\|_{W^{1,\infty}} \right)$$
$$+ \|\mathcal{V}_{23}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{23}^\nu(\check{\boldsymbol{\eta}})\|_{W^{1,\infty}} \leq C\mathfrak{R}_\star^\nu(\boldsymbol{\eta}, \check{\boldsymbol{\eta}}),$$

where

$$\mathfrak{R}_\star^\nu(\boldsymbol{\eta}, \check{\boldsymbol{\eta}}) := \left(\nu^{1/3} + \|\eta_1\|_{H_q^1} + \|\check{\eta}_1\|_{H_q^1} + \|\eta_2\|_{W^{1,\infty}} + \|\check{\eta}_2\|_{W^{1,\infty}} \right)$$
$$\left(\|\eta_1 - \check{\eta}_1\|_{H_q^1} + \|\eta_2 - \check{\eta}_2\|_{W^{1,\infty}} \right).$$

The terms $\mathcal{V}_{1k}^\nu$ were defined in (6.3) and $\mathcal{V}_{2k}^\nu$ in (6.4).

B.2.1. Lipschitz estimates on $\mathcal{V}_{11}^\nu$

We use the estimate on $\mathcal{M}^{(\nu)} - \mathcal{M}^{(0)}$ from Proposition 4.1 to obtain

$$\|\mathcal{V}_{11}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{11}^\nu(\check{\boldsymbol{\eta}})\|_{H_q^1} \leq C\nu^{1/3} \|\left(\mathcal{R}^\nu(\sigma + \eta_1) - \mathcal{R}^\nu(\sigma + \check{\eta}_1)\right)(\sigma + \check{\eta}_1)\|_{H_q^1}$$
$$+ C\nu^{1/3} \|\mathcal{R}^\nu(\sigma + \eta_1)(\eta_1 - \check{\eta}_1)\|_{H_q^1}.$$

We first estimate

$$\|\big(\mathcal{R}^\nu(\sigma + \eta_1) - \mathcal{R}^\nu(\sigma + \mathring{\eta}_1)\big)(\sigma + \mathring{\eta}_1)\|_{H_q^1} \le \|\partial_X\big[\mathcal{R}^\nu(\sigma + \eta_1) - \mathcal{R}^\nu(\sigma + \mathring{\eta}_1)\big]$$
$$(\sigma + \mathring{\eta}_1)\|_{L_q^2} + \|\big(\mathcal{R}^\nu(\sigma + \eta_1) - \mathcal{R}^\nu(\sigma + \mathring{\eta}_1)\big)\partial_X[\sigma + \mathring{\eta}_1]\|_{L_q^2},$$

where

$$\|\partial_X\big[\mathcal{R}^\nu(\sigma + \eta_1) - \mathcal{R}^\nu(\sigma + \mathring{\eta}_1)\big](\sigma + \mathring{\eta}_1)\|_{L_q^2}$$
$$\le \|\partial_X\big[\mathcal{R}^\nu(\sigma + \eta_1) - \mathcal{R}_1^\nu(\sigma + \mathring{\eta}_1)\big]\|_{L^\infty}\|\sigma + \mathring{\eta}_1\|_{L_q^2}$$
$$\le C\big(\nu^{1/2} + \|\eta_1\|_{H_q^1} + \|\mathring{\eta}_1\|_{H_q^1}\big)\|\eta_1 - \mathring{\eta}_1\|_{H_q^1}$$

by part (iv) of Lemma B.4 and

$$\|\big(\mathcal{R}_1^\nu(\sigma + \eta_1) - \mathcal{R}_1^\nu(\sigma + \mathring{\eta}_1)\big)\partial_X[\sigma + \mathring{\eta}_1]\|_{L_q^2},$$
$$\le \|\mathcal{R}_1^\nu(\sigma + \eta_1) - \mathcal{R}_1^\nu(\sigma + \mathring{\eta}_1)\|_{L^\infty}\|\partial_X[\sigma + \mathring{\eta}_1]\|_{L_q^2}$$
$$\le C\big(\nu^{1/2} + \|\eta_1\|_{H_q^1} + \|\mathring{\eta}_1\|_{H_q^1}\big)\|\eta_1 - \mathring{\eta}_1\|_{H_q^1}$$

by part (iii) of Lemma B.4.

Next we estimate

$$\|\mathcal{R}_1^\nu(\sigma + \eta_1)(\eta_1 - \mathring{\eta}_1)\|_{H_q^1} \le \|\partial_X\big[\mathcal{R}_1^\nu(\sigma + \eta_1)\big](\eta_1 - \mathring{\eta}_1)\|_{L_q^2}$$
$$+ \|\mathcal{R}_1^\nu(\sigma + \eta_1)\partial_X[\eta_1 - \mathring{\eta}_1]\|_{L_q^2},$$

where

$$\|\partial_X\big[\mathcal{R}_1^\nu(\sigma + \eta_1)\big](\eta_1 - \mathring{\eta}_1)\|_{L_q^2} \le \|\partial_X\big[\mathcal{R}_1^\nu(\sigma + \eta_1)\big]\|_{L^\infty}\|\eta_1 - \mathring{\eta}_1\|_{L_q^2}$$
$$\le C\|\sigma + \eta_1\|_{H_q^1}^2\|\eta_1 - \mathring{\eta}_1\|_{H_q^1} \le C\|\eta_1 - \mathring{\eta}_1\|_{H_q^1}$$

by part (ii) of Lemma B.4 and

$$\|\mathcal{R}_1^\nu(\sigma + \eta_1)\partial_X[\eta_1 - \mathring{\eta}_1]\|_{L_q^2} \le \|\mathcal{R}_1^\nu(\sigma + \eta_1)\|_{L^\infty}\|\partial_X[\eta_1 - \mathring{\eta}_1]\|_{L_q^2}$$
$$\le C\|\sigma + \eta_1\|_{H_q^1}^2\|\eta_1 - \mathring{\eta}_1\|_{H_q^1} \le C\|\eta_1 - \mathring{\eta}_1\|_{H_q^1}^2.$$

by part (i) of Lemma B.4.

B.2.2. Lipschitz estimates on $\mathcal{V}_{12}^\nu$

We use the smoothing property of $\mathcal{M}^{(0)}$ from Lemma 3.1 to bound

$$\|\mathcal{V}_{12}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{12}^\nu(\mathring{\boldsymbol{\eta}})\|_{H_q^1} \le C\|\big[\big(\mathcal{R}_1^\nu(\eta_1)$$

$$-\mathcal{R}_1^0(\eta_1)) - \left(\mathcal{R}_1^v(\grave{\eta}_1) - \mathcal{R}_1^0(\grave{\eta}_1)\right)](\sigma + \eta_1)\|_{L_q^2}$$
$$+C\|\left(\mathcal{R}_1^v(\grave{\eta}_1) - \mathcal{R}_1^0(\grave{\eta}_1)\right)(\eta_1 - \grave{\eta}_1)\|_{L_q^2}.$$

Call the two L_q^2-norm terms above I and II. We estimate

$$I \leq \|\left(\mathcal{R}_1^v(\eta_1) - \mathcal{R}_1^0(\eta_1)\right) - \left(\mathcal{R}_1^v(\grave{\eta}_1) - \mathcal{R}_1^0(\grave{\eta}_1)\right)\|_{L^\infty}\|\sigma + \eta_1\|_{L_q^2}$$
$$\leq C\left(v^{1/2} + \|\eta_1\|_{H_q^1} + \|\grave{\eta}_1\|_{H_q^1}\right)\|\eta_1 - \grave{\eta}_1\|_{H_q^1}.$$

by part (vi) of Lemma B.4 and

$$II \leq \|\mathcal{R}_1^v(\grave{\eta}_1) - \mathcal{R}_1^0(\grave{\eta}_1)\|_{L^\infty}\|\eta_1 - \grave{\eta}_1\|_{L_q^2} \leq Cv^{1/2}\|\grave{\eta}_1\|_{H_q^1}\|\eta_1 - \grave{\eta}_1\|_{H_q^1}$$

by part (vii) of Lemma B.4.

B.2.3. Lipschitz estimates on $\mathcal{V}_{13}^v$

We use again the smoothing property of $\mathcal{M}^{(0)}$ to bound

$$\|\mathcal{V}_{13}^v(\boldsymbol{\eta}) - \mathcal{V}_{13}^v(\grave{\boldsymbol{\eta}})\|_{H_q^1} \leq C\|\left(\mathcal{R}_1^0(\sigma + \eta_1) - \mathcal{R}_1^0(\sigma) - D\mathcal{R}_1^0(\sigma)\eta_1\right)\sigma\|_{L_q^2}$$
$$\leq C\|\mathcal{R}_1^0(\sigma + \eta_1) - \mathcal{R}_1^0(\sigma) - D\mathcal{R}_1^0(\sigma)\eta_1\|_{L^\infty}\|\sigma\|_{L_q^2}.$$

Next we will use the following 'difference of squares' estimate, which is proved using the fundamental theorem of calculus. We thank J. Douglas Wright for pointing out this lemma to us.

Lemma B.6 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces with $\mathcal{Z} \subseteq \mathcal{X}$ open and convex and with $0 \in \mathcal{Z}$. Let $f \in C^1(\mathcal{Z}, \mathcal{Y})$ with $Df(0) = 0$, and suppose*

$$\mathrm{Lip}_{\mathcal{Z}}(Df) := \sup_{\substack{x,\grave{x}\in\mathcal{Z} \\ x\neq\grave{x}}} \frac{\|Df(x) - Df(\grave{x})\|_{\mathbf{B}(\mathcal{X},\mathcal{Y})}}{\|x - \grave{x}\|_{\mathcal{X}}} < \infty.$$

Then

$$\|f(x) - f(\grave{x})\|_{\mathcal{Y}} \leq \frac{1}{2}\,\mathrm{Lip}_{\mathcal{Z}}(Df)\left(\|x\|_{\mathcal{X}} + \|\grave{x}\|_{\mathcal{X}}\right)\|x - \grave{x}\|_{\mathcal{X}}.$$

We apply this lemma to $f(\eta_1) := \mathcal{R}_1^0(\sigma)\eta_1 - \mathcal{R}_1^0(\sigma) - D\mathcal{R}_1^0(\sigma)\eta_1$, which is infinitely differentiable as a map from H_q^1 to $W^{1,\infty}$ by Remark 3.4, to conclude

$$\|\mathcal{R}_1^0(\sigma + \eta_1) - \mathcal{R}_1^0(\sigma) - D\mathcal{R}_1^0(\sigma)\eta_1\|_{L^\infty} \leq C\left(\|\eta_1\|_{H_q^1} + \|\grave{\eta}_1\|_{H_q^1}\right)\|\eta_1 - \grave{\eta}_1\|_{H_q^1}.$$

B.2.4. Lipschitz estimates on $\mathcal{V}_{14}^\nu$

We smooth with $\mathcal{M}^{(0)}$ once more, and then we use the fundamental theorem of calculus and the smoothness of $\mathcal{R}_1^0$ to rewrite

$$\|\mathcal{V}_{14}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{14}^\nu(\grave{\boldsymbol{\eta}})\|_{H_q^1} \le C\|I\|_{L_q^2} + C\|II\|_{L_q^2},$$

where

$$I := \left(\int_0^1 \left(D\mathcal{R}_1^0(\sigma + s\eta_1) - D\mathcal{R}_1^0(\sigma + s\grave{\eta}_1) \right) ds \right) \eta_1^2$$

and

$$II := \left(\int_0^1 D\mathcal{R}_1^0(\sigma + s\grave{\eta}_1) \, ds \right) (\eta_1 + \grave{\eta}_1)(\eta_1 - \grave{\eta}_1).$$

Then

$$\|I\|_{L_q^2} \le \left\| \int_0^1 \left(D\mathcal{R}_1^0(\sigma + s\eta_1) - D\mathcal{R}_1^0(\sigma + s\grave{\eta}_1) \right) ds \right\|_{L^\infty} \|\eta_1\|_{L_q^2}$$

and

$$\|II\|_{L_q^2} \le \left\| \int_0^1 D\mathcal{R}_1^0(\sigma + s\grave{\eta}_1) \, ds \right\|_{L^\infty} \|\eta_1 + \grave{\eta}_1\|_{L^\infty} \|\eta_1 - \grave{\eta}_1\|_{L_q^2}.$$

We conclude

$$\|I\|_{L_q^2} \le C\|\eta_1\|_{H_q^1} \|\eta_1 - \grave{\eta}_1\|_{H_q^1}$$

via a Lipschitz estimate on $D\mathcal{R}_1^0$ and

$$\|II\|_{L_q^2} \le C\big(\|\eta_1\|_{H_q^1} + \|\grave{\eta}_1\|_{H_q^1}\big)\|\eta_1 - \grave{\eta}_1\|_{H_q^1}$$

via the boundedness of $D\mathcal{R}_1^0$.

B.2.5. Lipschitz estimates on $\mathcal{V}_{15}^\nu$

We smooth with $\mathcal{M}^{(0)}$ to estimate

$$\|\mathcal{V}_{15}^\nu(\boldsymbol{\eta}) - \mathcal{V}_{15}^\nu(\grave{\boldsymbol{\eta}})\|_{H_q^1} \le C\nu^{1/2}\|\mathcal{N}^\nu(\sigma + \eta_1, \zeta + \eta_2) - \mathcal{N}^\nu(\sigma + \grave{\eta}_1, \zeta + \grave{\eta}_2)\|_{L_q^2}.$$

The desired estimate then follows from part (i) of Lemma B.5.

B.2.6. Lipschitz estimates on $\mathcal{V}_{21}^{\nu}$

This is a direct application of parts (v) and (vi). of Lemma B.3.

B.2.7. Lipschitz estimates on $\mathcal{V}_{22}^{\nu} \circ \mathfrak{N}_1^{\nu}$

We have

$$
\mathcal{V}_{22}^{\nu}(\mathfrak{N}_1^{\nu}(\boldsymbol{\eta}))(X) = \mathcal{P}_1^0(\sigma + \mathfrak{N}_1^{\nu}(\boldsymbol{\eta}))(X) - \mathcal{P}_1^0(\sigma)(X) = \frac{\alpha}{c_0} \int_X^{\infty} \mathfrak{N}_1^{\nu}(\boldsymbol{\eta})(V) \, dV.
$$
(B.5)

The desired Lipschitz estimate on $\mathcal{V}_{22}^{\nu}$ then follows at once from the Lipschitz estimate

$$
\|\mathfrak{N}_1^{\nu}(\boldsymbol{\eta}) - \mathfrak{N}_1^{\nu}(\grave{\boldsymbol{\eta}})\|_{H_q^1} \le C \mathfrak{R}_{\star}^{\nu}(\boldsymbol{\eta}, \grave{\boldsymbol{\eta}}),
$$

which we proved in Appendix B.2.1 through B.2.5. Without having substituted $\mathfrak{N}_1^{\nu}(\boldsymbol{\eta})$ for η_1 in the process of defining $\mathfrak{N}_2^{\nu}$ in (6.6), we would have only a useless $\mathcal{O}(1)$ estimate here.

B.2.8. Lipschitz estimates on $\mathcal{V}_{23}^{\nu}$

This is a direct application of part (i) of Lemma B.5.

B.3. Mapping estimates

We prove the mapping estimates that deliver part (i) of Lemma B.1 and rely mostly on the preceding Lipschitz estimates. Due to the boundedness of $\mathcal{T}^{-1}$, it suffices to show

$$
\sum_{k=1}^{5} \|\mathcal{V}_{1k}^{\nu}(\boldsymbol{\eta})\|_{H_q^1} + \|\mathcal{V}_{21}^{\nu}(\boldsymbol{\eta})\|_{W^{1,\infty}} + \|\mathcal{V}_{23}^{\nu}(\boldsymbol{\eta})\|_{W^{1,\infty}}
$$
$$
\le C\big(\nu^{1/3} + \|\eta_1\|_{H_q^1}^2 + \|\eta_2\|_{W^{1,\infty}}^2\big).
$$

B.3.1. Mapping estimates on $\mathcal{V}_{11}^{\nu}$

We estimate

$$
\|\mathcal{V}_{11}^{\nu}(\boldsymbol{\eta})\|_{H_q^1} \le \|\mathcal{V}_{11}^{\nu}(\boldsymbol{\eta}) - \mathcal{V}_{11}^{\nu}(0)\|_{H_q^1} + \|\mathcal{V}_{11}^{\nu}(0)\|_{H_q^1},
$$

where

$$
\|\mathcal{V}_{11}^{\nu}(\boldsymbol{\eta}) - \mathcal{V}_{11}^{\nu}(0)\|_{H_q^1} \le C\nu^{1/3} \|\eta_1\|_{H_q^1}^2
$$

by the Lipschitz estimates in Appendix B.2.1 and

$$\|\mathcal{V}_{11}^{\nu}(0)\|_{H_q^1} = \|\big(\mathcal{M}^{(\nu)} - \mathcal{M}^{(0)}\big)\big[\mathcal{R}_1^{\nu}(\sigma)\sigma\big]\|_{H_q^1} \leq C\nu^{1/3}$$

by Proposition 4.1.

B.3.2. Mapping estimates on $\mathcal{V}_{12}^{\nu}$

We estimate

$$\|\mathcal{V}_{12}^{\nu}(\boldsymbol{\eta})\|_{L_q^2} \leq \|\mathcal{V}_{12}^{\nu}(\boldsymbol{\eta}) - \mathcal{V}_{12}^{\nu}(0)\|_{L_q^2} + \|\mathcal{V}_{12}^{\nu}(0)\|_{L_q^2},$$

where

$$\|\mathcal{V}_{12}^{\nu}(\boldsymbol{\eta}) - \mathcal{V}_{12}^{\nu}(0)\|_{L_q^2} \leq C$$

by the Lipschitz estimates in Appendix B.2.1 and

$$\|\mathcal{V}_{12}^{\nu}(0)\|_{L_q^2} = \|\big(\mathcal{R}_1^{\nu}(\sigma) - \mathcal{R}_1^0(\sigma)\big)\sigma\|_{L_q^2} \leq \|\mathcal{R}_1^{\nu}(\sigma) - \mathcal{R}_1^0(\sigma)\|_{L^{\infty}}\|\sigma\|_{L_q^2} \leq C\nu^{1/2}$$

by part (vii) of Lemma B.4.

B.3.3. Mapping estimates on $\mathcal{V}_{13}^{\nu}$

Because $\mathcal{V}_{13}^{\nu}(0) = 0$, these follow from the Lipschitz estimates for $\mathcal{V}_{13}^{\nu}$ that we developed above in Appendix B.2.3.

B.3.4. Mapping estimates on $\mathcal{V}_{14}^{\nu}$

Because $\mathcal{V}_{14}^{\nu}(0) = 0$, these follow from the Lipschitz estimates for $\mathcal{V}_{14}^{\nu}$ that we developed above in Appendix B.2.3.

B.3.5. Mapping estimates on $\mathcal{V}_{15}^{\nu}$

The estimates are analogous to those in Appendix B.3.1, except now we use Lemma B.5 instead of the Lipschitz estimates in Appendix B.2.1.

B.3.6. Mapping estimates on $\mathcal{V}_{21}^{\nu}$

These estimates follow directly from parts (iii) and (iv) of Lemma B.3.

B.3.7. Mapping estimates on $\mathcal{V}_{22}^{\nu} \circ \mathfrak{N}_1^{\nu}$

We obtain these estimates by first rewriting $\mathcal{V}_{22}^{\nu} \circ \mathfrak{N}_1^{\nu}$ via the identity (B.5) and then using the mapping estimates on $\mathfrak{N}_1^{\nu}$ developed in Appendices B.3.1 through B.3.5.

B.3.8. Mapping estimates on $\mathcal{V}_{23}^{\nu}$

This estimate follows from part (ii) of Lemma B.5.

References

Adamowski M, Friml J (2015) PIN-dependent auxin transport: action, regulation, and evolution. Plant Cell 27:20–32

Allen HR, Ptashnyk M (2020) Mathematical modelling of auxin transport in plant tissues: flux meets signalling and growth. Bull Math Biol 82:1–35

Althuis R (2021) Auxin waves in a two-dimensional grid. BSc thesis, Leiden University. https://pub.math. leidenuniv.nl/hupkeshj/scriptie_rosalie.pdf

Aronson DG, Weinberger HF (1975) Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation, in Partial differential equations and related topics (Program, Tulane Univ., New Orleans, La., 1974). Lecture notes in mathematics, vol 446. Springer, Berlin, pp 5–49

Aronson DG, Weinberger HF (1978) Multidimensional nonlinear diffusion arising in population genetics. Adv Math 30:33–76

Autran D, Bassel GW, Chae E, Ezer D, Ferjani A, Fleck C, Hamant O, Hartmann FP, Jiao Y, Johnston IG, Kwiatkowska D, Lim BL, Mähönen AP, Morris RJ, Mulder BM, Nakayama N, Sozzani R, Strader LC, Tusscher Kt, Ueda M, Wolf S (2021) What is quantitative plant biology? Quant Plant Biol 2

Bayer EM, Smith RS, Mandel T, Nakayama N, Sauer M, Prusinkiewicz P, Kuhlemeier C (2009) Integration of transport-based models for phyllotaxis and midvein formation. Genes Dev 23:373–384

Beale JT (1980) Water waves generated by a pressure disturbance on a steady stream. Duke Math J 47:297–323

Benítez M, Hernández-Hernández V, Newman SA, Niklas KJ (2018) Dynamical patterning modules, biogeneric materials, and the evolution of multicellular plants. Front Plant Sci 9:871

Brillouin L (1953) Wave propagation in periodic structures. Dover Phoenix Editions, New York

Chen X, Guo J-S, Wu C-C (2008) Traveling waves in discrete periodic media for bistable dynamics. Arch Ration Mech Anal 189:189–236

Cieslak M, Owens A, Prusinkiewicz P (2021) Computational models of auxin-driven patterning in shoots. Cold Spring Harb Perspect Biol 14:a040097

Dauxois T (2008) Fermi, Pasta, Ulam, and a mysterious lady. Phys Today 61:55–57

Draelants D, Avitabile D, Vanroose W (2015) Localized auxin peaks in concentration-based transport models of the shoot apical meristem. J R Soc Interface 12:20141407

Emerenini BO, Hense BA, Kuttler C, Eberl HJ (2015) A mathematical model of quorum sensing induced biofilm detachment. PLoS ONE 10:e0132385–e0132385

Faver TE (2018) Nanopteron-stegoton traveling waves in mass and spring dimer Fermi–Pasta–Ulam–Tsingou lattices. Ph.D. thesis, Drexel University, Philadelphia, PA, May

Faver TE, Wright JD (2018) Exact diatomic Fermi–Pasta–Ulam–Tsingou solitary waves with optical band ripples at infinity. SIAM J Math Anal 50:182–250

Fendrych M, Leung J, Friml J (2016) TIR1/AFB-Aux/IAA auxin perception mediates rapid cell wall acidification and growth of Arabidopsis hypocotyls. Elife 5:e19048

Fermi E, Pasta J, Ulam S (1955) Studies of nonlinear problems. Lect Appl Math 12:143–56

Friesecke G, Pego RL (1999) Solitary waves on FPU lattices. I. Qualitative properties, renormalization and continuum limit. Nonlinearity 12:1601–1627

Friesecke G, Pego RL (2002) Solitary waves on FPU lattices. II. Linear implies nonlinear stability. Nonlinearity 15:1343–1359

Friesecke G, Pego RL (2004a) Solitary waves on Fermi–Pasta–Ulam lattices. III. Howland-type Floquet theory. Nonlinearity 17:207–227

Friesecke G, Pego RL (2004b) Solitary waves on Fermi–Pasta–Ulam lattices. IV. Proof of stability at low energy. Nonlinearity 17:229–251

Friesecke G, Wattis JAD (1994) Existence theorem for solitary waves on lattices. Commun Math Phys 161:391–418

Ghasemi M, Sonner S, Eberl HJ (2018) Time adaptive numerical solution of a highly non-linear degenerate cross-diffusion system arising in multi-species biofilm modelling. Eur J Appl Math 29:1035–1061

Hajný J, Prát T, Rydza N, Rodriguez L, Tan S, Verstraeten I, Domjan D, Mazur E, Smakowska-Luzan E, Smet W, Mor E, Nolf J, Yang B, Grunewald W, Molnár G, Belkhadir Y, Rybel BD, Friml J (2020) Receptor kinase module targets PIN-dependent auxin transport during canalization. Science 370:550–557

Hajný J, Tan S, Friml J (2022) Auxin canalization: from speculative models toward molecular players. Curr Opin Plant Biol 65:102174

Haskovec J, Jönsson H, Kreusser LM, Markowich P (2019) Auxin transport model for leaf venation. Proc R Soc A 475:20190015

Heisler MG, Jonsson H (2006) Modeling auxin transport and plant development. J Plant Growth Regul 25:302–312

Herrmann M, Matthies K (2015) Asymptotic formulas for solitary waves in the high-energy limit of FPU-type chains. Nonlinearity 28:2767–2789

Hochstrasser D, Mertens F, Büttner H (1989) Energy transport by lattice solitons in α-helical proteins. Phys Rev A 40:2602

Hoffman A, Wright JD (2017) Nanopteron solutions of diatomic Fermi–Pasta–Ulam–Tsingou lattices with small mass-ratio. Physica D 358:33–59

Holloway DM, Wenzel CL (2021) Polar auxin transport dynamics of primary and secondary vein patterning in dicot leaves. in silico Plants 3:diab030

Hupkes HJ, Sandstede B (2010) Travelling pulse solutions for the discrete FitzHugh–Nagumo system. SIAM J Appl Dyn Syst 9:827–882

Johnson MA, Wright JD (2020) Generalized solitary waves in the gravity-capillary Whitham equation. Stud Appl Math 144:102–130

Johnston ST, Baker RE, McElwain DS, Simpson MJ (2017) Co-operation, competition and crowding: a discrete framework linking Allee kinetics, nonlinear diffusion, shocks and sharp-fronted travelling waves. Sci Rep 7:1–19

Jones C, Kopell N, Langer R (1991) Construction of the FitzHugh–Nagumo pulse using differential forms. In: Aris R, Aronson DG, Swinney HL (eds) Patterns and dynamics in reactive media. Springer, New York, pp 101–115

Jönsson H, Heisler M, Shapiro B, Meyerowitz E, Mjolsness E (2006) An auxin-driven polarized transport model for phyllotaxis. Proc Natl Acad Sci USA 103:1633–1638

Julien JD, Pumir A, Boudaoud A (2019) Strain- or stress-sensing in mechanochemical patterning by the phytohormone auxin. Bull Math Biol 81:3342–3361

Keener JP (1987) Propagation and its failure in coupled systems of discrete excitable cells. SIAM J Appl Math 47:556–572

Kevrekidis PG (2011) Non-linear waves in lattices: past, present, future. IMA J Appl Math 76:389–423

Kneuper I, Teale W, Dawson JE, Tsugeki R, Katifori E, Palme K, Ditengou FA (2020) Auxin biosynthesis and cellular efflux act together to regulate leaf vein patterning. J Exp Bot 72:1151–1165

Li Y, van Heijster P, Simpson MJ, Wechselberger M (2021) Shock-fronted travelling waves in a reaction–diffusion model with nonlinear forward–backward–forward diffusion. Physica D 423:132916

Mallet-Paret J (1999) The global structure of traveling waves in spatially discrete dynamical systems. J Dyn Differ Equ 11:49–128

Merks RMH, Van de Peer Y, Inzé D, Beemster GTS (2007) Canalization without flux sensors: a traveling-wave hypothesis. Trends Plant Sci 12:384–390

Mitchison GJ (1980) A model for vein formation in higher plants. Proc R Soc Lond Ser B Biol Sci 207:79–109

Mitchison G (1981) The polar transport of auxin and vein patterns in plants. Philos Trans R Soc Lond B Biol Sci 295:461–471

Moser P (2021) The propagation of auxin waves and wave trains. B.Sc. thesis, Leiden University. https://hdl.handle.net/1887/3197145

Pankov A (2005) Travelling waves and periodic oscillations in Fermi–Pasta–Ulam lattices. Imperial College Press, Singapore

Paque S, Weijers D (2016) Q&a: Auxin: the plant molecule that influences almost anything. BMC Biol 14:67

Razavi MS, Shirani E, Kassab GS (2018) Scaling laws of flow rate, vessel blood volume, lengths, and transit times with number of capillaries. Front Physiol 9:581

Reinhardt D, Pesce E-R, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C (2003) Regulation of phyllotaxis by polar auxin transport. Nature 426:255–260

Rolland-Lagan A-G (2008) Vein patterning in growing leaves: axes and polarities. Curr Opin Genet Dev 18:348–353

Rolland-Lagan A-G, Prusinkiewicz P (2005) Reviewing models of auxin canalization in the context of leaf vein pattern formation in Arabidopsis. Plant J Cell Mol Biol 44:854–865

Sachs T (1975) The induction of transport channels by auxin. Planta 127:201–206

Sandstede B (2002) Stability of travelling waves. In: Fiedler B (ed) Handbook of dynamical systems, vol 2. Elsevier, Amsterdam, pp 983–1055

Scarpella E, Marcos D, Friml J, Berleth T (2006) Control of leaf vascular patterning by polar auxin transport. Genes Dev 20:1015–1027

Schmidt-Nielsen K, Knut S-N (1984) Scaling: why is animal size so important? Cambridge University Press, Cambridge

Shi B, Vernoux T (2018) Patterning at the shoot apical meristem and phyllotaxis. Curr Top Dev Biol 131:81–107

Shih Y-L, Huang L-T, Tu Y-M, Lee B-F, Bau Y-C, Hong CY, lin Lee H, Shih Y-P, Hsu M-F, Lu Z-X, Chen J-S, Chao L (2019) Active transport of membrane components by self-organization of the Min proteins. Biophys J 116:1469–1482

Smith RS, Guyomarc'h S, Mandel T, Reinhardt D, Kuhlemeier C, Prusinkiewicz P (2006) A plausible model of phyllotaxis. Proc Natl Acad Sci USA 103:1301–1306

Sonner S, Efendiev MA, Eberl HJ (2011) On the well-posedness of a mathematical model of quorum-sensing in patchy biofilm communities. Math Methods Appl Sci 34:1667–1684

Stefanov A, Wright JD (2020) Small amplitude traveling waves in the full-dispersion Whitham equation. J Dyn Differ Equ 32:85–99

van Berkel K, de Boer RJ, Scheres B, ten Tusscher K (2013) Polar auxin transport: models and mechanisms. Development 140:2253–2268

Verna C, Ravichandran SJ, Sawchuk MG, Linh NM, Scarpella E (2019) Coordination of tissue cell polarity by auxin transport and signaling. Elife 8:e51061

Walke ML, Farcot E, Traas J, Godin C (2013) The flux-based pin allocation mechanism can generate either canalyzed or diffuse distribution patterns depending on geometry and boundary conditions. PLoS ONE 8:e54802

West GB, Brown JH (2004) Life's universal scaling laws. Phys Today 57:36–43

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.