Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

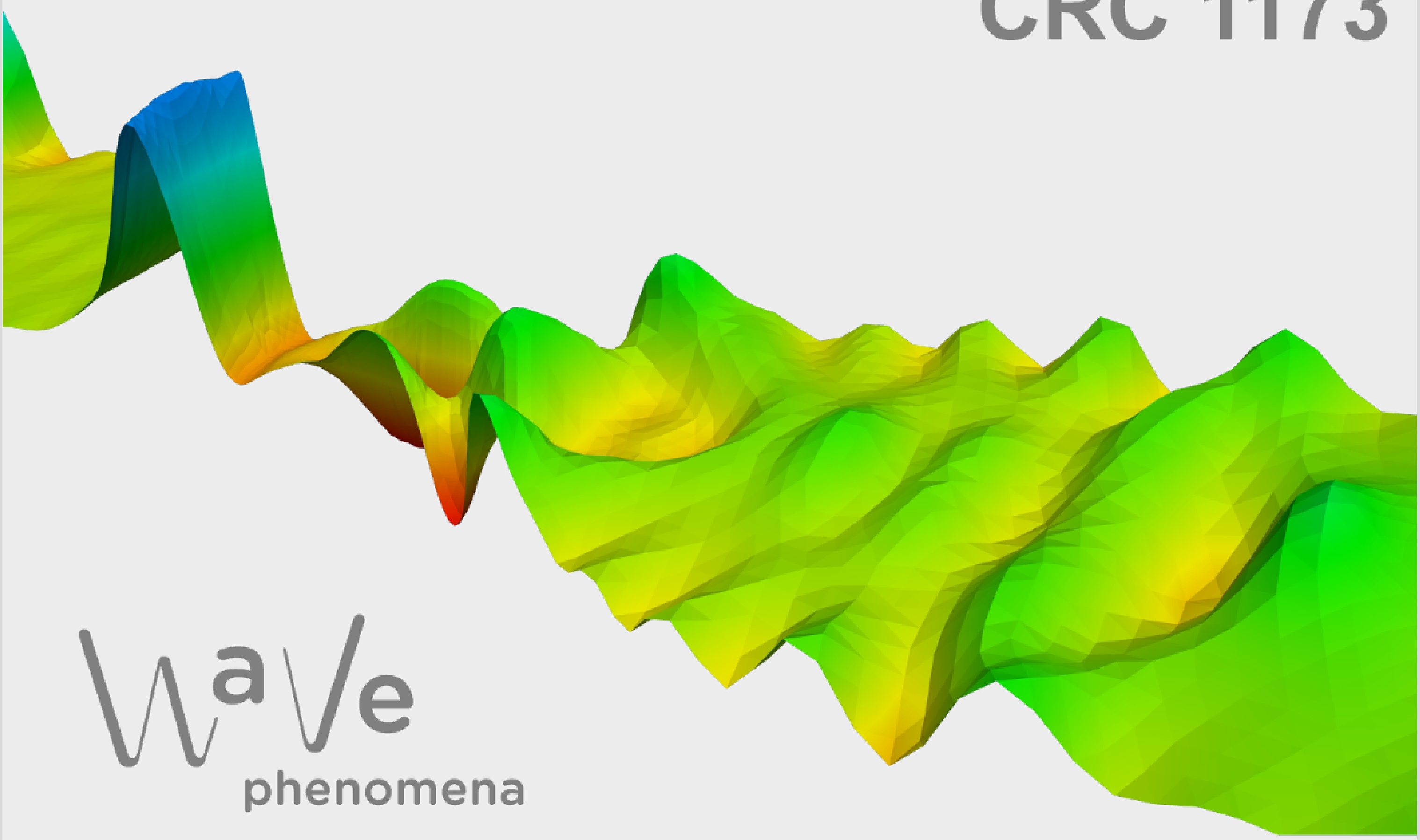
# Validity of the Whitham approximation for a complex cubic Klein–Gordon equation

Sarah Hofbauer, Xian Liao, Guido Schneider

CRC Preprint 2022/64, November 2022

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



**Participating universities**

Universität Stuttgart

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

TU WIEN

**Funded by**

DFG

ISSN 2365-662X

# Validity of the Whitham approximation for a complex cubic Klein-Gordon equation

Sarah Hofbauer[1], Xian Liao[2], Guido Schneider[1]

[1] Institut für Analysis, Dynamik und Modellierung,
Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

[2] Institute for Analysis, KIT, Englerstrasse 2, 76131 Karlsruhe, Germany

November 24, 2022

**Abstract**

The complex cubic Klein-Gordon (ccKG) equation possesses a family of periodic traveling wave solutions. Whitham's modulation equations (WME) can be derived by a multiple scaling perturbation analysis in order to describe slow modulations in time and space of these traveling wave solutions. We prove estimates between true solutions of the ccKG equation and their associated WME approximation. The bounds are obtained in Gevrey spaces and hold independently of the spectral stability of the underlying traveling wave solutions. The proof is based on a suitable choice of variables, Cauchy-Kovalevskaya theory, infinitely many near identity changes of variables, and energy estimates in Gevrey spaces. The analysis for the ccKG equation is more complicated than the analysis for the nonlinear Schrödinger (NLS) equation which has been handled in the existing literature, due to additional curves of eigenvalues leading to an additional oscillatory behavior.

## 1 Introduction

Whitham's modulation equations (WME) can be derived by a multiple scaling perturbation analysis, cf. [Whi74], with a small perturbation parameter

$0 < \delta \ll 1$, in order to describe slow modulations in time and space of periodic traveling wave solutions in dispersive and dissipative systems. In this introduction and also in the subsequent two sections we are going to explain the backgrounds, ideas, as well as some intuitive calculations in numerous remarks.

**Remark 1.1.** So far there are only few approximation results showing that the WME approximations make correct predictions about the dynamics of periodic traveling wave solutions in dispersive and dissipative systems. In [DS09] such estimates were obtained in Gevrey spaces for such waves of the NLS equation

$$\partial_\tau A = i\nu_1 \partial_\xi^2 A + i\nu_2 A|A|^2, \tag{1}$$

with $\tau \in \mathbb{R}$, $\xi \in \mathbb{R}$, $A(\xi, \tau) \in \mathbb{C}$, and coefficients $\nu_1$, $\nu_2 \in \mathbb{R}$, as original system. In [BKS20] such estimates were obtained in Gevrey spaces for such waves of a system of coupled NLS equations as original system. For spectrally stable waves of the NLS equation in [BKZ21] it was shown that WME even make correct predictions for initial conditions in Sobolev spaces. The only approximation result, we are aware of, for dissipative systems can be found in [HdRS21] where the validity of WME was shown for such waves in the amplitude system which appears at the first instability of the Marangoni problem consisting of a Ginzburg-Landau equation coupled to a diffusive conservation law.

**Remark 1.2.** Such approximation results are non-trivial since solutions of order $\mathcal{O}(1)$ have to be bounded on a long $\mathcal{O}(1/\delta)$-time scale. In general solutions of order $\mathcal{O}(1)$ are only bounded on an $\mathcal{O}(1)$-time scale.

**Remark 1.3.** As a next step in the direction of handling modulations of periodic wave trains for general dispersive systems, in this paper, we consider the complex cubic Klein-Gordon (ccKG) equation

$$\partial_t^2 u = \partial_x^2 u - u + \gamma u \left|u\right|^2, \tag{2}$$

with $t, x \in \mathbb{R}$, $\gamma \in \{-1, 1\}$, and $u(x,t) \in \mathbb{C}$. It possesses traveling wave solutions

$$u\left(x,t\right) = e^{r_{q,\mu} + iqx + i\mu t},$$

with $\mu, q, r_{q,\mu} \in \mathbb{R}$ satisfying

$$-\gamma e^{2r_{q,\mu}} = \mu^2 - q^2 - 1.$$

**Remark 1.4.** All these equations have in common that the nonlinearities possess an $S^1$-symmetry, i.e., with $u$ also $ue^{i\phi}$, with $\phi \in \mathbb{R}$, is a solution. As a consequence, the underlying periodic traveling wave solutions are harmonic which easily allows us to extract a local wave number variable which is necessary for the derivation of WME.

**Remark 1.5.** What makes the analysis for the ccKG equation more complicated than the analysis for the NLS equation are two additional curves of eigenvalues which lead to an additional oscillatory behavior. See Figure 1.

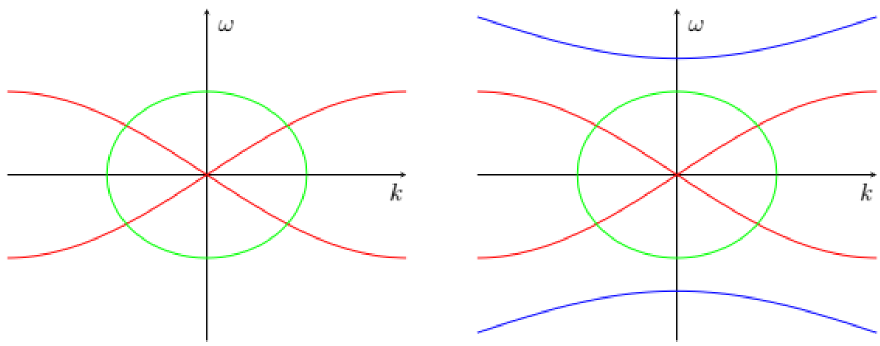



Figure 1: The left panel shows the spectral curves for the NLS equation. The right panel shows the spectral curves for the ccKG equation (2). WME describe the modes in the circles.

**Remark 1.6.** For notational simplicity, here in the introduction, we derive WME for the ccKG equation (2) for the wave train associated to $q = 0$, where we have $\mu^2 = 1 - \gamma e^{2r_{0,\mu}}$. We introduce polar coordinates

$$u = e^{r+i\varphi+r_{0,\mu}+i\mu t}, \tag{3}$$

with $r = r(x,t)$ and $\varphi = \varphi(x,t)$. Using

$$\begin{aligned} \partial_t u &= e^{r+i\varphi+r_{0,\mu}+i\mu t}(\partial_t r + i\partial_t \varphi + i\mu), \\ \partial_t^2 u &= e^{r+i\varphi+r_{0,\mu}+i\mu t}(\partial_t r + i\partial_t \varphi + i\mu)^2 + e^{r+i\varphi+r_{0,\mu}+i\mu t}(\partial_t^2 r + i\partial_t^2 \varphi), \end{aligned}$$

and similar expressions for $\partial_x u$ and $\partial_x^2 u$ we find by separating real and imaginary parts that

$$\begin{aligned} \partial_t^2 r - (\partial_t \varphi + \mu)^2 + (\partial_t r)^2 &= \partial_x^2 r - (\partial_x \varphi)^2 + (\partial_x r)^2 - 1 + \gamma e^{2r+2r_{0,\mu}}, \\ 2\,(\partial_t r)\,(\partial_t \varphi + \mu) + \partial_t^2 \varphi &= 2\,(\partial_x r)\,(\partial_x \varphi) + \partial_x^2 \varphi. \end{aligned}$$

We introduce the local temporal wave number $\vartheta = \partial_t \varphi$ and the local spatial wave number $\psi = \partial_x \varphi$ for which we obtain

$$\begin{aligned}
\partial_t^2 r &= \partial_x^2 r + \vartheta^2 + 2\mu\vartheta - (\partial_t r)^2 - \psi^2 + (\partial_x r)^2 + \gamma e^{2r_{0,\mu}}(e^{2r} - 1), && (4)\\
\partial_t \vartheta &= 2(\partial_x r)\psi + \partial_x \psi - 2(\partial_t r)(\vartheta + \mu), && (5)\\
\partial_t \psi &= \partial_x \vartheta, && (6)
\end{aligned}$$

by using

$$(\partial_t \varphi + \mu)^2 = (\vartheta + \mu)^2 = \vartheta^2 + 2\mu\vartheta + \mu^2 = \vartheta^2 + 2\mu\vartheta + 1 - \gamma e^{2r_{0,\mu}}.$$

For the derivation of WME we make the long wave ansatz

$$(r, \psi, \vartheta)(x, t) = (\check{r}, \check{\psi}, \check{\vartheta})(\delta x, \delta t), \tag{7}$$

with $0 < \delta \ll 1$ a small perturbation parameter. Ignoring higher order terms yields the system

$$\begin{aligned}
0 &= \check{\vartheta}^2 + 2\mu\check{\vartheta} - \check{\psi}^2 + \gamma e^{2r_{0,\mu}}(e^{2\check{r}} - 1), && (8)\\
\partial_T \check{\vartheta} &= 2(\partial_X \check{r})\check{\psi} + \partial_X \check{\psi} - 2(\partial_T \check{r})(\check{\vartheta} + \mu), && (9)\\
\partial_T \check{\psi} &= \partial_X \check{\vartheta}. && (10)
\end{aligned}$$

Since the second equation contains derivatives of $\check{\vartheta}$ and $\check{r}$ w.r.t. $T$, it turns out be advantageous to work with the variables

$$\check{v} = 2\mu\check{\vartheta} + b\check{r} \qquad \text{and} \qquad \check{w} = \check{\vartheta} + 2\mu\check{r}, \tag{11}$$

respectively

$$\check{\vartheta} = (-2\mu\check{v} + b\check{w})/D \qquad \text{and} \qquad \check{r} = (\check{v} - 2\mu\check{w})/D,$$

with $b = 2\gamma e^{2r_{0,\mu}} = 2 - 2\mu^2$ and $D = -4\mu^2 - b = -(6\mu^2 - 2) = 2 - 6\mu^2$. Equation (8) is then of the form

$$\check{v} = f_v(\check{v}, \check{w}, \check{\psi}),$$

with $f_v$ at least quadratic in its arguments. For $\check{w}$ and $\check{\psi}$ sufficiently small this equation can be solved with respect to $\check{v}$, i.e., there exists a nonlinear function $g_v$ such that

$$\check{v} = g_v(\check{w}, \check{\psi}), \tag{12}$$

with $g_v$ at least quadratic in its arguments. Using

$$\partial_T \check{\vartheta} + 2\mu \partial_T \check{r} = 2(\partial_X \check{r})\check{\psi} + \partial_X \check{\psi} - 2(\partial_T \check{r})\check{\vartheta}$$

we rewrite the $\check{\vartheta}$-equation into

$$\begin{aligned}
\partial_T \check{w} &= 2(\partial_X(a_1\check{v} + a_2\check{w}))\check{\psi} + \partial_X \check{\psi} - 2(\partial_T(a_1\check{v} + a_2\check{w}))(a_3\check{v} + a_4\check{w}), \\
&= 2\check{\psi}(\partial_X(a_1\check{v} + a_2\check{w})) + \partial_X \check{\psi} - 2(a_3\check{v} + a_4\check{w})a_2\partial_T \check{w} \\
&\quad -2(a_3\check{v} + a_4\check{w})a_1(\ell_1(\check{w}, \check{\psi})\partial_T \check{w} + \ell_2(\check{w}, \check{\psi})\partial_T \check{\psi}),
\end{aligned}$$

with $\ell_1$ and $\ell_2$ at least linear in its arguments and constant coefficients $a_j$. We can replace $\partial_T \check{\psi}$ by the right hand side of the third equation which is of the form

$$\partial_T \check{\psi} = \partial_X(a_3\check{v} + a_4\check{w}).$$

Hence, for $\check{w}$ and $\check{\psi}$ sufficiently small the second equation can be solved with respect to $\partial_T \check{w}$, i.e., there exists a nonlinear function $g_w$ such that

$$\partial_T \check{w} = g_w(\check{v}, \check{w}, \check{\psi}, \partial_X \check{v}, \partial_X \check{w}, \partial_X \check{\psi}),$$

where $g_w$ is of the form

$$g_w = g_{w,1}(\check{v}, \check{w}, \check{\psi})\partial_X \check{v} + g_{w,2}(\check{v}, \check{w}, \check{\psi})\partial_X \check{w} + g_{w,3}(\check{v}, \check{w}, \check{\psi})\partial_X \check{\psi}.$$

Eliminating $\check{v}$ by the above expression finally yields WME given by

$$\begin{aligned}
\partial_T \check{w} &= g_w(g_v(\check{w}, \check{\psi}), \check{w}, \check{\psi}, \partial_X g_v(\check{w}, \check{\psi}), \partial_X \check{w}, \partial_X \check{\psi}), && (13) \\
\partial_T \check{\psi} &= \partial_X(a_3 g_v(\check{w}, \check{\psi}) + a_4\check{w}), && (14)
\end{aligned}$$

describing the modes in the circles in the right panel of Figure 1, where the right-hand side of the $\check{w}$-equation (13) can be written as

$$g_{w,4}(\check{w}, \check{\psi})\partial_X \check{w} + g_{w,5}(\check{w}, \check{\psi})\partial_X \check{\psi}. \qquad (15)$$

**Remark 1.7.** Depending on the values of $\mu$ and $\gamma$ WME (13)-(14) can be well- or ill-posed in Sobolev spaces. In the first case, it turns out that equivalently the periodic wave train is spectrally stable, in the second case spectrally unstable. The first situation is called the Benjamin-Feir stable and the second situation the Benjamin-Feir unstable case, cf. Remark 1.11 and Section 2.1.

In the following we prove estimates between true solutions of the ccKG equation and their associated WME approximation. The bounds are obtained in Gevrey spaces and hold independently of the spectral stability of the underlying traveling wave solutions.

**Definition 1.8.** *The Gevrey spaces*

$$G_\sigma^m = \{u \in L^2 : \|u\|_{G_\sigma^m} < \infty\}$$

*are Hilbert spaces equipped with the inner product*

$$(u, v)_{G_\sigma^m} = \int_{\mathbb{R}} e^{2\sigma(1+|k|)} \left(1 + |k|^2\right)^m \widehat{u}(k)\overline{\widehat{v}(k)} dk,$$

*for* $\sigma \geq 0$ *and* $m \geq 0$.

**Remark 1.9.** Since the right hand sides of WME (13)-(14) only contain first derivatives local existence and uniqueness of solutions in Gevrey spaces for WME is well known by the Cauchy-Kowalevskaya theorem. See Section 4.

Our approximation result, formulated for $q = 0$, is as follows.

**Theorem 1.10.** *Let* $|\mu| > 1/\sqrt{3}$, $\sigma_0 > 0$ *and* $m \geq 3$. *Then for all* $T_0$ *and* $C_1$ *there exist* $C_2$, $T_1$, $\delta_0 > 0$ *such that for all* $\delta \in (0, \delta_0)$ *the following holds. Let* $(\check{w}_{app}, \check{\psi}_{app}) \in C([0, T_0], G_{\sigma_0}^{m+1}) \cap C^1((0, T_0], G_{\sigma_0}^m)$ *be a solution of WME* (13)-(14) *satisfying*

$$\sup_{T \in [0, T_0]} \|(\check{w}_{app}, \check{\psi}_{app})\|_{G_{\sigma_0}^{m+1}} \leq C_1,$$

*let* $\check{v}_{app}$ *be the corresponding solution to the algebraic equation* (12) *and let* $(\check{r}_{app}, \check{\vartheta}_{app}, \check{\psi}_{app})$ *be the approximation constructed from* $(\check{v}_{app}, \check{w}_{app}, \check{\psi}_{app})$.
*Then there exist solutions* $(r, \vartheta, \psi)$ *of* (4)-(6) *with*

$$\sup_{t \in [0, T_1/\delta]} \sup_{x \in \mathbb{R}} |(r, \vartheta, \psi)(x, t) - (\check{r}_{app}, \check{\vartheta}_{app}, \check{\psi}_{app})(\delta x, \delta t)| \leq C_2 \delta.$$

**Remark 1.11.** The approximation result covers the Benjamin-Feir stable case, $|\mu| \geq 1$, and the Benjamin-Feir unstable case, $|\mu| \in (1/\sqrt{3}, 1)$, cf. Figure 2. For $|\mu| \leq 1/\sqrt{3}$ it cannot be expected that WME make correct predictions, cf. Figure 3 and Remark 9.2.

**Remark 1.12.** As already said, the above validity result is a nontrivial task. The WME approximation and the associated solution are of order $\mathcal{O}(1)$ for $\delta \to 0$. Therefore, a simple application of Gronwall's inequality would only provide the boundedness of the solutions on an $\mathcal{O}(1)$-time scale, but not on the natural $\mathcal{O}(1/\delta)$-time scale of the WME approximation.

**Remark 1.13.** There is a number of counterexamples where formally derived amplitude equations make wrong predictions about the dynamics of original systems on the natural time scale of the amplitude equations, cf. [Sch95, SSZ15, HS20, BSSZ20, FS22].

**Remark 1.14.** Although Theorem 1.10 is not optimal in the sense that the possible approximation time $T_1/\delta$ is possibly smaller than $T_0/\delta$, we do establish an approximation result on the natural $\mathcal{O}(1/\delta)$-time scale of the WME approximation.

The plan of the paper is as follows. In the next section we go on with some further remarks. We start with the Benjamin-Feir instability, i.e., in Section 2.1 we explain that there are stable and unstable wave trains. Modulations of small amplitude wave trains of the ccKG equation (2) can be described by an NLS approximation. Therefore, in Section 2.2 we relate our approximation result to the associated approximation results for the NLS equation. The WME approximation is a long wave limit approximation like the KdV approximation or the inviscid Burgers approximation. Hence, in Section 2.3 we relate our result to other long wave approximation results and formulate the associated approximation results for modulations of wave trains in the ccKG equation (2). Finally, in Section 2.4 we explain the ideas of the proof of Theorem 1.10. In Section 3 we redo some calculations for $q \neq 0$ and plot a few spectral curves which look different from the spectral curves for $q = 0$. The rest of the paper is about the proof of the main theorem 1.10. We use Cauchy-Kovalevskaya theory in Section 4 to obtain local existence and uniqueness of solutions to WME (13)-(14). After diagonalisation of the error equations in Section 5, we have to use infinitely many normal form transformations in Section 6 to get rid of the new oscillatory terms appearing in the right panel of Figure 1, cf. [DKS16]. After some preparations in Section 7 we obtain a system for which in Section 8 we can use energy estimates, similar to the one used for WME in Section 4, to control the solutions close to the wave number $k = 0$. At wave numbers bounded away from $k = 0$ an artificial damping is available to control the solutions due to the use of a time-dependent scale of Gevrey spaces. This finally yields the validity of the main theorem 1.10. We close this paper with Section 9 where we discuss related questions such as difficulties and strategies to obtain estimates in Sobolev spaces. In an Appendix we collect some calculations about the spectral curves plotted subsequently in Figure 5 for the case $q \neq 0$ and give a proof for some estimates used in the previous sections.

# 2 Some further remarks

As in the introduction all further explanations in this section are still made for the wave trains with the wave number $q = 0$.

## 2.1 The Benjamin-Feir instability

In this section we explain that there are stable and unstable wave trains. The so called Benjamin-Feir instability is a long wave instability.

**Remark 2.1.** The linearization of (4)-(6) is given by

$$\begin{aligned} \partial_t^2 r &= \partial_x^2 r + 2\mu\vartheta + 2\gamma e^{2r_{0,\mu}} r, \\ \partial_t \vartheta &= \partial_x \psi - 2\mu(\partial_t r), \\ \partial_t \psi &= \partial_x \vartheta, \end{aligned}$$

where for $q = 0$ we have

$$-\gamma e^{2r_{0,\mu}} = \mu^2 - 1.$$

A Fourier ansatz yields the dispersion relations

$$\begin{aligned} -\omega^2 \widehat{r} &= -k^2 \widehat{r} + 2\mu\widehat{\vartheta} - 2(\mu^2 - 1)\widehat{r}, \\ i\omega\widehat{\vartheta} &= ik\widehat{\psi} - 2\mu(i\omega\widehat{r}), \\ i\omega\widehat{\psi} &= ik\widehat{\vartheta}. \end{aligned}$$

For $k = 0$ we find the dispersion relation

$$-\omega^2(-\omega^2 + 6\mu^2 - 2) = 0.$$

Hence, we have two eigenvalues zero and two eigenvalues bounded away from zero, which is exactly the spectral situation necessary for the derivation of WME.

**Remark 2.2.** For the calculation of the eigenvalues we have to solve

$$\det \begin{pmatrix} \omega^2 - k^2 - 2(\mu^2 - 1) & 2\mu & 0 \\ -2\mu i\omega & -i\omega & ik \\ 0 & ik & -i\omega \end{pmatrix} = 0.$$

We find

$$(\omega^2 - k^2 - 2(\mu^2 - 1))(-\omega^2 + k^2) + 4\mu^2\omega^2 = 0,$$

respectively,

$$\omega^4 - \omega^2(2k^2 + 6\mu^2 - 2) + k^4 + 2k^2(\mu^2 - 1) = 0,$$

and so

$$2\omega_{1,2}^2 = (2k^2 + 6\mu^2 - 2) \pm \sqrt{(2k^2 + 6\mu^2 - 2)^2 - 4(k^4 + 2k^2(\mu^2 - 1))},$$

which yields

$$\omega_{1,2}^2 = (k^2 + 3\mu^2 - 1) \pm \sqrt{4k^2\mu^2 + (3\mu^2 - 1)^2}.$$

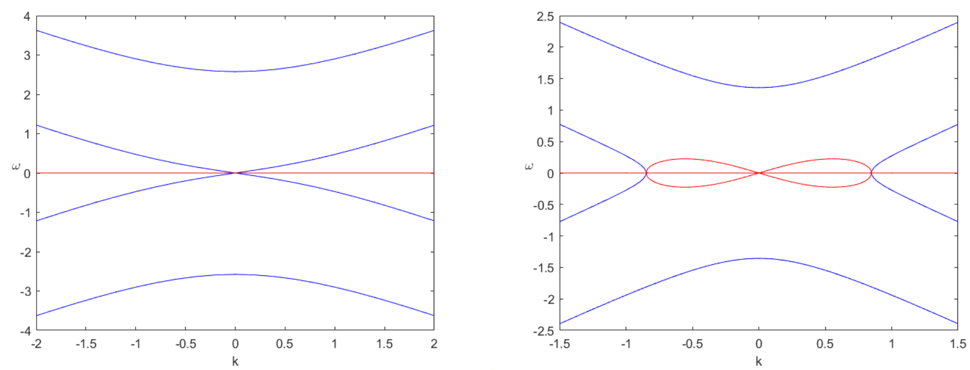



Figure 2: The left panel shows the spectral curves $\pm i\omega_{1,2}$ of (4)-(6) for $\mu = 1.2 \geq 1$. They are purely imaginary (in blue) since the real part (in red) of $\pm i\omega_{1,2}$ vanishes identically. The right panel shows the imaginary part of the spectral curves of (4)-(6) for $\mu = 0.8 \in (1/\sqrt{3}, 1)$ (in blue). The eigenvalues with vanishing imaginary part have non-zero real part (in red), i.e., there is a so called Benjamin-Feir instability.

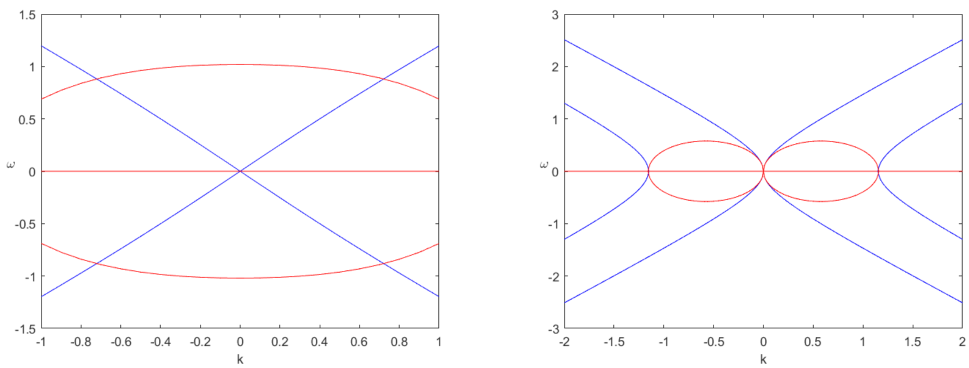



Figure 3: The left panel shows the real (in red) and imaginary part (in blue) of the spectral curves $\pm i\omega_{1,2}$ of (4)-(6) for $\mu = 0.4 < 1/\sqrt{3}$. The right panel shows the same at the threshold $\mu = 1/\sqrt{3}$, cf. Remark 1.11.

Figure 2 and Figure 3 show that the traveling wave solutions are only spectrally stable for $|\mu| \geq 1$. In this region an approximation result with initial conditions in Sobolov spaces would be desirable. For $|\mu| \in (1/\sqrt{3}, 1)$ we have a Benjamin-Feir instability and so only an approximation result with initial conditions in Gevrey spaces can be expected. For $|\mu| < 1/\sqrt{3}$ the modes associated to $\omega = 0$ are imaginary again, however, for $k = 0$ there are now modes with positive growth rates and so it cannot be expected that the WME approximation makes correct predictions, cf. Figure 3.

## 2.2 The NLS limit

In the following remarks we explain how the previous WME approximation results from [DS09, BKZ21] for the NLS equation are related to our result for the ccKG equation stated in Theorem 1.10.

**Remark 2.3.** Inserting the multiple scaling ansatz

$$u(x,t) = \varepsilon A(\varepsilon(x-ct), \varepsilon^2 t)e^{i(k_0 x - \omega_0 t)}$$

into the ccKG equation (2) and then equating the coefficients in front of $\varepsilon^n e^{i(k_0 x - \omega_0 t)}$ to zero for $n = 1, 2, 3$ gives the dispersion relation

$$\omega_0^2 = k_0^2 + 1,$$

the group velocity

$$c = k_0/\omega_0,$$

and shows that in lowest order $A$ has to satisfy the NLS equation

$$2i\omega_0 \partial_\tau A = (1 - c^2)\partial_\xi^2 A + \gamma A|A|^2.$$

**Remark 2.4.** The normalized NLS equation

$$\partial_\tau U = i\partial_\xi^2 U + \gamma i U|U|^2, \tag{16}$$

with $\tau \in \mathbb{R}$, $\xi \in \mathbb{R}$, $U(\xi, \tau) \in \mathbb{C}$, and $\gamma = \pm 1$ possesses traveling wave solutions

$$U_{\varrho,q}(\xi, \tau) = e^{\varrho + i\widetilde{\omega}\tau + i\phi_0 + iq\xi}, \tag{17}$$

with $\varrho, \widetilde{\omega}, \phi_0, q \in \mathbb{R}$ satisfying

$$\widetilde{\omega} = -q^2 + \gamma e^{2\varrho}. \tag{18}$$

WME describe slow modulations in time and in space of these waves. For notational simplicity in this remark we restrict ourselves in the following to modulations of the wave train to $q = \varrho = \phi_0 = 0$. For the derivation we introduce polar coordinates, with radius also in exponential form, in a uniformly rotating frame. The NLS equation in such polar coordinates

$$U(\xi, \tau) = e^{r(\xi,\tau) + i\phi(\xi,\tau) + i\gamma\tau}$$

is then given by

$$\begin{aligned} \partial_\tau r &= -\partial_\xi^2 \phi - 2(\partial_\xi r)(\partial_\xi \phi), & (19) \\ \partial_\tau \phi &= \partial_\xi^2 r - (\partial_\xi \phi)^2 + (\partial_\xi r)^2 + \gamma(e^{2r} - 1). & (20) \end{aligned}$$

Introducing the local spatial wave number $\partial_\xi \phi = \psi$ yields

$$\begin{aligned} \partial_\tau r &= -\partial_\xi \psi - 2(\partial_\xi r)\psi, & (21) \\ \partial_\tau \psi &= \partial_\xi^3 r - \partial_\xi(\psi^2) + \partial_\xi(\partial_\xi r)^2 + 2\gamma e^{2r}\partial_\xi r. & (22) \end{aligned}$$

The long wave ansatz

$$r(\xi, \tau) = \check{r}(\delta\xi, \delta\tau), \quad \psi(\xi, \tau) = \check{\psi}(\delta\xi, \delta\tau), \tag{23}$$

with $0 < \delta \ll 1$, leads to

$$\begin{aligned} \partial_T \check{r} &= -\partial_X \check{\psi} - 2(\partial_X \check{r})\check{\psi}, & (24) \\ \partial_T \check{\psi} &= \delta^2 \partial_X^3 \check{r} - \partial_X(\check{\psi}^2) + \delta^2 \partial_X (\partial_X \check{r})^2 + 2\gamma e^{2\check{r}} \partial_X \check{r}, & (25) \end{aligned}$$

where $T = \delta\tau$ and $X = \delta\xi$. Ignoring the higher order terms gives WME

$$\begin{aligned} \partial_T \check{r} &= -\partial_X \check{\psi} - 2(\partial_X \check{r})\check{\psi}, & (26) \\ \partial_T \check{\psi} &= -\partial_X(\check{\psi}^2) + 2\gamma e^{2\check{r}} \partial_X \check{r}. & (27) \end{aligned}$$

In case $\gamma = 1$ we recover the Benjamin-Feir instability, cf. [BM95], i.e., the linearization

$$\partial_T \check{r} = -\partial_X \check{\psi}, \qquad \partial_T \check{\psi} = 2\gamma \partial_X \check{r},$$

is an elliptic system and so ill-posed in Sobolev spaces for $\gamma = 1$. However, even for $\gamma = 1$ WME still possess local in time solutions in the space of functions analytic in a strip around the real axis in the complex plane, i.e., in so called Gevrey spaces by the Cauchy-Kowalevskaya theorem.

**Remark 2.5.** We derive WME for the ccKG equation (2) in the NLS limit introduced in the two previous remarks. In this limit we consider modulations of the traveling wave solution

$$u(x,t) = \varepsilon e^{i\omega t}$$

of the ccKG equation (2) with

$$\omega = \sqrt{1 - \gamma\varepsilon^2} = 1 - \frac{\gamma}{2}\varepsilon^2 + \mathcal{O}(\varepsilon^4).$$

For the derivation of WME we introduce polar coordinates

$$u(x,t) = \varepsilon e^{r(\varepsilon x, \varepsilon^2 t) + i\varphi(\varepsilon x, \varepsilon^2 t) + i\omega t}.$$

Using

$$\begin{aligned} \partial_t u &= \varepsilon e^{r + i\varphi + it - i\gamma\varepsilon^2 t/2 + \mathcal{O}(\varepsilon^4)} (\varepsilon^2 \partial_\tau r + i\varepsilon^2 \partial_\tau \varphi + i - i\varepsilon^2\gamma/2 + \mathcal{O}(\varepsilon^4)), \\ \partial_t^2 u &= \varepsilon e^{r + i\varphi + it - i\gamma\varepsilon^2 t/2 + \mathcal{O}(\varepsilon^4)} (\varepsilon^2 \partial_\tau r + i\varepsilon^2 \partial_\tau \varphi + i - i\varepsilon^2\gamma/2 + \mathcal{O}(\varepsilon^4))^2 \\ &\quad + \varepsilon e^{r + i\varphi + it - i\gamma\varepsilon t/2 + \mathcal{O}(\varepsilon^4)} (\varepsilon^4 \partial_\tau^2 r + i\varepsilon^4 \partial_\tau^2 \varphi), \end{aligned}$$

and

$$\begin{aligned}
\partial_x u &= \varepsilon e^{r+i\varphi+it-i\gamma\varepsilon^2 t/2+\mathcal{O}(\varepsilon^4)}(\varepsilon\partial_\xi r + i\varepsilon\partial_\xi\varphi), \\
\partial_x^2 u &= \varepsilon e^{r+i\varphi+it-i\gamma\varepsilon^2 t/2+\mathcal{O}(\varepsilon^4)}(\varepsilon\partial_\xi r + i\varepsilon\partial_\xi\varphi)^2 \\
&\quad +\varepsilon e^{r+i\varphi+it-i\gamma\varepsilon^2 t/2+\mathcal{O}(\varepsilon^4)}(\varepsilon^2\partial_\xi^2 r + i\varepsilon^2\partial_\xi^2\varphi),
\end{aligned}$$

and separating real and imaginary parts we find

$$\begin{aligned}
&\varepsilon^4\partial_\tau^2 r - \left(\varepsilon^2\partial_\tau\varphi + 1 - \gamma\varepsilon^2/2 + \mathcal{O}(\varepsilon^4)\right)^2 + \left(\varepsilon^2\partial_\tau r\right)^2 \\
&\qquad = \varepsilon^2\partial_\xi^2 r - (\varepsilon\partial_\xi\varphi)^2 + (\varepsilon\partial_\xi r)^2 - 1 + \varepsilon^2\gamma e^{2r}, \\
&2\varepsilon^2\left(\partial_\tau r\right)\left(\varepsilon^2\partial_\tau\varphi + 1 - \gamma\varepsilon^2/2 + \mathcal{O}(\varepsilon^4)\right) + \varepsilon^4\partial_\tau^2\varphi \\
&\qquad = 2\left(\varepsilon\partial_\xi r\right)\left(\varepsilon\partial_\xi\varphi\right) + \varepsilon^2\partial_\xi^2\varphi,
\end{aligned}$$

where $\xi = \varepsilon x$ and $\tau = \varepsilon^2 t$. In lowest order we obtain

$$\begin{aligned}
-2\partial_\tau\varphi &= \partial_\xi^2 r - (\partial_\xi\varphi)^2 + (\partial_\xi r)^2 + \gamma(e^{2r} - 1), \\
2\partial_\tau r &= 2\left(\partial_\xi r\right)\left(\partial_\xi\varphi\right) + \partial_\xi^2\varphi,
\end{aligned}$$

which corresponds up to some rescaling to (19)-(20). Following the rest of Remark 2.4 allows us to recover WME in the NLS limit for the ccKG equation (2).

**Remark 2.6.** In order to relate the equations (19)-(20) to WME (13)-(14) for the ccKG equation we use the variables from above, namely $\psi = \partial_\xi\phi$ and $\vartheta = \partial_\tau\phi$. We find

$$\begin{aligned}
\partial_\tau r &= -\partial_\xi\psi - 2(\partial_\xi r)\psi, \\
\vartheta &= \partial_\xi^2 r - \psi^2 + (\partial_\xi r)^2 + \gamma(e^{2r} - 1), \\
\partial_\tau\psi &= \partial_\xi\vartheta.
\end{aligned}$$

Diffentiating the $\vartheta$-equation w.r.t. $\tau$ and replacing then $\partial_\tau r$ and $\partial_\tau\psi$ on the new right-hand side by the right-hand sides of the $\partial_\tau r$- and $\partial_\tau\psi$-equations gives the systems from above.

**Remark 2.7.** Finally we show how the Benjamin-Feir instability criterion for the ccKG equation and for the associated NLS equation fit together. With the notations of Remark 1.6 we have $\mu^2 = 1 - \gamma e^{2r_{0,\mu}}$ and get

$$\check{r}^*(\check{\vartheta}, \check{\psi}) = -\frac{\mu}{\gamma e^{2r_{0,\mu}}}\check{\vartheta} + h.o.t..$$

We find

$$\partial_T \check{r}^* = -\frac{\mu}{\gamma e^{2r_{0,\mu}}}\partial_T \check{\vartheta} + h.o.t.,$$
$$\partial_X \check{r}^* = -\frac{\mu}{\gamma e^{2r_{0,\mu}}}\partial_X \check{\vartheta} + h.o.t..$$

Inserting this in the above equations (8)-(10) yields

$$\begin{aligned}\partial_T \check{\vartheta} &= 2\left(-\frac{\mu}{\gamma e^{2r_{0,\mu}}}\partial_X \check{\vartheta}\right)\check{\psi} + \partial_X \check{\psi} - 2\left(-\frac{\mu}{\gamma e^{2r_{0,\mu}}}\partial_T \check{\vartheta}\right)(\check{\vartheta}+\mu) \\ &= -\frac{2\mu}{\gamma e^{2r_{0,\mu}}}\left(\partial_T \check{\psi}\right)\check{\psi} + \partial_X \check{\psi} + \frac{2\mu}{\gamma e^{2r_{0,\mu}}}\left(\partial_T \check{\vartheta}\right)\check{\vartheta} + \frac{2\mu}{\gamma e^{2r_{0,\mu}}}\left(\partial_T \check{\vartheta}\right)\mu.\end{aligned}$$

The linearization of this equation is given by

$$\partial_T \check{\vartheta} = \partial_X \check{\psi} + \frac{2\mu^2}{\gamma e^{2r_{0,\mu}}}\partial_T \check{\vartheta} = \partial_X \check{\psi} + \frac{2\left(1-\gamma e^{2r_{0,\mu}}\right)}{\gamma e^{2r_{0,\mu}}}\partial_T \check{\vartheta}.$$

Hence we find

$$\partial_T \check{\vartheta} = \frac{1}{\left(3 - \frac{2}{\gamma e^{2r_{0,\mu}}}\right)}\partial_X \check{\psi} = \frac{1}{\left(3 - 2\gamma^{-1}e^{-2r_{0,\mu}}\right)}\partial_X \check{\psi}.$$

In the NLS limit we have $r_{0,\mu} \to -\infty$ and so $3 \ll e^{-2r_{0,\mu}}$. Since additionally $\partial_T \check{\psi} = \partial_X \check{\vartheta}$, then $\gamma$ decides about the stability and instability. The so called Benjamin-Feir instability occurs for $3 - 2\gamma^{-1}e^{-2r_{0,\mu}} < 0$ and is possible for $\gamma = 1$. In this case the system is ill-posed in Sobolev spaces.

## 2.3 Long wave limit approximations

WME appear as a long wave approximation. Other long wave approximations are the KdV approximation or the inviscid Burgers equation. The KdV approximation describes long waves of amplitude $\mathcal{O}(\delta^2)$ on an $\mathcal{O}(1/\delta^3)$-time scale whereas, as we have seen, the WME approximation describes long waves of amplitude $\mathcal{O}(1)$ on an $\mathcal{O}(1/\delta)$-time scale.

**Remark 2.8.** In order to obtain a KdV equation

$$\partial_T A = \nu_1 \partial_X^3 A + \nu_2 A \partial_X A, \tag{28}$$

with coefficients $\nu_1$, $\nu_2 \in \mathbb{R}$, for (4)-(6) with $|\mu| > 1$ we make the ansatz

$$\begin{pmatrix} r_{kdv} \\ \vartheta_{kdv} \\ \psi_{kdv} \end{pmatrix} (x,t) = \delta^2 A(\delta(x-ct), \delta^3 t)V,$$

where $c \in \mathbb{R}$ is the group velocity and $V \in \mathbb{R}^3$ an eigenvector to the eigenvalue 0 associated to one of the two curves $\omega_{\pm 1}$ plotted in the circle of the right panel of Figure 1. Then similar to Theorem 1.10 the following approximation result can be established.

**Theorem 2.9.** *Let $|\mu| > 1$, $\sigma_0 > 0$ and $m \geq 5$. Then for all $T_0$ and $C_1$ there exist $C_2$, $T_1$, $\delta_0 > 0$ such that for all $\delta \in (0, \delta_0)$ the following holds. Let $A \in C([0,T_0], G^{m+3}_{\sigma_0}) \cap C^1((0,T_0], G^m_{\sigma_0})$ be a solution of the KdV equation* (28) *satisfying*

$$\sup_{T \in [0,T_0]} \|A\|_{G^{m+3}_{\sigma_0}} \leq C_1.$$

*Then there exist solutions $(r, \vartheta, \psi)$ of* (4)-(6) *with*

$$\sup_{t \in [0,T_1/\delta^3]} \sup_{x \in \mathbb{R}} |(r,\vartheta,\psi)(x,t) - (r_{kdv}, \vartheta_{kdv}, \psi_{kdv})(x,t)| \leq C_2 \delta^3.$$

**Remark 2.10.** We explain in the subsequent Remark 2.15 how the proof of Theorem 1.10 has to be modified for proving Theorem 2.9. We refrain from formulating a similar result for the approximation by an inviscid Burgers equation, cf. [BDS19].

**Remark 2.11.** The spectral situation, plotted in the left panel of Figure 1, appears for various systems with a spatially homogeneous background state and so for this spectral situation various KdV approximation results exist, for instance for the water wave problem, cf. [Cra85, SW00b, Dül12], or the FPU-system, cf. [SW00a]. Less results are available for the spectral situation plotted in the right panel of Figure 1. Only recently, methods have been developed for the description of the long wave limit of spatially homogeneous systems by KdV approximations, cf. [CS11, Sch20], and WME approximations for such systems. In the justification analysis the new oscillatory modes are eliminated by some normal form transformations. In the justification analysis of the WME approximation a new serious difficulty occurs, namely, the fact that due to the scaling of the WME ansatz infinitely many normal form transformations have to be performed, cf. [DKS16]. In

[BDS19] for a Boussinesq equation with spatially periodic coefficients the validity of the WME approximation was established with a suitable chosen energy.

**Remark 2.12.** The spectral situation, plotted in Figure 1 a), also appears in the situation described in Remark 1.1, and so beside the already mentioned WME approximation results, cf. [DS09, BKZ21] also KdV approximation results, cf. [BGSS09, BGSS10, CR10, CDS14], do exist for the NLS equation. However, slow modulations in time and space of periodic traveling wave solutions with a spectral situation as plotted in the right panel of Figure 1 have not been considered before.

## 2.4 Idea of the proof

**Remark 2.13.** The strategy of the proof is as follows. By Cauchy-Kovalevskaya theory in Gevrey spaces we have local existence and uniqueness of solutions to WME (4)-(6). Another application of the Cauchy-Kovalevskaya theory yields the local existence of higher-order approximations in Gevrey spaces. These higher-order approximations are necessary for the proof of the main result, Theorem 1.10. The solutions of the error equations are controlled with methods from [DKS16], described above. Since we have to perform infinitely many normal form transformations we have to show the convergence of this procedure. Energy estimates for the limit system provide the final argument to finish the proof of Theorem 1.10.

**Remark 2.14.** Although our proof is based on the overall idea of [DKS16], there are a number of differences between the analysis from [DKS16] and the analysis of the present paper. In the present paper the normal form transformations are only made in a neighborhood of wave numbers at $k = 0$. Therefore, in the present paper the validity of non-resonance conditions is only necessary in this neighborhood but not on the whole real line like in [DKS16]. However, by this restriction due to some incompatibility of some Fourier modes supports infinitely many new terms are created whose convergence additionally has to be shown. Moreover, since the Benjamin-Feir unstable situation is included the estimates from [DKS16] has to be transfered from Sobolev spaces to Gevrey spaces.

**Remark 2.15.** The KdV approximation result in Theorem 2.9 can be proven similarly as the Whitham approximation result in Theorem 1.10. However, due to the smaller size of the solutions one normal form transformation is sufficient for the elimination of the oscillatory terms. Finally we can use the same energy estimates as for the WME approximation.

# 3 The case $q \neq 0$

In this section we derive the evolution equations in case $q \neq 0$ and investigate the linear stability of the associated wave trains. Calculations for determining the stability regions in the $(\mu, q)$-parameter plane can be found in Appendix A.1.

## 3.1 The evolution equations

In this section we redo the calculations from Remark 1.6 for the case $q \neq 0$.

**Remark 3.1.** We introduce polar coordinates

$$u = e^{r+i\varphi+r_{q,\mu}+i\mu t+iqx},$$

with $r = r(x,t)$ and $\varphi = \varphi(x,t)$. Inserting this into the ccKG equation (2) and separating real and imaginary parts finally gives

$$\begin{aligned}
\partial_t^2 r &= -(\partial_t r)^2 + (\partial_t \varphi)^2 + 2\mu\partial_t\varphi + (\partial_x r)^2 - (\partial_x\varphi)^2 - 2q\partial_x\varphi + \partial_x^2 r \\
&\quad +\gamma e^{2r_{q,\mu}}\left(e^{2r}-1\right), \\
\partial_t^2\varphi &= -2\partial_t r\partial_t\varphi - 2\mu\partial_t r + 2\partial_x r\partial_x\varphi + 2q\partial_x r + \partial_x^2\varphi.
\end{aligned}$$

As in Remark 1.6 we introduce the local spatial wave number $\psi = \partial_x\varphi$ and the local temporal wave number $\vartheta = \partial_t\varphi$ for which we obtain the evolutionary system

$$\begin{aligned}
\partial_t^2 r &= \partial_x^2 r + \vartheta^2 + 2\mu\vartheta - (\partial_t r)^2 - \psi^2 + (\partial_x r)^2 - 2q\psi \qquad (29)\\
&\quad +\gamma e^{2r_{q,\mu}}\left(e^{2r}-1\right), \\
\partial_t\vartheta &= 2(\partial_x r)(\psi+q) + \partial_x\psi - 2(\partial_t r)(\vartheta+\mu), \qquad (30)\\
\partial_t\psi &= \partial_x\vartheta. \qquad (31)
\end{aligned}$$

## 3.2 Linear stability analysis

In this section we redo the linear stability analysis from Section 2.1 for the periodic wave trains in case $q \neq 0$, i.e., we consider the linear stability of $(r, \vartheta, \psi) = (0, 0, 0)$ of (29)-(31).

**Remark 3.2.** The linearization of (29)-(31) at the origin is given by

$$\begin{aligned}\partial_t^2 r &= \partial_x^2 r + 2\mu\vartheta - 2q\psi + 2\gamma e^{2r_{0,\mu}} r,\\ \partial_t \vartheta &= \partial_x \psi - 2\mu(\partial_t r) + 2q(\partial_x r),\\ \partial_t \psi &= \partial_x \vartheta,\end{aligned}$$

which yields the spectral problem

$$\begin{aligned}-\omega^2 \widehat{r} &= -k^2 \widehat{r} + 2\mu\widehat{\vartheta} - 2q\widehat{\psi} - 2(\mu^2 - 1 - q^2)\widehat{r},\\ i\omega\widehat{\vartheta} &= ik\widehat{\psi} - 2\mu(i\omega\widehat{r}) + 2qik\widehat{r},\\ i\omega\widehat{\psi} &= ik\widehat{\vartheta}\end{aligned}$$

where we used $\mu^2 = 1 + q^2 - \gamma e^{2r_{q,\mu}}$.

**Remark 3.3.** For the calculation of the eigenvalues we have to solve

$$\det \begin{pmatrix} \omega^2 - k^2 - 2(\mu^2 - 1 - q^2) & 2\mu & -2q \\ -2\mu i\omega + 2qik & -i\omega & ik \\ 0 & ik & -i\omega \end{pmatrix} = 0.$$

We find

$$(\omega^2 - k^2 - 2(\mu^2 - 1 - q^2))(-\omega^2 + k^2) - (-2\mu i\omega + 2qik)^2 = 0,$$

respectively

$$\omega^4 - \omega^2(2k^2 + 6\mu^2 - 2 - 2q^2) + \omega(8\mu qk) + k^4 + 2k^2(\mu^2 - 1 - 3q^2) = 0,$$

which no longer can be solved explicitly w.r.t. $\omega$.

**Remark 3.4.** Figure 4 shows the different stability/instability regions in the $(\mu, q)$-parameter plane. In the (yellow) area $\mathcal{P}_{\text{stab}}$ the spectral curves show a similar behavior as the ones in the left panel of Figure 2. In that case WME approximations can be derived with the same techniques as for $q = 0$. In the (white) area $\mathcal{P}_{\text{rest}}$ there are eigenvalues with positive real part at the wave number $k = 0$. As we will see below in this region it cannot be expected that the WME approximation makes correct predictions. A typical spectral curve for the parameter region $\mathcal{P}_{\text{benj}}$ is shown in the left panel of Figure 5. It shows a Benjamin-Feir instability for $q \neq 0$. The right panel of Figure 5 shows a spectrally stable situation which does not occur in this form for $q = 0$. Since the derivation of WME needs a spectral situation as shown in Figure 2 we concentrate in the following on parameters outside the parameter region $\mathcal{P}_{\text{rest}}$, cf. Remark 9.2.

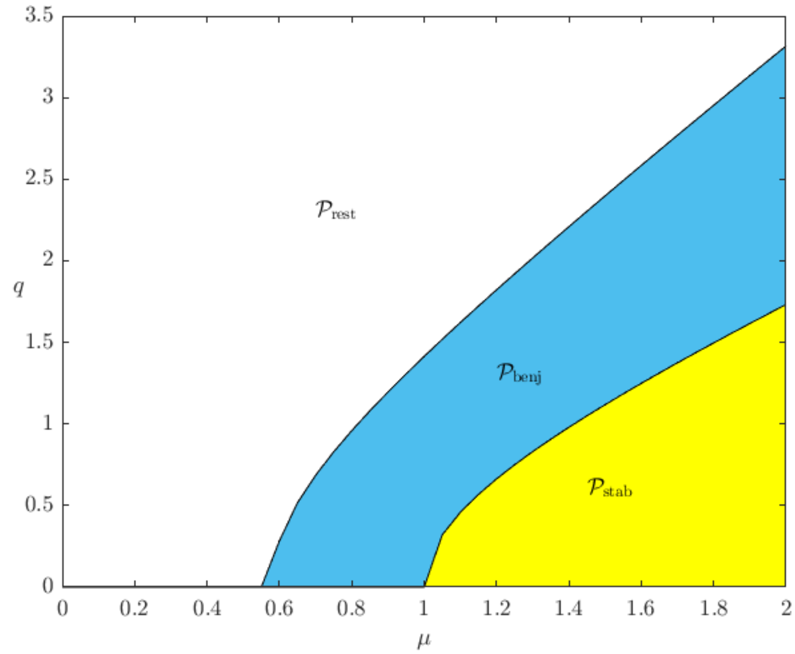



Figure 4: In the $(\mu, q)$-parameter plane we identify regions where the spectral curves look qualitatively different. The parameter region $\mathcal{P}_{\text{benj}}$ is determined by the inequality $1 + q^2 - 3\mu^2 \leq 0$ and the parameter region $\mathcal{P}_{\text{stab}}$ by $1 + 3q^2 - \mu^2 \leq 0$. See Appendix A.1 for some details.

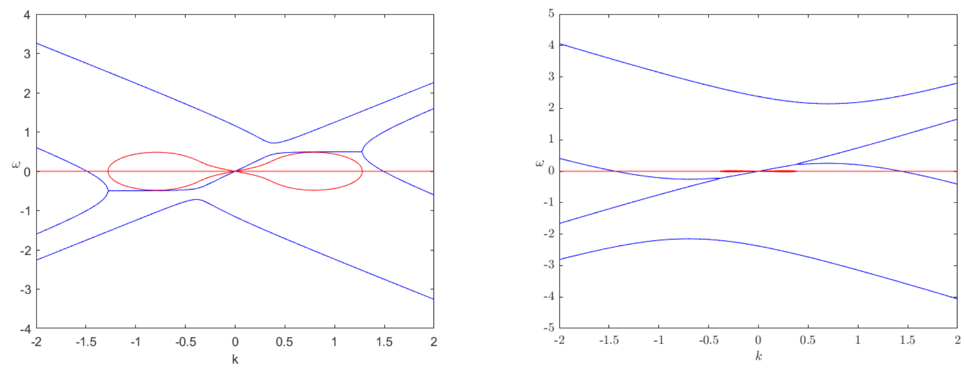



Figure 5: The left panel shows the real (in red) and imaginary part (in blue) of the spectral curves $\pm i\omega_{1,2}$ of (29)-(31) for $\mu = 0.8$ and $q = 0.5$ which is located in $\mathcal{P}_{\text{benj}}$. The right panel shows the same for $\mu = 1.2$ and $q = 0.7$ which is close to the boundary of $\mathcal{P}_{\text{benj}}$ and $\mathcal{P}_{\text{stab}}$.

# 4 The improved WME approximation

For estimating the error made by the WME approximation we need that the residual terms, i.e., the terms which do not cancel after inserting the WME approximation into the ccKG equation (2) are sufficiently small. The residual can be made smaller by adding higher order terms to the WME approximation. This section contains the construction of such an improved WME approximation. The subsequent analysis is an adaption of [HdRS21, Section 2]. The local existence and uniqueness of solutions of the approximation equations is guaranteed by an application of the Cauchy-Kovalevskaya theory in Gevrey spaces.

## 4.1 Some preparations

In the next Remark 4.1 we collect some inequalities which we use in the following.

**Remark 4.1.** a) We use that $G_\sigma^m$ is an algebra for $m > 1/2$. In addition, if $u, v \in G_\sigma^m$ then $uv \in G_\sigma^m$ and

$$\|uv\|_{G_\sigma^m} \leq C_m \|u\|_{G_\sigma^m} \|v\|_{G_\sigma^m}, \tag{32}$$

where the constant $C_m > 0$ is independent of $\sigma \geq 0$. In case that $u$ and $v$ are vector-valued, the product is replaced by an inner product on $\mathbb{R}^d$. Formula (32) can be improved to

$$\|uv\|_{G_\sigma^{m_1}} \leq C_{m_1,m_2}(\|u\|_{G_\sigma^{m_1}}\|v\|_{G_\sigma^{m_2}} + \|u\|_{G_\sigma^{m_2}}\|v\|_{G_\sigma^{m_1}}), \tag{33}$$

which holds for all $\sigma \geq 0$ and $m_j > 1/2$ for $j = 1, 2$ where the constant $C_{m_1,m_2}$ is independent of $\sigma \geq 0$.

b) Let $\phi$ be any entire function $\phi$ with $\phi(0) = 0$. Then for any $m > 1/2$ there exists an entire function $\phi_m(z)$ which is monotonically increasing on $\mathbb{R}_+$ and satisfies $\phi_m(0) = 0$ such that we have

$$\|\phi(u)\|_{G_\sigma^m} \leq \phi_m(\|u\|_{G_\sigma^m}) \tag{34}$$

for all $u \in G_\sigma^m$.

c) Functions $u \in G_\sigma^m$ can be extended to functions that are analytic on the strip $\{z \in \mathbb{C} : |\mathrm{Im}(z)| < \sigma\}$ by the Paley-Wiener Theorem, cf. [RS75, Theorem IX.13]. It is easy to see that for any $\sigma_1 > \sigma_2 \geq 0$ and any $m \geq 0$ we have the continuous embedding $G_{\sigma_1}^0 \hookrightarrow G_{\sigma_2}^m$.

d) Since $\|u(\delta x)\|_{L^2(dx)} = \mathcal{O}(\delta^{-1/2})$ the WME approximation will be of order $\mathcal{O}(\delta^{-1/2})$ in the $L^2$-based spaces $G_\sigma^m$. In order to estimate the WME approximation without this loss of powers of $\delta$ we introduce

**Definition 4.2.** *The spaces*

$$\mathcal{W}_\sigma^m = \{u \in C_b^0 : \|u\|_{\mathcal{W}_\sigma^m} < \infty\}$$

*are equipped with the norm*

$$\|u\|_{\mathcal{W}_\sigma^m} = \int_{\mathbb{R}} e^{\sigma(1+|k|)} \left(1 + |k|^2\right)^{m/2} |\widehat{u}(k)| dk,$$

*for $\sigma \geq 0$ and $m \geq 0$.*

In the following we use

$$\|uv\|_{G_\sigma^m} \leq C_m \|u\|_{\mathcal{W}_\sigma^m} \|v\|_{G_\sigma^m}, \tag{35}$$

for $u \in \mathcal{W}_\sigma^m$ and $v \in G_\sigma^m$, with $m, \sigma \geq 0$, where the constant $C_m > 0$ is independent of $\sigma \geq 0$.

## 4.2 The structure of the problem

Our starting point in this section is the System (4)-(6) on which the transform (11) is applied. The equations for $v$, $w$ and $\psi$ are obtained as above in Remark 1.6, i.e.

$$\begin{aligned}
\partial_t^2(a_1 v + a_2 w) &= \partial_x^2(a_1 v + a_2 w) + (a_3 v + a_4 w)^2 + v \\
&\quad - (\partial_t(a_1 v + a_2 w))^2 - \psi^2 + (\partial_x(a_1 v + a_2 w))^2 \\
&\quad + \gamma e^{2r_{0,\mu}}(e^{2(a_1 v + a_2 w)} - 1 - 2(a_1 v + a_2 w)), && (36)\\
\partial_t w &= 2\psi(\partial_x(a_1 v + a_2 w)) + \partial_x \psi \\
&\quad -2(\partial_t(a_1 v + a_2 w))(a_3 v + a_4 w), && (37)\\
\partial_t \psi &= \partial_x(a_3 v + a_4 w). && (38)
\end{aligned}$$

We switch to the $(X,T)$-coordinates in (36)-(38)

$$\begin{aligned}
\delta^2 \partial_T^2(a_1 \check{v} + a_2 \check{w}) &= \delta^2 \partial_X^2(a_1 \check{v} + a_2 \check{w}) + (a_3 \check{v} + a_4 \check{w})^2 + \check{v} && (39)\\
&\quad -\delta^2 \left(\partial_T(a_1 \check{v} + a_2 \check{w})\right)^2 - \check{\psi}^2 \\
&\quad +\delta^2 \left(\partial_X(a_1 \check{v} + a_2 \check{w})\right)^2 \\
&\quad +\gamma e^{2r_{0,\mu}}(e^{2(a_1 \check{v} + a_2 \check{w})} - 1 - 2(a_1 \check{v} + a_2 \check{w})), \\
\partial_T \check{w} &= 2\check{\psi}(\partial_X(a_1 \check{v} + a_2 \check{w})) + \partial_X \check{\psi} && (40)\\
&\quad -2(\partial_T(a_1 \check{v} + a_2 \check{w}))(a_3 \check{v} + a_4 \check{w}), \\
\partial_T \check{\psi} &= \partial_X(a_3 \check{v} + a_4 \check{w}). && (41)
\end{aligned}$$

The resulting system for $v$, $w$, and $\psi$ is then of the form

$$\begin{aligned}
0 &= \mathbf{M}_v(\mathbf{v},\mathbf{u}) + \delta^2 \mathbf{F}_v(D_X^2\mathbf{v}, D_X^2\mathbf{u}, D_T^2\mathbf{v}, D_T^2\mathbf{u}), \\
\partial_T \mathbf{u} &= \mathbf{M}_u(\mathbf{v},\mathbf{u})\partial_X(\mathbf{v},\mathbf{u}) + \mathbf{M}_{T,u}(\mathbf{v},\mathbf{u})\partial_T(\mathbf{v},\mathbf{u}), && (42)
\end{aligned}$$

with $\mathbf{v} = \check{v}$ and $\mathbf{u} = (\check{w}, \check{\psi})$, where $\mathbf{M}_v(\mathbf{v},\mathbf{u})$, $\mathbf{M}_u(\mathbf{u},\mathbf{v})$ and $\mathbf{M}_{T,u}(\mathbf{u},\mathbf{v})$ are entire (matrix-valued) functions of their arguments. The functions $\mathbf{F}_v$ is polynomial in $D_X^2\mathbf{v} = (\mathbf{v}, \partial_X\mathbf{v}, \partial_X^2\mathbf{v})$, $D_X^2\mathbf{u} = (\mathbf{u}, \partial_X\mathbf{u}, \partial_X^2\mathbf{u})$, $D_T\mathbf{v} = (\mathbf{v}, \partial_T\mathbf{v})$ and $D_T\mathbf{u} = (\mathbf{u}, \partial_T\mathbf{u})$, and the additional property that $\mathbf{F}_v$ is linear in $\partial_T^2\mathbf{v}$ and $\partial_T^2\mathbf{u}$.

In lowest order for $\delta = 0$ we have

$$\begin{aligned}
0 &= \mathbf{M}_v(\mathbf{v},\mathbf{u}), && (43)\\
\partial_T \mathbf{u} &= \mathbf{M}_u(\mathbf{v},\mathbf{u})\partial_X(\mathbf{v},\mathbf{u}) + \mathbf{M}_{T,u}(\mathbf{v},\mathbf{u})\partial_T(\mathbf{v},\mathbf{u}). && (44)
\end{aligned}$$

By construction for $\mathbf{u}$ sufficiently small the equation (43) can be solved by the implicit function theorem with respect to $\mathbf{v} = \mathbf{v}^*(\mathbf{u})$. Inserting this into (44) yields

$$\partial_T \mathbf{u} = \mathbf{M}_u(\mathbf{u})\partial_X \mathbf{u} + \mathbf{M}_{T,u}(\mathbf{u})\partial_T \mathbf{u}. \tag{45}$$

With the help of Neumann's series for $\check{v}$, $\check{w}$ and $\check{\psi}$ sufficiently small, we can solve (45) w.r.t. $\partial_T \mathbf{u}$, i.e., we obtain

$$\partial_T \mathbf{u} = \mathbf{M}(\mathbf{u})\partial_X \mathbf{u}. \tag{46}$$

Our system is now in a form where we find a local existence and uniqueness result by using the Cauchy-Kovalevskaya theory for Gevrey spaces similarly to [HdRS21] and [BKS20].

## 4.3 Cauchy-Kovalevskaya theory in Gevrey spaces

We have the quasilinear abstract Cauchy problem of the form

$$\partial_T \mathbf{u} = \mathbf{M}(\mathbf{u})\partial_X \mathbf{u}, \quad \mathbf{u}\big|_{T=0} = \mathbf{u}_0\,, \qquad X \in \mathbb{R},\, T \geq 0, \tag{47}$$

where $\mathbf{u} = \mathbf{u}(X,T)$ is an unknown function taking values in $\mathbb{R}^d$. The initial condition $\mathbf{u}_0$ lies in the Gevrey space $G^m_{\sigma_0}$ for some $m > 1$ and $\sigma_0 > 0$, and $\mathbf{M}(\mathbf{u})$ is an entire matrix-valued function. The following Cauchy-Kovalevskaya theorem provides local existence and uniqueness of solutions in Gevrey spaces for (47).

**Theorem 4.3.** *Let $m > 1$ and $R, \sigma_0 > 0$. Then, for every $\mathbf{u}_0 \in G^m_{\sigma_0}$ with $2\|\mathbf{u}_0\|_{G^m_{\sigma_0}} < R$ and $\sigma_1 \in (0, \sigma_0)$, there exists an $\eta = \eta(R, m, \sigma_0, \sigma_1) > 0$ such that for $T_0 = (\sigma_0 - \sigma_1)/\eta$ there exists a local solution $\mathbf{u} \in C^1\big((0, T_0], G^{m-1}_{\sigma_1}\big) \cap C\big([0, T_0], G^m_{\sigma_1}\big)$ to (47), satisfying*

$$\sup_{T \in [0,T_0]} \|\mathbf{u}(T)\|_{G^m_{\sigma_1}} \leq R. \tag{48}$$

For a proof we refer to the existing literature, cf. [Saf95, Theorem 1.1]. As preparation for the subsequent error estimates we would like to show for this simple example how to obtain estimates in a time-dependent scale of Gevrey spaces.

Let $|k|_{op} := \sqrt{-\partial_x^2}$. Multiplication of (47) by

$$e^{2\sigma(T)(1+|k|_{op})} \left(1+|k|_{op}^2\right)^m \mathbf{u},$$

where $\sigma(T) = \sigma_0 - \eta T$, and integration w.r.t. $X \in \mathbb{R}$ leads to

$$\begin{aligned}\frac{1}{2}\frac{d}{dT}\|\mathbf{u}\|^2_{G^m_{\sigma(T)}} &+ \eta\|(1+|k|_{op})^{1/2}\mathbf{u}\|^2_{G^m_{\sigma(T)}} \\ &= \operatorname{Re}\Big(((\mathbf{M}(\mathbf{u}) - \mathbf{M}(0))\partial_X \mathbf{u}, \mathbf{u})_{G^m_{\sigma(T)}} + (\mathbf{M}(0)\partial_X \mathbf{u}, \mathbf{u})_{G^m_{\sigma(T)}}\Big).\end{aligned}$$

By the Cauchy-Schwarz like inequality

$$\operatorname{Re}(\mathbf{u}, \mathbf{v})_{G^m_\sigma} \leq \|\mathbf{u}\|_{G^{m-1/2}_\sigma} \|\mathbf{v}\|_{G^{m+1/2}_\sigma}, \tag{49}$$

Remark 4.1 with (32) and (34), and the assumption $m - \frac{1}{2} > \frac{1}{2}$ we have

$$\frac{1}{2}\frac{d}{dT}\|\mathbf{u}\|^2_{G^m_{\sigma(T)}} + \eta\|\mathbf{u}\|^2_{G^{m+1/2}_{\sigma(T)}} \leq \|\mathbf{M}(0)\|\|\mathbf{u}\|^2_{G^{m+1/2}_{\sigma(T)}} + \phi_m(\|\mathbf{u}\|_{G^{m-1/2}_{\sigma(T)}})\|\mathbf{u}\|^2_{G^{m+1/2}_{\sigma(T)}},$$

with $\phi_m$ an entire function which is monotonically increasing on $\mathbb{R}_+$ and satisfies $\phi_m(0) = 0$. Finally, we obtain

$$\frac{1}{2}\frac{d}{dT}\|\mathbf{u}\|^2_{G^m_{\sigma(T)}} + \Big(\eta - \|\mathbf{M}(0)\| - \phi_m(\|\mathbf{u}\|_{G^m_{\sigma(T)}})\Big)\|\mathbf{u}\|^2_{G^{m+1/2}_{\sigma(T)}} \leq 0. \tag{50}$$

Choosing $\eta$ so large that

$$\eta > \|\mathbf{M}(0)\| + \phi_m(R)$$

finally yields (48).

## 4.4 Approximate solutions for the perturbed problem

The residual contains the terms which do not cancel after inserting the approximation into the original system. Adding higher-order terms to the WME approximation (13)-(14) allows us to make the residual sufficiently small for our purposes. Like in [BKS20, HdRS21] we consider an improved approximation $(\mathbf{v}, \mathbf{u})$ of the form

$$\begin{aligned}\mathbf{v}(X,T,\delta) &= \mathbf{v}^0(X,T) + \delta^2\mathbf{v}^1(X,T) + \delta^4\mathbf{v}^2(X,T) + h.o.t., \\ \mathbf{u}(X,T,\delta) &= \mathbf{u}^0(X,T) + \delta^2\mathbf{u}^1(X,T) + \delta^4\mathbf{u}^2(X,T) + h.o.t..\end{aligned}$$

We insert the ansatz into (42) and equate the coefficient in front of the $\delta^s$ to zero. In lowest order, i.e., here at $\delta^0$ and $\delta^1$, we get

$$\begin{aligned} 0 &= \mathbf{M}_v(\mathbf{v}^0, \mathbf{u}^0), \\ \partial_T \mathbf{u}^0 &= \mathbf{M}_u(\mathbf{v}^0, \mathbf{u}^0)\partial_X(\mathbf{v}^0, \mathbf{u}^0) + \mathbf{M}_{T,u}(\mathbf{v}^0, \mathbf{u}^0)\partial_T(\mathbf{v}^0, \mathbf{u}^0). \end{aligned}$$

As above the first equation can be solved w.r.t. to $\mathbf{v}^0$ for $\mathbf{u}^0$ sufficiently small. Inserting the solution $\mathbf{v}^0 = \mathbf{v}^0(\mathbf{u}^0)$ in the second equation and solving then by Neumann's series, for $\mathbf{u^0}$ sufficiently small, the second equation w.r.t. $\partial_T \mathbf{u}^0$ yields

$$\partial_T \mathbf{u}^0 = \mathbf{M}(\mathbf{u}^0)\partial_X(\mathbf{u}^0), \quad \text{with} \ \ \mathbf{u}^0|_{T=0} = \mathbf{u}_0, \tag{51}$$

which coincides with the equation (47) studied earlier. The governing equations for $(\mathbf{v}^n, \mathbf{u}^n), n \in \mathbb{N}$, arise at $\delta^{2n}$ and $\delta^{2n+1}$. We obtain linear inhomogeneous equations of the form

$$\begin{aligned} 0 &= \widetilde{\mathbf{M}}_v(\mathbf{v}^n, \mathbf{v}^0, \mathbf{u}^n, \mathbf{u}^0) \\ &\quad + \mathbf{F}_{v,n}\left(D_X^2 \mathbf{v}^0, D_X^2 \mathbf{u}^0, D_T^2 \mathbf{v}^0, D_T^2 \mathbf{u}^0, ... \right. \\ &\qquad\qquad \left. ..., D_X^2 \mathbf{v}^{n-1}, D_X^2 \mathbf{u}^{n-1}, D_T^2 \mathbf{v}^{n-1}, D_T^2 \mathbf{u}^{n-1}\right), \\ \partial_T \mathbf{u}^n &= \mathbf{M}_u(\mathbf{v}^0, \mathbf{u}^0)\partial_X(\mathbf{v}^n, \mathbf{u}^n) + D\mathbf{M}_u(\mathbf{v}^0, \mathbf{u}^0)[(\mathbf{v}^n, \mathbf{u}^n)]\partial_X(\mathbf{v}^0, \mathbf{u}^0) \\ &\quad + \mathbf{M}_{T,u}(\mathbf{v}^0, \mathbf{u}^0)\partial_T(\mathbf{v}^n, \mathbf{u}^n) + D\mathbf{M}_{T,u}(\mathbf{v}^0, \mathbf{u}^0)[(\mathbf{v}^n, \mathbf{u}^n)]\partial_T(\mathbf{v}^0, \mathbf{u}^0) \\ &\quad + \mathbf{F}_{u,n}(D_X^2 \mathbf{v}^0, D_X^2 \mathbf{u}^0, D_T \mathbf{v}_0, D_T \mathbf{u}_0, ... \\ &\qquad\qquad ..., D_X^2 \mathbf{v}^{n-1}, D_X^2 \mathbf{u}^{n-1}, D_T \mathbf{v}^{n-1}, D_T \mathbf{u}^{n-1}). \end{aligned}$$

Herein, $D\mathbf{M}_u(\mathbf{v}^0, \mathbf{u}^0)[(\mathbf{v}^n, \mathbf{u}^n)]$ denotes the linearization of the map $(v, u) \mapsto \mathbf{M}_u(v, u)$ in the point $(\mathbf{v}^0, \mathbf{u}^0)$ applied to $(\mathbf{v}^n, \mathbf{u}^n)$. $D\mathbf{M}_{T,u}(\mathbf{v}^0, \mathbf{u}^0)[(\mathbf{v}^n, \mathbf{u}^n)]$ is analogously defined. As above we can use the implicit function theorem to solve the first equation with respect to

$$\begin{aligned} \mathbf{v}^n &= \mathbf{v}^n\left(D_X^2 \mathbf{v}^0, D_X \mathbf{u}^0, D_T^2 \mathbf{v}^0, D_T^2 \mathbf{u}^0, ..., D_X^2 \mathbf{v}^{n-1}, D_X \mathbf{u}^{n-1}, \mathbf{u}^n, D_T^2 \mathbf{v}^{n-1}, \right. \\ &\qquad \left. D_T^2 \mathbf{u}^{n-1}\right), \end{aligned}$$

for sufficiently small and sufficiently smooth data, where $\mathbf{v}^n$ is analytic in a neighborhood of 0 and where $\mathbf{v}^n = 0$ for $(D_X^2 \mathbf{v}^0, \ldots, D_T^2 \mathbf{u}^{n-1}) = 0$. By assumption, $\mathbf{v}^{n-1}$ and therefore $D_T^2 \mathbf{v}^{n-1}$ and $D_X^2 \mathbf{v}^{n-1}$ are completely described in terms of $(\mathbf{v}^0, \mathbf{u}^0, ..., \mathbf{v}^{n-2}, \mathbf{u}^{n-2}, \mathbf{u}^{n-1})$ and its temporal and spatial derivatives. Iteratively $(\mathbf{v}^0, ..., \mathbf{v}^{n-1}, \mathbf{v}^n)$ is determined in terms of $(\mathbf{u}^0, ..., \mathbf{u}^{n-1}, \mathbf{u}^n)$.

Thus, we get the equation

$$
\begin{aligned}
\partial_T \mathbf{u}^n &= \widetilde{\mathbf{M}}_u(\mathbf{u}^0)\partial_X \mathbf{u}^n + \widetilde{D\mathbf{M}}_u(\mathbf{u}^0)[\mathbf{u}^n]\partial_X \mathbf{u}^0 \\
&\quad + \widetilde{\mathbf{M}}_{T,u}(\mathbf{u}^0)\partial_T \mathbf{u}^n + \widetilde{D\mathbf{M}}_{T,u}(\mathbf{u}^0)[(\mathbf{u}^n)]\partial_T(\mathbf{u}^0) \\
&\quad + \mathbf{F}_n(D_X^2\mathbf{v}^0, D_X^2\mathbf{u}^0, D_T\mathbf{v}^0, D_T\mathbf{u}^0, \ldots \\
&\qquad \ldots, D_X^2\mathbf{v}^{n-1}, D_X^2\mathbf{u}^{n-1}, D_T\mathbf{v}^{n-1}, D_T\mathbf{u}^{n-1}),
\end{aligned} \tag{52}
$$

with zero inital data for $n \geq 1$. Here, $\widetilde{\mathbf{M}}_u$, $\widetilde{\mathbf{M}}_{T,u}$, $\mathbf{F}_n$ and the linearizations $\widetilde{D\mathbf{M}}_u(\mathbf{u}^0)$ and $\widetilde{D\mathbf{M}}_{T,u}(\mathbf{u}^0)$ are entire (matrix-valued) functions and $\mathbf{F}_n(0) = 0$. Similar to above we apply Neumann's series to solve (52) w.r.t $\partial_T \mathbf{u}^n$ for sufficiently small and smooth inital data. This finally yields

$$
\begin{aligned}
\partial_T \mathbf{u}^n &= \widetilde{\mathbf{M}}_u^*(\mathbf{u}^0)\partial_X \mathbf{u}^n + \widetilde{D\mathbf{M}}_u^*(\mathbf{u}^0)[\mathbf{u}^n]\partial_X \mathbf{u}^0 \\
&\quad + \widetilde{D\mathbf{M}}_{T,u}^*(\mathbf{u}^0)[(\mathbf{u}^n)]\partial_T(\mathbf{u}^0) \\
&\quad + \mathbf{F}_n^*(D_X^2\mathbf{v}^0, D_X^2\mathbf{u}^0, D_T\mathbf{v}^0, D_T\mathbf{u}^0, \ldots \\
&\qquad \ldots, D_X^2\mathbf{v}^{n-1}, D_X^2\mathbf{u}^{n-1}, D_T\mathbf{v}^{n-1}, D_T\mathbf{u}^{n-1}).
\end{aligned} \tag{53}
$$

Line for line as in [HdRS21] we obtain

**Theorem 4.4.** *Let $m > 1$ and $\sigma_0 > 0$. Suppose there exists a local solution*

$$\mathbf{u}^0 \in C^1\big((0, T_0], G_{\sigma_0}^{m-1}\big) \cap C\big([0, T_0], G_{\sigma_0}^m\big),$$

*to* (51)*. Then, for every $\sigma_1 \in (0, \sigma_0)$, $n \in \mathbb{N}$, and for all $0 < k \leq n$, there exist $T_1 = T_1(\sigma_1, k+1) \leq T_1(\sigma_1, k) \leq T_0$ and solutions*

$$\mathbf{u}^k \in C^1\big((0, T_1], G_{\sigma_1}^{m-1}\big) \cap C\big([0, T_1], G_{\sigma_1}^m\big),$$

*to* (53)*.*

Then the n-th order approximations are given by

$$
\begin{aligned}
\widetilde{\mathbf{v}}^n(T) &= \mathbf{v}^0(T) + \delta^2\mathbf{v}^1(T) + \cdots + \delta^{2n}\mathbf{v}^n(T)\,, \\
\widetilde{\mathbf{u}}^n(T) &= \mathbf{u}^0(T) + \delta^2\mathbf{u}^1(T) + \cdots + \delta^{2n}\mathbf{u}^n(T)\,,
\end{aligned}
$$

with corresponding residuals

$$
\begin{aligned}
\mathsf{Res}_v^n(T) &= \mathbf{M}_v(\widetilde{\mathbf{v}}^n, \widetilde{\mathbf{u}}^n) + \delta^2\mathbf{F}_v(D_X^2\widetilde{\mathbf{v}}^n, D_X^2\widetilde{\mathbf{u}}^n, D_T^2\widetilde{\mathbf{v}}^n, D_T^2\widetilde{\mathbf{u}}^n), \\
\mathsf{Res}_u^n(T) &= -\partial_T\widetilde{\mathbf{u}}^n + \mathbf{M}_u(\widetilde{\mathbf{v}}^n, \widetilde{\mathbf{u}}^n)\partial_X(\widetilde{\mathbf{v}}^n, \widetilde{\mathbf{u}}^n) + \mathbf{M}_{T,u}(\widetilde{\mathbf{v}}^n, \widetilde{\mathbf{u}}^n)\partial_T(\widetilde{\mathbf{v}}^n, \widetilde{\mathbf{u}}^n) \\
&\quad + \delta^2\mathbf{F}_u(D_X^2\widetilde{\mathbf{v}}^n, D_X^2\widetilde{\mathbf{u}}^n, D_T\widetilde{\mathbf{v}}^n, D_T\widetilde{\mathbf{u}}^n).
\end{aligned}
$$

By the above construction and Theorem 4.4 we directly obtain

**Corollary 4.5.** *Assume that the hypotheses of Theorem 4.4 are met. Then, for every* $n \in \mathbb{N}$*, the approximate solutions* $(\widetilde{\mathbf{v}}^n, \widetilde{\mathbf{u}}^n)$ *and residuals* $(\mathsf{Res}_v^n, \mathsf{Res}_u^n)$ *are in* $C\big([0,T_1], G^m_{\tilde\sigma_1}\big)$ *for all* $\tilde\sigma_1 \in [0,\sigma_1)$*. Further, there exists a constant* $C > 0$ *such that we have*

$$\sup_{T\in[0,T_1]} \|(\widetilde{\mathbf{v}}^n(T), \widetilde{\mathbf{u}}^n(T)) - (\mathbf{v}^0(T), \mathbf{u}^0(T))\|_{G^m_{\tilde\sigma_1}} \le C\delta^2,$$

$$\sup_{T\in[0,T_1]} \|(\mathsf{Res}_v^n, \mathsf{Res}_u^n)(T)\|_{G^m_{\tilde\sigma_1}} \le C\delta^{2n+2}.$$

# 5 The error equations

For notational simplicity the following analysis is carried out for $q = 0$. It will be obvious that the proof will also work in the parameter regimes $\mathcal{P}_{\text{benj}}$, and $\mathcal{P}_{\text{stab}}$, cf. Figure 4. For estimating the difference between the WME approximation and true solutions of (4)-(6) on the long $\mathcal{O}(1/\delta)$-timescale, we separate the modes in a neighborhood of the wave number $k = 0$ from the the modes bounded away from the wave number $k = 0$. In the neighborhood of $k = 0$ we use normal form transformations and energy estimates similar to the ones in Section 4 to get rid of the terms of order $\mathcal{O}(1)$ in the equations for the error. Outside this neighborhood we use the artificial damping obtained from the time-dependent scale of Gevrey spaces.

Our starting point is system (4)-(6) which we write as first order system

$$\partial_t V = LV + N(V), \tag{54}$$

where

$$V = \begin{pmatrix} r \\ \widetilde{r} \\ \vartheta \\ \psi \end{pmatrix}, \qquad LV = \begin{pmatrix} \widetilde{r} \\ \partial_x^2 r + 2\mu\vartheta + 2(\mu^2-1)r \\ \partial_x\psi - 2\mu\widetilde{r} \\ \partial_x\vartheta \end{pmatrix},$$

and

$$N(V) = \begin{pmatrix} 0 \\ \vartheta^2 - \widetilde{r}^2 - \psi^2 + (\partial_x r)^2 + (1-\mu^2)(e^{2r} - 1 - 2r) \\ 2(\partial_x r)\psi - 2\widetilde{r}\vartheta \\ 0 \end{pmatrix}.$$

We introduce the error function $\mathcal{R}$ made by the associated WME approximation $\Psi$ through $V = \Psi + \delta^{3/2}\mathcal{R}$. The error function $\mathcal{R}$ satisfies

$$\partial_t \mathcal{R} = L\mathcal{R} + B(\Psi, \mathcal{R}) + \delta^{3/2} G(\mathcal{R}) + \delta^{-3/2}\text{Res}_{\mathcal{R}},$$

where $B(\Psi,\mathcal{R})$ stand for the $\Psi$-dependent terms which are linear in $\mathcal{R}$, and $G(\mathcal{R})$ for the terms which are non-linear in $\mathcal{R}$, i.e.,

$$B(\Psi,\mathcal{R}) = DN(\Psi)\mathcal{R}, \qquad \delta^{3/2}G(\mathcal{R}) = N(\Psi+\delta^{3/2}\mathcal{R}) - N(\Psi) - DN(\Psi)\mathcal{R}.$$

In our notation we suppress the fact $G$ depends on $\Psi$, too. For the separation of the modes we introduce the mode projections

$$\widehat{E_{\delta_c}}(k) = \begin{cases} 1, & |k| \leq \delta_c, \\ 0, & |k| > \delta_c, \end{cases}$$

and $\widehat{E^c_{\delta_c}}(k) = 1 - \widehat{E_{\delta_c}}(k)$ for $\delta_c > 0$ independent of $0 < \delta \ll 1$.

We split $\mathcal{R}$ with the help of $E_{\delta_c}$, i.e., let $R = E_{\delta_c}\mathcal{R}$ and $R^c = E^c_{\delta_c}\mathcal{R}$. The new error functions $R$ and $R^c$ satisfy

$$\begin{aligned} \partial_t R &= LR + E_{\delta_c}B(\Psi,R) + E_{\delta_c}B(\Psi,R^c) \qquad (55)\\ &\quad + \delta^{3/2}E_{\delta_c}G(R+R^c) + \delta^{-3/2}E_{\delta_c}\mathrm{Res}_{\mathcal{R}}, \\ \partial_t R^c &= LR^c + E^c_{\delta_c}B(\Psi,R) + E^c_{\delta_c}B(\Psi,R^c) \qquad (56)\\ &\quad + \delta^{3/2}E^c_{\delta_c}G(R+R^c) + \delta^{-3/2}E^c_{\delta_c}\mathrm{Res}_{\mathcal{R}}. \end{aligned}$$

For controlling the error functions $R$ and $R^c$ on the long $\mathcal{O}(1/\delta)$-time scale we have to get rid of the $\mathcal{O}(1)$-terms on the right-hand side. The terms

$$E^c_{\delta_c}B(\Psi,R) + E^c_{\delta_c}B(\Psi,R^c)$$

in (56) can be controlled with the artificial damping obtained from the time-dependent scale of Gevrey spaces. We use normal form transformations and energy estimates to get rid of the terms $E_{\delta_c}B(\Psi,R) + E_{\delta_c}B(\Psi,R^c)$ in (55).

For the normal form transformations in a neighborhood around $k = 0$ it is advantageous to diagonalize the linearized system in this neighborhood. A diagonalization is possible, since $L$ is of the form

$$L(k) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 2(\mu^2-1)+\mathcal{O}(k^2) & 0 & 2\mu & 0 \\ 0 & -2\mu & 0 & ik \\ 0 & 0 & ik & 0 \end{pmatrix}$$

for $k \to 0$. There are two eigenvalues $\omega_{\pm 1}$ of order $\mathcal{O}(1)$ and two eigenvalues $\omega_{\pm 2}$ of order $\mathcal{O}(k)$. The associated eigenvectors are of the form

$$\varphi_{\pm 1}(0) = \begin{pmatrix} \mathcal{O}(1) \\ \mathcal{O}(1) \\ \mathcal{O}(1) \\ 0 \end{pmatrix}, \qquad \varphi_2(0) = \begin{pmatrix} \mathcal{O}(1) \\ 0 \\ \mathcal{O}(1) \\ 0 \end{pmatrix}, \qquad \varphi_{-2}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathcal{O}(1) \end{pmatrix}.$$

For $R^* = S^{-1}R$, with the matrix $S(k) = (\varphi_1(k), \varphi_{-1}(k), \varphi_2(k), \varphi_{-2}(k))$ for $|k| \leq \delta_c$, we find

$$\begin{aligned} \partial_t R^* &= \Lambda R^* + E_{\delta_c} S^{-1} B(\Psi, SR^*) + E_{\delta_c} S^{-1} B(\Psi, R^c) \\ &\quad + \delta^{3/2} E_{\delta_c} S^{-1} G(SR^* + R^c) + \delta^{-3/2} E_{\delta_c} S^{-1} \mathrm{Res}_{\mathcal{R}}, \end{aligned} \tag{57}$$

where

$$\Lambda(k) = S^{-1}(k)L(k)S(k) = \begin{pmatrix} i\omega_1(k) & 0 & 0 & \\ 0 & i\omega_{-1}(k) & 0 & 0 \\ 0 & 0 & i\omega_2(k) & 0 \\ 0 & 0 & 0 & i\omega_{-2}(k) \end{pmatrix}.$$

**Remark 5.1.** Since $\Psi$ is strongly concentrated at the wave number $k = 0$ the part $E^c_{\delta_c}\Psi$ is $\mathcal{O}(\delta^s)$ in the spaces used subsequently if $\Psi$ is chosen $s$ times more differentiable than the error. The approximation $\Psi$ appears in the equations for the error not only linearly but also nonlinearly but due to (32) also $E^c_{\delta_c}$ applied on the nonlinear terms in $\Psi$ is $\mathcal{O}(\delta^s)$ if $\Psi$ is chosen $s$ times more differentiable than the error. Hence, if we separate $\Psi$ in $E_{\delta_c}\Psi$ and $E^c_{\delta_c}\Psi$, a product of these terms is already $\mathcal{O}(\delta^s)$ if at least one $E^c_{\delta_c}\Psi$-factor appears. Hence, terms with no $E^c_{\delta_c}\Psi$-factor are of order $\mathcal{O}(1)$. Again due to the concentration of $\Psi$ at the wave number $k = 0$ if $E^c_{\delta_c}$ is applied to such a term, i.e., to an entire function of $E_{\delta_c}\Psi$, we again have an $\mathcal{O}(\delta^s)$-order. Therefore, we can restrict the support of the $\mathcal{O}(1)$-terms to the support of $E_{\delta_c}$. In our notation these $\mathcal{O}(1)$-terms will be denoted by $\Psi_0$. The $\mathcal{O}(\delta^s)$-terms will be denoted by $\Psi_r$.

**Example 5.2.** Let $\psi_1$ be the first component of $\Psi$. Then $E_{\delta_c}(E_{\delta_c}\psi_1 \cdot E_{\delta_c}\psi_1)$ belongs to $\Psi_0$. The terms $E^c_{\delta_c}(E_{\delta_c}\psi_1 \cdot E_{\delta_c}\psi_1)$, $E^c_{\delta_c}\psi_1 \cdot E_{\delta_c}\psi_1$, $E_{\delta_c}\psi_1 \cdot E^c_{\delta_c}\psi_1$, and $E^c_{\delta_c}\psi_1 \cdot E^c_{\delta_c}\psi_1$ are at least of order $\mathcal{O}(\delta^s)$ and belong to $\Psi_r$.

We use Remark 5.1 to extract higher order terms from $E_{\delta_c} S^{-1} B(\Psi, SR^*) + E_{\delta_c} S^{-1} B(\Psi, R^c)$ such that only a number of lower order terms with a bounded Fourier support have to be eliminated by normal form transformations. Hence we rewrite (57) into

$$\partial_t R^* = \Lambda R^* + E_{\delta_c} S^{-1} B(\Psi_0, SR^*) + E_{\delta_c} S^{-1} B(\Psi_0, R^c) + H^* \tag{58}$$

where

$$\begin{aligned} H^* &= E_{\delta_c} S^{-1} B(\Psi_r, SR^*) + E_{\delta_c} S^{-1} B(\Psi_r, R^c) \\ &\quad + \delta^{3/2} E_{\delta_c} S^{-1} G(SR^* + R^c) + \delta^{-3/2} E_{\delta_c} S^{-1} \mathrm{Res}_{\mathcal{R}}. \end{aligned}$$

**Remark 5.3.** As continuation of Remark 5.1 we remind that in physical space $E_{\delta_c}S^{-1}B(\Psi_0, SR^*) + E_{\delta_c}S^{-1}B(\Psi_0, R^c)$ mainly consists of products of $E_{\delta_c}\Psi_0$ multiplied with one $R^*$, which has been rearranged in such a way that $E_{\delta_c}$ is applied on the product of the $E_{\delta_c}\Psi_0$ and $E^c_{\delta_c}$ applied on this product has been moved to $H^*$.

**Example 5.4.** Let $\psi_1$ resp. $R_1$ be the first component of $\Psi$ resp. $R$. Then $E_{\delta_c}\psi_1 \cdot E_{\delta_c}\psi_1 \cdot R_1$ is written as $E_{\delta_c}(E_{\delta_c}\psi_1 \cdot E_{\delta_c}\psi_1) \cdot R_1$ which is a part of $E_{\delta_c}S^{-1}B(\Psi_0, SR^*)$ and $E^c_{\delta_c}(E_{\delta_c}\psi_1 \cdot E_{\delta_c}\psi_1) \cdot R_1$ which is a part of $H^*$.

**Remark 5.5.** The terms in $E_{\delta_c}S^{-1}B(\Psi_0, SR^*) + E_{\delta_c}S^{-1}B(\Psi_0, R^c)$ are a mode projection applied on a product of a term with Fourier support in $[-\delta_c, \delta_c]$ with an error term. In Fourier space the term is mainly a convolution with some kernel $K$, i.e.,

$$\widehat{E_{\delta_c}}(k) \int K(k, k-m, m)\Psi_0(k-m)R(m)dm.$$

This term is non-zero if $|k| \leq \delta_c$ and $|k-m| \leq \delta_c$. As a consequence only $|m| \leq 2\delta_c$ has to be considered. Hence w.l.o.g. we write the term $E_{\delta_c}S^{-1}B(\Psi_0, SR^*) + E_{\delta_c}S^{-1}B(\Psi_0, R^c)$ as $E_{\delta_c}S^{-1}B(\Psi_0, SR)$ which can always be achieved if $\delta_c > 0$ is chosen sufficiently small.

With the previous remarks we write (58) as

$$\partial_t R^* = \Lambda R^* + E_{\delta_c}S^{-1}B(\Psi_0, SR^*) + H^* \tag{59}$$

and use normal form transformations to simplify $E_{\delta_c}S^{-1}B(\Psi_0, SR)$ as far as possible for applying subsequently energy estimates to get rid of the remaining terms. The term $H^*$ does not make problems to prove bounds for $R$ on the long $\mathcal{O}(1/\delta)$-time scale since all terms in $H^*$ are at least of order $\mathcal{O}(\delta)$.

We separate (59) in its components. With $R^*$, respectively $R$, written as $(R_1, R_{-1}, R_2, R_{-2})$, with some slight abuse of notation, we have

$$\partial_t R_1 = i\omega_1 R_1 + \sum_{j=\pm1,\pm2} B_{1,j}(\Psi_0, R_j) + H_1^*, \tag{60}$$

$$\partial_t R_2 = i\omega_2 R_2 + \sum_{j=\pm1,\pm2} B_{2,j}(\Psi_0, R_j) + H_2^*, \tag{61}$$

and similarly for $R_{-1}$ und $R_{-2}$, where

$$B_{j_1,j_2}(\Psi_0, R_{j_2})(k) = \varphi^*_{j_1}(k) \cdot E_{\delta_c}S^{-1}B(\Psi_0, (\varphi^*_{j_2} \cdot SR^*)\varphi_{j_2})(k),$$

and where the $\varphi^*_j(k)$ are the adjoint eigenvectors of $L(k)$, with the property $(\varphi^*_i(k), \varphi_j(k))_{\mathbb{C}^4} = \delta_{ij}$.

# 6 The series of normal form transformations

Our goal is to prove an $\mathcal{O}(1)$-bound for $R_{\pm 1}$, $R_{\pm 2}$, and $R^c$ on an $\mathcal{O}(1/\delta)$-time scale. As already said, as next step on this path, we simplify (60)-(61) by eliminating all non-resonant terms of order $\mathcal{O}(1)$ by near-identity changes of variables. System (60)-(61) has a similiar structure as [DKS16, System (21)-(24)] and so it can be expected that the non-resonant terms $B_{1,\pm 2}$ and $B_{2,\pm 1}$ can be eliminated with a convergent infinite series of normal form transformations.

To illustrate the procedure, we show how to obtain the first change of variables close to the identity. We set

$$\begin{aligned} R_{1,1} &= R_1 + \sum_{j=\pm 2} M_{1,j}(\Psi_0, R_j), \\ R_{2,1} &= R_2 + \sum_{j=\pm 1} M_{2,j}(\Psi_0, R_j). \end{aligned}$$

The operators $M_{i,j}$ and $B_{i,j}$ are linear in the error functions $R_j$ and possess a convolution structure, i.e.,

$$M_{i,j}(\Psi_0, R_j)(k) = \int m_{i,j}(k, k-m, m)\Psi_0(k-m)R_j(m)\mathrm{d}m,$$

with kernel $m_{i,j}$ and similarly for $B_{i,j}$ with kernel $b_{i,j}$.

We differentiate $R_{1,1}$ w.r.t. time and obtain

$$\begin{aligned} \partial_t R_{1,1} &= \partial_t R_1 + \sum_{j=\pm 2} (M_{1,j}(\partial_t \Psi_0, R_j) + M_{1,j}(\Psi_0, \partial_t R_j)) \\ &= i\omega_1 R_1 + \sum_{j=\pm 1,\pm 2} B_{1,j}(\Psi_0, R_j) + \mathcal{O}(\delta) \\ &\quad + \sum_{j=\pm 2} M_{1,j}\left(\Psi_0, i\omega_j R_j + \sum_{j_1=\pm 1,\pm 2} B_{j,j_1}(\Psi_0, R_{j_1})\right) + \mathcal{O}(\delta) \\ &= i\omega_1 R_{1,1} - i\omega_1 \sum_{j=\pm 2} M_{1,j}(\Psi_0, R_j) + \sum_{j=\pm 1,\pm 2} B_{1,j}(\Psi_0, R_j) + \mathcal{O}(\delta) \\ &\quad + \sum_{j=\pm 2} M_{1,j}\left(\Psi_0, i\omega_j R_j + \sum_{j_1=\pm 1,\pm 2} B_{j,j_1}(\Psi_0, R_{j_1})\right) + \mathcal{O}(\delta) \end{aligned}$$

where we used that $\partial_t \Psi_0 = \mathcal{O}(\delta)$ due to the long wave character of $\Psi_0$. In order to eliminate the terms $B_{1,j}$ for $j = \pm 2$ we choose $M_{1,j}$ to satisfy

$$\begin{aligned} -i\omega_1 M_{1,2}(\Psi_0, R_2) + M_{1,2}(\Psi_0, i\omega_2 R_2) + B_{1,2}(\Psi_0, R_2) &= 0, \\ -i\omega_1 M_{1,-2}(\Psi_0, R_{-2}) + M_{1,-2}(\Psi_0, i\omega_{-2} R_{-2}) + B_{1,-2}(\Psi_0, R_{-2}) &= 0, \end{aligned}$$

i.e., we set

$$\begin{aligned} m_{1,2}(k, k-m, m) &= \frac{b_{1,2}(k, k-m, m)}{i\omega_1(k) - i\omega_2(m)}, \\ m_{1,-2}(k, k-m, m) &= \frac{b_{1,-2}(k, k-m, m)}{i\omega_1(k) - i\omega_{-2}(m)}. \end{aligned}$$

Since $|k| \leq \delta_c$ and $|m| \leq 2\delta_c$ the denominator is non-zero for $\delta_c > 0$ sufficiently small, and so $m_{1,\pm 2}$ is well-defined and bounded. As a consequence the $M_{1,\pm 2}$ are bounded mappings in all $G_\sigma^m$-spaces. After this transformation we thus have

$$\begin{aligned} \partial_t R_{1,1} &= i\omega_1 R_{1,1} + \sum_{j=\pm 1} B_{1,j}(\Psi_0, R_j) + \mathcal{O}(\delta) \\ &\quad + \sum_{j=\pm 2} M_{1,j}\left(\Psi_0, \sum_{j_1=\pm 1,\pm 2} B_{j,j_1}(\Psi_0, R_{j_1})\right) + \mathcal{O}(\delta). \end{aligned}$$

We do exactly the same with $R_{2,1}$ and obtain

$$\begin{aligned} m_{2,1}(k, k-m, m) &= \frac{b_{2,1}(k, k-m, m)}{i\omega_2(k) - i\omega_1(m)}, \\ m_{2,-1}(k, k-m, m) &= \frac{b_{2,-1}(k, k-m, m)}{i\omega_2(k) - i\omega_{-1}(m)}, \end{aligned}$$

such that finally

$$\begin{aligned} \partial_t R_{2,1} &= i\omega_2 R_{2,1} + \sum_{j=\pm 2} B_{2,j}(\Psi_0, R_j) + \mathcal{O}(\delta) \\ &\quad + \sum_{j=\pm 1} M_{2,j}\left(\Psi_0, \sum_{j_1=\pm 1,\pm 2} B_{j,j_1}(\Psi_0, R_{j_1})\right) + \mathcal{O}(\delta). \end{aligned}$$

Since the bilinear functions $B$ are of order $\mathcal{O}(1)$ the norm of the normal form transformations $M$ is bounded by the norm of $\Psi$ and $R$. This yields the invertibility of the near identity change of variables with the help of Neumann series for $\Psi = \mathcal{O}(1)$, but sufficiently small.

However, new terms of order $\mathcal{O}(1)$ are created by this strategy and so this procedure must be performed again and again. Convergence of these finally infinitely many transformations holds since only the first resonant terms are of order $\mathcal{O}(\|\Psi\|)$. The newly created terms by the second normal form transformation are at most of order $\mathcal{O}(\|\Psi\|^2)$, then $\mathcal{O}(\|\Psi\|^3)$ by the third transformation, etc., such that a geometric series in $\|\Psi\|$ can be used as convergent majorant for $\|\Psi\|$ sufficiently small. Since the construction is very similar to [DKS16] we refrain from repeating the complete analysis and refer to this paper for more details. We only mention that the convergence of the $\Psi$-terms is controlled by the norm

$$\|f\|_{X^{\sigma,m,\delta}} = \int \sup_{k\in\mathbb{R}} |f(k,l)| \left(1 + \left(\frac{l}{\delta}\right)^2\right)^{\frac{m}{2}} \exp\left(\sigma\left(\frac{l}{\delta}\right)\right) dl.$$

See Appendix A.2 for more details about this point. There is one technical difference between [DKS16] and the present approach. Before performing the next transformation we have to prepare our system according to Remark 5.3 and Example 5.4 such that we have a sequence of problems

$$\partial_t R_{1,n} = i\omega_1 R_{1,n} + \sum_{j=\pm1,\pm2} B_{1,j,n}(\Psi_{0,n}, R_{j,n}) + H^*_{1,n}, \tag{62}$$

$$\partial_t R_{2,n} = i\omega_2 R_{2,n} + \sum_{j=\pm1,\pm2} B_{2,j,n}(\Psi_{0,n}, R_{j,n}) + H^*_{2,n}, \tag{63}$$

and a sequence of normal form transformations

$$R_{1,n+1} = R_{1,n} + \sum_{j=\pm2} M_{1,j,n}(\Psi_{0,n}, R_{j,n}),$$
$$R_{2,n+1} = R_{2,n} + \sum_{j=\pm1} M_{2,j,n}(\Psi_{0,n}, R_{j,n}).$$

The terms which move to $H^*_{j,n}$ when transforming $\Psi_{0,n-1}$ into $\Psi_{0,n}$, according to Remark 5.3 and Example 5.4, are of order $\mathcal{O}(\|\Psi\|^n)$ such that convergence holds for these terms, too.

The limit system has the following structure

$$\partial_t R_{1,\infty} = i\omega_1 R_{1,\infty} + \sum_{j=\pm 1} B_{1,j,\infty}(\Psi_{0,\infty}, R_{j,\infty}) + H^*_{1,\infty}, \tag{64}$$

$$\partial_t R_{2,\infty} = i\omega_2 R_{2,\infty} + \sum_{j=\pm 2} B_{2,j,\infty}(\Psi_{0,\infty}, R_{j,\infty}) + H^*_{2,\infty}, \tag{65}$$

and similar for $R_{-1,\infty}$ and $R_{-2,\infty}$. For $j \in \{\pm 1, \pm 2\}$ the nonlinear terms obey the estimates

$$\|H^*_{j,\infty}\|_{G^m_\sigma} \leq C\delta \left(\|(R_{1,\infty}, R_{2,\infty})\|_{G^m_\sigma} + \delta^{1/2}(\|(R_{1,\infty}, R_{2,\infty})\|^2_{G^m_\sigma} + \|R^c\|^2_{G^m_\sigma}) + 1\right).$$

# 7 Some further preparations

Before performing the energy estimates for obtaining an $\mathcal{O}(1)$-bound for $R_{\pm 1,\infty}$, $R_{\pm 2,\infty}$, and $R^c$ on an $\mathcal{O}(1/\delta)$-time scale we need two additional preparations.

**Remark 7.1.** For notational simplicity it turns out to be advantageous if all components of $R_{\cdot,\infty}$ and $R^c$ have the same regularity. This is automatically fulfilled for $R_{\pm 1,\infty}$ and $R_{\pm 2,\infty}$ since they all have a compact support in Fourier space. However, for $R^c$ this is not the case and so we introduce the multiplication operator $\mathcal{M}$ defined by its symbol $\widehat{\mathcal{M}}(k) = (1+k^2)^{1/2}$ in Fourier space. Since in Equation (61) we have $\widetilde{r} \in G^m_\sigma$, $\vartheta \in G^m_\sigma$, and $\psi \in G^m_\sigma$, but $r \in G^{m+1}_\sigma$, we introduce $\mathbf{r} = \mathcal{M} r \in G^m_\sigma$ and find

$$\partial_t \mathbf{V} = \mathbf{L}\mathbf{V} + \mathbf{N}(\mathbf{V}), \tag{66}$$

where

$$\mathbf{V} = \begin{pmatrix} \mathbf{r} \\ \widetilde{r} \\ \vartheta \\ \psi \end{pmatrix}, \qquad \mathbf{L}\mathbf{V} = \begin{pmatrix} \mathcal{M}\widetilde{r} \\ \partial_x^2(\mathcal{M}^{-1}\mathbf{r}) + 2\mu\vartheta + 2(\mu^2-1)(\mathcal{M}^{-1}\mathbf{r}) \\ \partial_x\psi - 2\mu\widetilde{r} \\ \partial_x\vartheta \end{pmatrix},$$

and

$$\mathbf{N}(\mathbf{V}) = \begin{pmatrix} 0 \\ \vartheta^2 - \widetilde{r}^2 - \psi^2 + \left(\partial_x(\mathcal{M}^{-1}\mathbf{r})\right)^2 + (1-\mu^2)(e^{2(\mathcal{M}^{-1}\mathbf{r})} - 1 - 2(\mathcal{M}^{-1}\mathbf{r})) \\ 2(\partial_x(\mathcal{M}^{-1}\mathbf{r}))\psi - 2\widetilde{r}\vartheta \\ 0 \end{pmatrix}.$$

For the transformed error part $\mathbf{R}^c$ we obtain a system of the form

$$\partial_t \mathbf{R}^c = \mathbf{L}\mathbf{R}^c + H_c^*, \tag{67}$$

with

$$\|H_c^*\|_{G_\sigma^m} \leq C\|\mathbf{R}^c\|_{G_\sigma^m} + \delta(\|(R_{1,\infty}, R_{2,\infty})\|_{G_\sigma^m} + \delta^{1/2}\|\mathbf{R}^c\|_{G_\sigma^m}^2 + 1).$$

In this estimate the terms linear in $R_{\pm 1,\infty}$ and $R_{\pm 2,\infty}$ are at least of order $\mathcal{O}(\delta)$ since these terms coming from the procedure described in Remark 5.1.

**Remark 7.2.** In order to apply the ideas from Section 4 in the energy estimates of $R_{j,\infty}$ we need an additional structure in the limit system (64)-(65). Comparing the derivation of (65) with the derivation of (13)-(14) in combination with (15) we see that the terms $\sum_{j=\pm 2} B_{2,j,\infty}(\Psi_{0,\infty}, R_{j,\infty})$ are the linearization of (13)-(14) around $\Psi_{0,\infty}$. Therefore, they can be written as

$$\sum_{j=\pm 2} B_{2,j,\infty}^*(\Psi_{0,\infty})\partial_x R_{j,\infty} + \mathcal{O}(\delta).$$

The terms where a derivative of (15) falls on $\Psi_{0,\infty}$ gives an additional $\mathcal{O}(\delta)$ such that such a term will be included in $H_{2,\infty}^*$.

# 8 Error estimates in Gevrey spaces

Now we come back to the full system consisting of (64)-(65) and (67). In order to get rid of the remaining $\mathcal{O}(1)$-terms in these equations which are problematic to come to an $\mathcal{O}(1/\delta)$-scale we use as in Section 4 a time-dependent scale of Gevrey spaces to handle the $R_{2,\infty}$-variable exploiting the property from Remark 7.2, use classical energy estimates to handle the $R_{1,\infty}$-variable, and use the artificial damping coming from the time-dependent scale of Gevrey spaces to get rid of the $\mathbf{R}^c$ variable.

Technically, we multiply the equation of $R_{j,\infty}$ with

$$e^{2\sigma(t)|k|_{op}}(1+|k|_{op}^2)^m R_{j,\infty},$$

and the equation of $\mathbf{R}^c$ with

$$e^{2\sigma(t)|k|_{op}}(1+|k|_{op}^2)^m \mathbf{R}^c,$$

where $\sigma(t) = \sigma_0/\delta - \eta t$ and integrate then w.r.t. $x$ where as before $|k|_{op} = \sqrt{-\partial_x^2}$. For

$$E(t) = \|R_{1,\infty}\|^2_{G^m_{\sigma(t)}} + \|R_{2,\infty}\|^2_{G^m_{\sigma(t)}} + \|\mathbf{R}^c\|^2_{G^m_{\sigma(t)}}$$

we find

$$\frac{1}{2}\frac{d}{dt}E = \mathrm{Re}\sum_{j=1}^{11} s_j,$$

where

$$\begin{aligned}
s_1 &= -\eta\| |k|^{1/2}_{op} R_{1,\infty}\|^2_{G^m_{\sigma(t)}}, \\
s_2 &= (R_{1,\infty}, i\omega_1 R_{1,\infty})_{G^m_{\sigma(t)}}, \\
s_3 &= \left(R_{1,\infty}, \sum_{j=\pm 1} B_{1,j,\infty}(\Psi_{0,\infty}, R_{j,\infty})\right)_{G^m_{\sigma(t)}}, \\
s_4 &= (R_{1,\infty}, H^*_{1,\infty})_{G^m_{\sigma(t)}}, \\
s_5 &= -\eta\| |k|^{1/2}_{op} R_{2,\infty}\|^2_{G^m_{\sigma(t)}}, \\
s_6 &= (R_{2,\infty}, i\omega_2 R_{2,\infty})_{G^m_{\sigma(t)}}, \\
s_7 &= \left(R_{2,\infty}, \sum_{j=\pm 2} B_{2,j,\infty}(\Psi_{0,\infty}, R_{j,\infty})\right)_{G^m_{\sigma(t)}}, \\
s_8 &= (R_{2,\infty}, H^*_{2,\infty})_{G^m_{\sigma(t)}}, \\
s_9 &= -\eta\| |k|^{1/2}_{op} \mathbf{R}^c\|^2_{G^m_{\sigma(t)}}, \\
s_{10} &= (\mathbf{R}^c, \mathbf{L}\mathbf{R}^c)_{G^m_{\sigma(t)}}, \\
s_{11} &= (\mathbf{R}^c, H^*_c)_{G^m_{\sigma(t)}}.
\end{aligned}$$

In the following we estimate the 'bad' terms $s_2$, $s_3$, $s_4$, $s_6$, $s_7$, $s_8$, $s_{10}$, and $s_{11}$ by the 'good' artificial damping terms $s_1$, $s_5$, and $s_9$. Before we do so, we explain where the good terms come from.

$\mathbf{s_1}$, $\mathbf{s_5}$, $\mathbf{s_9}$: We have for instance

$$\begin{aligned}
&\int \partial_t R_{1,\infty} e^{2\sigma(t)|k|_{op}} (1+|k|_{op}^2)^m R_{1,\infty} \mathrm{d}x \\
=&\frac{1}{2}\partial_t \int e^{2\sigma(t)|k|_{op}} (1+|k|_{op}^2)^m R_{1,\infty}^2 \mathrm{d}x \\
&+\eta \int e^{2\sigma(t)|k|_{op}} (1+|k|_{op}^2)^m |k|_{op} R_{1,\infty}^2 \mathrm{d}x \\
=&\frac{1}{2}\partial_t \|R_{1,\infty}\|^2_{G^m_{\sigma(t)}} + \eta \||k|_{op}^{1/2} R_{1,\infty}\|^2_{G^m_{\sigma(t)}}.
\end{aligned}$$

$\mathbf{s_2}$: Since $i\omega_1$ is a skew-symmetric operator in the parameter regimes under consideration, we have

$$s_2 = 0.$$

$\mathbf{s_6}$: In the Benjamin-Feir stable situation, cf. left panel of Figure 2, $i\omega_2$ is a skew-symmetric operator which yields

$$s_6 = 0.$$

In the Benjamin-Feir unstable situation, cf. right panel of Figure 2, $i\omega_2$ grows at most as $C|k|$ such that

$$|s_6| \leq C_6 \||k|_{op}^{1/2} R_{2,\infty}\|^2_{G^m_{\sigma(t)}}.$$

Next we go on with the higher order terms.

$\mathbf{s_4}$, $\mathbf{s_8}$, $\mathbf{s_{11}}$: We find

$$\begin{aligned}
|s_4| &\leq C\|R_{1,\infty}\|_{G^m_\sigma} \|H^*_{1,\infty}\|_{G^m_{\sigma(t)}} \\
&\leq C_4 \delta \left( \|(R_{1,\infty}, R_{2,\infty})\|^2_{G^m_{\sigma(t)}} + \delta^{1/2} (\|(R_{1,\infty}, R_{2,\infty})\|^3_{G^m_{\sigma(t)}} + \|\mathbf{R}^c\|^3_{G^m_{\sigma(t)}}) + 1 \right),
\end{aligned}$$

where we used $a \leq 1 + a^2$. Similarly, we obtain

$$\begin{aligned}
|s_8| &\leq C\|R_{2,\infty}\|_{G^m_{\sigma(t)}} \|H^*_{2,\infty}\|_{G^m_{\sigma(t)}} \\
&\leq C_8 \delta \left( \|(R_{1,\infty}, R_{2,\infty})\|^2_{G^m_{\sigma(t)}} + \delta^{1/2} (\|(R_{1,\infty}, R_{2,\infty})\|^3_{G^m_{\sigma(t)}} + \|\mathbf{R}^c\|^3_{G^m_{\sigma(t)}}) + 1 \right).
\end{aligned}$$

Finally, we estimate

$$\begin{aligned}
|s_{11}| &\leq C\|\mathbf{R}^c\|_{G^m_{\sigma(t)}} \|H^*_c\|_{G^m_{\sigma(t)}} \\
&\leq C_{11} (\|\mathbf{R}^c\|^2_{G^m_{\sigma(t)}} + \delta (\|(R_{1,\infty}, R_{2,\infty})\|_{G^m_{\sigma(t)}} \|\mathbf{R}^c\|_{G^m_{\sigma(t)}} + \delta^{1/2} \|\mathbf{R}^c\|^3_{G^m_{\sigma(t)}} + 1)).
\end{aligned}$$

It remains to estimate $s_3$, $s_7$, and $s_{10}$.

$\mathbf{s_{10}}$: We start with the energy estimates for the linear term $\mathbf{LR}^c$. The parameter regions we are working in, include the possibility of Benjamin-Feir unstable wave trains. Hence, the eigenvalues of $\mathbf{L}$ can be bounded from above by $C|k|$, cf. the right panel of Figure 2, and so we obtain the rather rough estimate

$$|s_{10}| \leq C_{10} \| |k|_{op}^{1/2} \mathbf{R}^c \|^2_{G^m_{\sigma(t)}}.$$

$\mathbf{s_7}$: With Remark 7.2 the term $s_7$ can be rewritten as

$$s_7 = (R_{2,\infty}, \sum_{j=\pm 2} B^*_{2,j,\infty}(\Psi_{0,\infty}) \partial_x R_{j,\infty} + \mathcal{O}(\delta))_{G^m_{\sigma(t)}}.$$

Since this term was constructed by infinitely many near identity changes we use the convergence in the $X^{m,\delta}$-spaces. More details about these estimates can be found in the Appendix A.2. Integration by parts with $|k|_{op}^{1/2}$ yields

$$\begin{aligned} |s_7| \leq\ & C(\|\Psi_{0,\infty}\|_{\mathcal{W}^m_{\sigma(t)}} \| |k|_{op}^{1/2} R_{\pm 2,\infty}\|^2_{G^m_{\sigma(t)}} \\ & + \| |k|_{op}^{1/2} \Psi_{0,\infty}\|_{\mathcal{W}^m_{\sigma(t)}} \|R_{\pm 2,\infty}\|_{G^m_{\sigma(t)}} \| |k|_{op}^{1/2} R_{\pm 2,\infty}\|_{G^m_{\sigma(t)}} + \mathcal{O}(\delta)). \end{aligned}$$

Next we use that $\| |k|_{op}^{1/2} \Psi_{0,\infty}\|_{\mathcal{W}^m_{\sigma(t)}} = \mathcal{O}(\delta^{1/2})$ and $\delta^{1/2} ab \leq a^2 + \delta b^2$ such that

$$\begin{aligned} |s_7| \leq\ & C(\| |k|_{op}^{1/2} R_{\pm 2,\infty}\|^2_{G^m_{\sigma(t)}} \\ & + \delta^{1/2} \|R_{\pm 2,\infty}\|_{G^m_{\sigma(t)}} \| |k|_{op}^{1/2} R_{\pm 2,\infty}\|_{G^m_{\sigma(t)}} + \mathcal{O}(\delta)) \\ \leq\ & C_7(\| |k|_{op}^{1/2} R_{\pm 2,\infty}\|^2_{G^m_{\sigma(t)}} + \delta \|R_{\pm 2,\infty}\|^2_{G^m_{\sigma(t)}}). \end{aligned}$$

$\mathbf{s_3}$: It remains to estimate

$$\begin{aligned} \mathrm{Re} s_3 =\ & \mathrm{Re}(R_{1,\infty}, B_{1,1,\infty}(\Psi_{0,\infty}, R_{1,\infty}))_{G^m_{\sigma(t)}} \\ & + \mathrm{Re}(R_{1,\infty}, B_{1,-1,\infty}(\Psi_{0,\infty}, R_{-1,\infty}))_{G^m_{\sigma(t)}} \\ & + \mathrm{Re}(R_{-1,\infty}, B_{-1,1,\infty}(\Psi_{0,\infty}, R_{1,\infty}))_{G^m_{\sigma(t)}} \\ & + \mathrm{Re}(R_{-1,\infty}, B_{-1,-1,\infty}(\Psi_{0,\infty}, R_{-1,\infty}))_{G^m_{\sigma(t)}} \end{aligned}$$

We have the following representation

$$\begin{aligned} s_{j,j_1} &= (R_{j,\infty}, B_{j,j_1,\infty}(\Psi_{0,\infty}, R_{j_1,\infty}))_{G^m_{\sigma(t)}} \\ &= \int\int R_{j,\infty}(k) b_{j,j_1,\infty}(k, k-m, m) \Psi_{0,\infty,j,j_1}(k-m) R_{j_1,\infty}(m) \mathrm{d}m \mathrm{d}k. \end{aligned}$$

Since $\Psi_{0,\infty,j,j_1}(k-m)$ is strongly concentrated at the wave number $k=0$ we have

$$\begin{aligned} s_{j,j_1} &= \int\int R_{j,\infty}(k) b_{j,j_1,\infty}(k,0,k)\Psi_{0,\infty,j,j_1}(k-m) R_{j_1,\infty}(m)\mathrm{d}m\mathrm{d}k + \mathcal{O}(\delta) \\ &= s_{j,j_1,a} + s_{j,j_1,b}, \end{aligned}$$

with

$$\begin{aligned} s_{j,j_1,a} &= \int\int R_{j,\infty}(k) b_{j,j_1,\infty}(0,0,0)\Psi_{0,\infty,j,j_1}(k-m) R_{j_1,\infty}(m)\mathrm{d}m\mathrm{d}k + \mathcal{O}(\delta) \\ s_{j,j_1,b} &= \int\int R_{j,\infty}(k)(b_{j,j_1,\infty}(k,0,k) - b_{j,j_1,\infty}(0,0,0)) \\ &\qquad \times\Psi_{0,\infty,j,j_1}(k-m) R_{j_1,\infty}(m)\mathrm{d}m\mathrm{d}k + \mathcal{O}(\delta). \end{aligned}$$

Using $|b_{j,j_1,\infty}(k,0,k) - b_{j,j_1,\infty}(0,0,0)| \leq C|k|$ the term $s_{j,j_1,b}$ can be estimated similar as $s_7$ by

$$\begin{aligned} |s_{j,j_1,b}| &\leq C(\||k|_{op}^{1/2} R_{\pm1,\infty}\|^2_{G^m_{\sigma(t)}} \\ &\qquad +\delta^{1/2}\|R_{\pm1,\infty}\|_{G^m_{\sigma(t)}} \||k|_{op}^{1/2} R_{\pm1,\infty}\|_{G^m_{\sigma(t)}} + \mathcal{O}(\delta)) \\ &\leq C_3(\||k|_{op}^{1/2} R_{\pm1,\infty}\|^2_{G^m_{\sigma(t)}} + \delta\|R_{\pm1,\infty}\|^2_{G^m_{\sigma(t)}}). \end{aligned}$$

Similar to [Sch20] we obtain

$$\begin{aligned} \mathrm{Re}(s_{1,1,a}) &= 2i\int\int R_{-1,\infty}(k) b_{1,1,\infty}(0,0,0)\Psi_{0,\infty,1,1}(k-m) R_{1,\infty}(m)\mathrm{d}m\mathrm{d}k \\ &\quad -2i\int\int R_{1,\infty}(k)\overline{b_{1,1,\infty}}(0,0,0)\overline{\Psi_{0,\infty,1,1}}(k-m) R_{-1,\infty}(m)\mathrm{d}m\mathrm{d}k, \end{aligned}$$

for $G^0_{\sigma(t)}$. By definition we have $\overline{\Psi_{0,\infty,1,1}}(k-m) = \Psi_{0,\infty,1,1}(m-k)$ and $\overline{b_{1,1,\infty}}(0,0,0) = b_{1,1,\infty}(0,0,0) \in \mathbb{R}$. Therefore, the terms cancel each other by interchanging the role of $k$ and $m$ in the second term, i.e.

$$\begin{aligned} \mathrm{Re}(s_{1,1,a}) &= 2i\int\int R_{-1,\infty}(k) b_{1,1,\infty}(0,0,0)\Psi_{0,\infty,1,1}(k-m) R_{1,\infty}(m)\mathrm{d}m\mathrm{d}k \\ &\quad -2i\int\int R_{1,\infty}(m) b_{1,1,\infty}(0,0,0)\Psi_{0,\infty,1,1}(k-m) R_{-1,\infty}(k)\mathrm{d}m\mathrm{d}k \\ &= 0. \end{aligned}$$

Doing the same calculations for $s_{-1,1,a}$, $s_{1,-1,a}$ and $s_{-1,-1,a}$ yields

$$\mathrm{Re}(s_{1,1,a}) + \mathrm{Re}(s_{-1,1,a}) + \mathrm{Re}(s_{1,-1,a}) + \mathrm{Re}(s_{-1,-1,a}) = 0.$$

For $G^m_{\sigma(t)}$ we can use the fact that whenever a derivative falls on $\Psi_{0,\infty,j,j_1}$ we gain an additional power of $\delta$. The term where every $m$ derivatives fall on $R_{\pm 1,\infty}$ in the second component of the scalar product can be estimated line for line as in the case $G^0_{\sigma(t)}$

**The final estimates:** The previous estimates yield

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}E \leq\; & s_1 + s_5 + s_9 + |s_3| + |s_4| + |s_6| + |s_7| + |s_8| + |s_{10}| + |s_{11}| \\
\leq\; & -\eta \| |k|^{1/2}_{op} R_{1,\infty}\|^2_{G^m_{\sigma(t)}} - \eta \| |k|^{1/2}_{op} R_{2,\infty}\|^2_{G^m_{\sigma(t)}} - \eta \| |k|^{1/2}_{op} \mathbf{R}^c\|^2_{G^m_{\sigma(t)}} \\
& + C_6 \| |k|^{1/2}_{op} R_{2,\infty}\|^2_{G^m_{\sigma(t)}} \\
& + (C_4 + C_8)\delta \big( \|(R_{1,\infty}, R_{2,\infty})\|^2_{G^m_\sigma} \\
& \quad + \delta^{1/2} (\|(R_{1,\infty}, R_{2,\infty})\|^3_{G^m_\sigma} + \|\mathbf{R}^c\|^3_{G^m_\sigma}) + 1 \big) \\
& + C_{11}(\|\mathbf{R}^c\|^2_{G^m_\sigma} \\
& \quad + \delta(\|(R_{1,\infty}, R_{2,\infty})\|_{G^m_\sigma} \|\mathbf{R}^c\|_{X^m_\sigma} + \delta^{1/2}\|\mathbf{R}^c\|^3_{G^m_\sigma} + 1)) \\
& + C_{10} \| |k|^{1/2}_{op} \mathbf{R}^c\|^2_{G^m_{\sigma(t)}} + C_7 (\| |k|^{1/2}_{op} R_{\pm 2,\infty}\|^2_{G^m_{\sigma(t)}} + \delta \|R_{\pm 2,\infty}\|^2_{G^m_{\sigma(t)}}) \\
& + C_3 (\| |k|^{1/2}_{op} R_{\pm 1,\infty}\|^2_{G^m_{\sigma(t)}} + \delta \|R_{\pm 1,\infty}\|^2_{G^m_{\sigma(t)}}) \\
\leq\; & (-\eta + C_3) \|(|k|_{op})^{1/2} R_{1,\infty}\|^2_{G^m_{\sigma(t)}} \\
& + (-\eta + C_6 + C_7) \| |k|^{1/2}_{op} R_{2,\infty}\|^2_{G^m_{\sigma(t)}} \\
& + (-\eta + C_{10}) \| |k|^{1/2}_{op} \mathbf{R}^c\|^2_{G^m_{\sigma(t)}} \\
& + (C_3 + C_4 + C_7 + C_8)\delta \left(E + \delta^{1/2}(E^{3/2} + 1)\right) \\
\leq\; & (C_3 + C_4 + C_7 + C_8)\delta \left(E + \delta^{1/2}(E^{3/2} + 1)\right),
\end{aligned}$$

if $\eta > 0$ is chosen so large that

$$-\eta + C_3 < 0, \qquad -\eta + C_6 + C_7 < 0, \qquad -\eta + C_{10} < 0.$$

After this we choose $\delta > 0$ so small that

$$\delta^{1/2} E^{1/2} \leq 1 \tag{68}$$

is fulfilled. Hence we have

$$\frac{\mathrm{d}}{\mathrm{d}t}E \leq (C+1)\delta E + C\delta.$$

With the help of Gronwalls inequality we obtain

$$E(t) \leq (E(0) + C\delta t)e^{(C+1)\delta}t \leq (E(0) + CT_0)e^{(C+1)T_0} = M = \mathcal{O}(1).$$

The constant $M$ is independent of $\eta$, $T_0$ and $0 < \delta \ll 1$. We choose $\delta_0 > 0$ sufficiently small such that $\delta_0^{1/2} M^{1/2} < 1$ is fullfilled. This guarantees the validity of (68). This proves our main theorem 1.10.

# 9 Discussion

**Remark 9.1.** Since the Benjamin-Feir instability occurs for $\gamma = 1$ respectively $\mu < 1$ we have no hope to replace the Gevrey spaces by classical Sobolev spaces, like the plot of the spectral curves in the right panel of Figure 2 shows. However, it is a natural question if this might be possible for $\gamma = -1$ respectively $\mu > 1$. In that case the wave train is spectrally stable and such a result can be found in [BKZ21] for the NLS equation. At this point we have no answer if this is also possible for the ccKG equation and have to postpone this question to future research.

**Remark 9.2.** Although in the parameter region $\mathcal{P}_{\text{rest}}$, cf. the left panel of Figure 3, WME can be derived, it cannot be expected that the associated WME approximation makes correct predictions on the long $\mathcal{O}(1/\delta)$-time scale. In the left panel of Figure 3 we have a smooth curve of eigenvalues with positive real part of order $\mathcal{O}(1)$ at the wave number $k = 0$. This leads to growth rates of order $\mathcal{O}(\exp(1/\delta))$ on the long $\mathcal{O}(1/\delta)$-time scale. Therefore, to come to the long $\mathcal{O}(1/\delta)$-time scale, by nonlinear interaction and initially only terms of order $\mathcal{O}(\exp(-1/\delta))$ can be allowed. However, this is not the case and so we expect that the WME approximation fails in $\mathcal{P}_{\text{rest}}$ to make correct predictions.

**Remark 9.3.** We finally remark that the reconstruction of the solution in physical variables (3) requires the spatial integration of the local wave number $\psi = \partial_x \varphi$ to reconstruct the phase $\varphi$. As a consequence in the original $u$-variable only a local in space approximation result can be obtained. The size of the spatial domain where the WME approximation makes correct predictions is proportional to the inverse order of the higher order WME approximation constructed in Section 4.4. For details see for instance [DS09, BKS20].

# A Appendix

## A.1 Stability regions for $q \neq 0$

In this section we consider the case $q \neq 0$ and explain where parameter regions plotted in Figure 4 come from.

**Remark A.1.** To analyze $q \neq 0$ we look at the system (29)-(31), derive WME, and redo the calculations from Remark 2.7. We have

$$\gamma e^{2r_{q,\mu}} = 1 + q^2 - \mu^2,$$

and as above, we make the long wave ansatz

$$(r, \psi, \vartheta)(x,t) = (\check{r}, \check{\psi}, \check{\vartheta})(\delta x, \delta t),$$

with $0 < \delta \ll 1$ a small perturbation perimeter. Ignoring higher order terms yields the system

$$\begin{aligned} 0 &= \check{\vartheta}^2 + 2\mu\check{\vartheta} - \check{\psi}^2 - 2q\check{\psi} + \gamma e^{2r_{q,\mu}}(e^{2\check{r}} - 1), \\ \partial_T \check{\vartheta} &= 2(\partial_X \check{r})(\check{\psi} + q) + \partial_X \check{\psi} - 2(\partial_T \check{r})(\check{\vartheta} + \mu), \\ \partial_T \check{\psi} &= \partial_X \check{\vartheta}. \end{aligned}$$

For $\check{\vartheta}$ and $\check{\psi}$ small we can solve the first equation w.r.t. $\check{r}$ and get

$$\check{r}^*(\check{\vartheta}, \check{\psi}) = -\frac{\mu}{\gamma e^{2r_{q,\mu}}}\check{\vartheta} + \frac{q}{\gamma e^{2r_{q,\mu}}}\check{\psi} + h.o.t.$$

with the partial temporal and spatial derivatives

$$\begin{aligned} \partial_T \check{r}^* &= -\frac{\mu}{\gamma e^{2r_{q,\mu}}}\partial_T\check{\vartheta} + \frac{q}{\gamma e^{2r_{q,\mu}}}\partial_T\check{\psi} + h.o.t., \\ \partial_X \check{r}^* &= -\frac{\mu}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\vartheta} + \frac{q}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\psi} + h.o.t.. \end{aligned}$$

Inserting this in the equations for $\check{\vartheta}$ and $\check{\psi}$ yields

$$\begin{aligned}
\partial_T \check{\vartheta} =& 2\left(-\frac{\mu}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\vartheta} + \frac{q}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\psi}\right)(\check{\psi}+q) + \partial_X\check{\psi} \\
&- 2\left(-\frac{\mu}{\gamma e^{2r_{q,\mu}}}\partial_T\check{\vartheta} + \frac{q}{\gamma e^{2r_{q,\mu}}}\partial_T\check{\psi}\right)(\check{\vartheta}+\mu) \\
=& -\frac{2\mu}{\gamma e^{2r_{q,\mu}}}\left(\partial_X\check{\vartheta}\right)(\check{\psi}+q) + \frac{2q}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\psi}(\check{\psi}+q) + \partial_X\check{\psi} \\
&+ \frac{2\mu}{\gamma e^{2r_{q,\mu}}}\left(\partial_T\check{\vartheta}\right)(\check{\vartheta}+\mu) - \frac{2q}{\gamma e^{2r_{q,\mu}}}\left(\partial_T\check{\psi}\right)(\check{\vartheta}+\mu).
\end{aligned}$$

The linearization of this equation is given by

$$\partial_T\check{\vartheta} = \partial_X\check{\psi} + \frac{2\mu^2}{\gamma e^{2r_{q,\mu}}}\partial_T\check{\vartheta} - \frac{4q\mu}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\vartheta} + \frac{2q^2}{\gamma e^{2r_{q,\mu}}}\partial_X\check{\psi}.$$

Hence we find

$$\begin{aligned}
\left(1-\frac{2\mu^2}{\gamma e^{2r_{q,\mu}}}\right)\partial_T\check{\vartheta} = \left(\frac{1+q^2-\mu^2-2\mu^2}{1+q^2-\mu^2}\right)\partial_T\check{\vartheta} &= \left(\frac{1+q^2-3\mu^2}{1+q^2-\mu^2}\right)\partial_T\check{\vartheta} \\
&= -\frac{4q\mu}{1+q^2-\mu^2}\partial_X\check{\vartheta} + \frac{1+q^2-\mu^2+2q^2}{1+q^2-\mu^2}\partial_X\check{\psi} \\
&= -\frac{4q\mu}{1+q^2-\mu^2}\partial_X\check{\vartheta} + \frac{1+3q^2-\mu^2}{1+q^2-\mu^2}\partial_X\check{\psi},
\end{aligned}$$

and so finally in the long wave limit

$$\begin{aligned}
\partial_T\check{\vartheta} &= -\frac{4q\mu}{1+q^2-3\mu^2}\partial_X\check{\vartheta} + \frac{1+3q^2-\mu^2}{1+q^2-3\mu^2}\partial_X\check{\psi}, \\
\partial_T\check{\psi} &= \partial_X\check{\vartheta}.
\end{aligned}$$

Hence, the sign of $\frac{1+3q^2-\mu^2}{1+q^2-3\mu^2}$ determines the stability or instability of the wave train w.r.t. long wave perturbations. If $1+3q^2-\mu^2$ and $1+q^2-3\mu^2$ are both smaller than zero the spectral curves look similar the ones in the left panel of Figure 2. Spectral curves for other parameter values are plotted in Figure 5.

## A.2 Some technical estimates

Since the near identity changes of variables converges in $X^{\sigma,m,\delta}$-spaces we would like to provide some more details about the role of the $X^{\sigma,m,\delta}$-spaces in the estimates of the resonant terms $s_3$ and $s_7$ in the Gevrey spaces $G^m_{\sigma(t)}$. In the following we concentrate on the term

$$s_7 = (R_{2,\infty}, \sum_{j=\pm 2} B^*_{2,j,\infty}(\Psi_{0,\infty})\partial_x R_{j,\infty} + \mathcal{O}(\delta))_{G^m_{\sigma(t)}},$$

where $B^*_{2,j,\infty}(\Psi_{0,\infty})$, which is completely determined by $\Psi_{0,\infty}$, is the limit of the resonant terms after the infinitely many normal form transformations. We have

$$\begin{aligned}
\| & \sum_{j=\pm 2} B^*_{2,j,\infty}(\Psi_{0,\infty})\partial_x R_{j,\infty}\|_{G^m_{\sigma(t)}} \\
\leq \; & C\|e^{\sigma(t)(|k|)} \int \sum_{j=\pm 2} ((B^*_{2,j,\infty}(\Psi_{0,\infty}))(k,k-l)|l| R_{j,\infty}(l) dl)\|_{L^2_m} \\
\leq \; & C\| \sum_{j=\pm 2} \int (e^{\sigma(t)(|k-l|)}(B^*_{2,j,\infty}(\Psi_{0,\infty}))(k,k-l) e^{\sigma(t)(|l|)}|l| R_{j,\infty}(l))\|_{L^2_m}.
\end{aligned}$$

By the inequality [DKS16, eq. (45)] we obtain

$$\begin{aligned}
& \| \sum_{j=\pm 2} B^*_{2,j,\infty}(\Psi_{0,\infty})\partial_x R_{j,\infty}\|_{G^m_{\sigma(t)}} \\
& \quad \leq C\|B^*_{2,\pm 2,\infty}(\Psi_{0,\infty})\|_{X^{\sigma,m,\delta}} \|e^{\sigma(t)(|k|)}|k| R_{\pm 2,\infty})\|_{L^2_m} \\
& \quad \leq C\|B^*_{2,\pm 2,\infty}(\Psi_{0,\infty})\|_{X^{\sigma,m,\delta}} \||k|_{op} R_{\pm 2,\infty})\|_{G^m_{\sigma(t)}} \\
& \quad \leq C\|\Psi_{0,\infty}\|_{\mathcal{W}^m_{\sigma(t)}} \||k|_{op} R_{\pm 2,\infty}\|_{G^m_{\sigma(t)}}.
\end{aligned}$$

In the last step of this estimate we use that the construction is built such that the limit of the resonant terms in the norm of $X^{m,\delta}$ can be estimated by $\|\Psi_0\|$ and therefore by $\|\Psi_0\|_{\mathcal{W}^m_{\sigma(t)}}$. We can follow the steps for the estimates of $s_7$ in Section 8. Integration by parts with $|k|^{1/2}$ yields with the estimates above

$$
\begin{aligned}
|s_7| &= |(R_{2,\infty}, \sum_{j=\pm 2} B^*_{2,j,\infty}(\Psi_{0,\infty})\partial_x R_{j,\infty} + \mathcal{O}(\delta))_{G^m_{\sigma(t)}}| \\
&\leq |(|k|^{1/2}_{op} R_{2,\infty}, \sum_{j=\pm 2} B^*_{2,j,\infty}(\Psi_{0,\infty})|k|^{1/2}_{op} R_{j,\infty} + \mathcal{O}(\delta))_{G^m_{\sigma(t)}}| \\
&\quad + |(R_{2,\infty}, \sum_{j=\pm 2} |k|^{1/2}_{op} B^*_{2,j,\infty}(\Psi_{0,\infty})|k|^{1/2}_{op} R_{j,\infty} + \mathcal{O}(\delta))_{G^m_{\sigma(t)}}| \\
&\leq C\Big( \|\Psi_{0,\infty}\|_{\mathcal{W}^m_{\sigma(t)}} \||k|^{1/2}_{op} R_{\pm 2,\infty})\|_{G^m_{\sigma(t)}} \\
&\quad + \||k|^{1/2}_{op}\Psi_{0,\infty}\|_{\mathcal{W}^m_{\sigma(t)}} \||k|^{1/2}_{op} R_{\pm 2,\infty})\|_{G^m_{\sigma(t)}} + \mathcal{O}(\delta)\Big) .
\end{aligned}
$$

The remaining rest of the estimate for $s_7$ follows as in Section 8.

**Acknowledgement.** The paper is partially supported by the Deutsche Forschungsgemeinschaft DFG through the SFB 1173 "Wave phenomena" Project-ID 258734477.

# References

[BDS19] Roman Bauer, Wolf-Patrick Düll, and Guido Schneider. The Korteweg-de Vries, Burgers and Whitham limits for a spatially periodic Boussinesq model. *Proc. R. Soc. Edinb., Sect. A, Math.*, 149(1):191–217, 2019.

[BGSS09] Fabrice Béthuel, Philippe Gravejat, Jean-Claude Saut, and Didier Smets. On the Korteweg-de Vries long-wave approximation of the Gross-Pitaevskii equation. I. *Int. Math. Res. Not.*, 2009(14):2700–2748, 2009.

[BGSS10] Fabrice Béthuel, Philippe Gravejat, Jean-Claude Saut, and Didier Smets. On the Korteweg-de Vries long-wave approximation of the Gross-Pitaevskii equation. II. *Commun. Partial Differ. Equations*, 35(1):113–164, 2010.

[BKS20] Thomas J. Bridges, Anna Kostianko, and Guido Schneider. A proof of validity for multiphase Whitham modulation theory. *Proc. R. Soc. Lond., A, Math. Phys. Eng. Sci.*, 476(2243):19, 2020. Id/No 20200203.

[BKZ21] Thomas J. Bridges, Anna Kostianko, and Sergey Zelik. Validity of the hyperbolic Whitham modulation equations in Sobolev spaces. *J. Differ. Equations*, 274:971–995, 2021.

[BM95] Thomas J. Bridges and Alexander Mielke. A proof of the Benjamin-Feir instability. *Arch. Ration. Mech. Anal.*, 133(2):145–198, 1995.

[BSSZ20] Simon Baumstark, Guido Schneider, Katharina Schratz, and Dominik Zimmermann. Effective slow dynamics models for a class of dispersive systems. *J. Dyn. Differ. Equations*, 32(4):1867–1899, 2020.

[CDS14] Martina Chirilus-Bruckner, Wolf-Patrick Düll, and Guido Schneider. Validity of the KdV equation for the modulation of periodic traveling waves in the NLS equation. *J. Math. Anal. Appl.*, 414(1):166–175, 2014.

[CR10] D. Chiron and F. Rousset. The KdV/KP-I limit of the nonlinear Schrödinger equation. *SIAM J. Math. Anal.*, 42(1):64–96, 2010.

[Cra85] Walter Craig. An existence theory for water waves and the Boussinesq and Korteweg- deVries scaling limits. *Commun. Partial Differ. Equations*, 10:787–1003, 1985.

[CS11] Christopher Chong and Guido Schneider. The validity of the KdV approximation in case of resonances arising from periodic media. *J. Math. Anal. Appl.*, 383(2):330–336, 2011.

[DKS16] Wolf-Patrick Düll, Kourosh Sanei Kashani, and Guido Schneider. The validity of Whitham's approximation for a Klein-Gordon-Boussinesq model. *SIAM J. Math. Anal.*, 48(6):4311–4334, 2016.

[DS09] Wolf-Patrick Düll and Guido Schneider. Validity of Whitham's equations for the modulation of periodic traveling waves in the NLS equation. *J. Nonlinear Sci.*, 19(5):453–466, 2009.

[Dül12] Wolf-Patrick Düll. Validity of the korteweg–de vries approximation for the two-dimensional water wave problem in the arc length formulation. *Communications on Pure and Applied Mathematics*, 65(3):381–429, 2012.

[FS22] Reika Fukuizumi and Guido Schneider. Interchanging space and time in nonlinear optics modeling and dispersion management models. *Journal of Nonlinear Science*, 32(3):1–39, 2022.

[HdRS21] Tobias Haas, Björn de Rijk, and Guido Schneider. Validity of Whitham's modulation equations for dissipative systems with a conservation law: Phase dynamics in a generalized Ginzburg-Landau system. *Indiana Journal of Mathematics*, accepted, 07.06.2021.

[HS20] Tobias Haas and Guido Schneider. Failure of the n-wave interaction approximation without imposing periodic boundary conditions. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 100(6):e201900230, 2020.

[RS75] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II: Fourier Analysis, Self-Adjointness*, volume 2. Elsevier, 1975.

[Saf95] Mikhail V Safonov. The abstract cauchy-kovalevskaya theorem in a weighted banach space. *Communications on Pure and Applied Mathematics*, 48(6):629–637, 1995.

[Sch95] Guido Schneider. Validity and limitation of the Newell-Whitehead equation. *Math. Nachr.*, 176:249–263, 1995.

[Sch20] Guido Schneider. The KdV approximation for a system with unstable resonances. *Math. Methods Appl. Sci.*, 43(6):3185–3199, 2020.

[SSZ15] Guido Schneider, Danish Ali Sunny, and Dominik Zimmermann. The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions. *J. Dyn. Differ. Equations*, 27(3-4):1077–1099, 2015.

[SW00a] Guido Schneider and C. Eugene Wayne. Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. In *International conference on differential equations. Proceedings of the conference, Equadiff '99, Berlin, Germany, August 1–7, 1999. Vol. 1*, pages 390–404. Singapore: World Scientific, 2000.

[SW00b] Guido Schneider and C. Eugene Wayne. The long-wave limit for the water wave problem. I: The case of zero surface tension. *Commun. Pure Appl. Math.*, 53(12):1475–1535, 2000.

[Whi74] G. B. Whitham. *Linear and nonlinear waves.* John Wiley & Sons, Hoboken, NJ, 1974.