Available online at www.sciencedirect.com

ScienceDirect

J. Differential Equations 262 (2017) 5409–5424

Journal of Differential Equations

www.elsevier.com/locate/jde

A rigorous justification of the Matthews–Cox approximation for the Nikolaevskiy equation

Dominik Zimmermann

IADM, Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

Received 30 June 2016; revised 13 January 2017
Available online 7 February 2017

Abstract

The Nikolaevskiy equation is an example of a pattern forming system with marginally stable long modes. It has the unusual property that the typical Ginzburg–Landau scaling ansatz for the description of propagating patterns does not yield asymptotically consistent amplitude equations. Instead, another scaling proposed by Matthews and Cox can be used to formally derive a consistent system of modulation equations. We give a rigorous proof that this system makes correct predictions about the dynamics of the Nikolaevskiy equation.

Keywords: Amplitude equations; Multiscale analysis; Approximation; Pattern formation

1. Introduction

The Nikolaevskiy partial differential equation, given by

$$\partial_t u + u\partial_x u = -\partial_x^2 \left[r - (1 + \partial_x^2)^2 \right] u,$$

($x \in \mathbb{R}$, $t \geq 0$, $u(x,t) \in \mathbb{R}$) was proposed as a one-dimensional model for seismic waves in viscoelastic media, see [1]. It also serves as a paradigmatic model for a pattern forming system with Galilean invariance, see [4]. For our multiscale analysis near the onset of pattern formation, i.e., in the case $0 < r \ll 1$, it is convenient to introduce a small parameter $\varepsilon > 0$, such that $r = \varepsilon^2$, and write the Nikolaevskiy equation as

E-mail address: Dominik.Zimmermann@mathematik.uni-stuttgart.de.

http://dx.doi.org/10.1016/j.jde.2017.02.005
0022-0396/

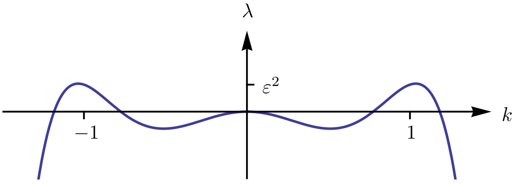

Fig. 1. Linear dispersion relation for the Nikolaevskiy equation: Turing instability with marginally stable long modes.

$$\partial_t u = L_\varepsilon u - \frac{1}{2}\partial_x(u^2), \qquad \text{where} \qquad L_\varepsilon = \partial_x^2(1+\partial_x^2)^2 - \varepsilon^2\partial_x^2. \tag{1}$$

Looking at the linear dispersion relation,

$$\lambda = -k^2(1-k^2)^2 + \varepsilon^2 k^2,$$

for modes $u(x,t) = e^{ikx+\lambda t}$, we see that for $\varepsilon > 0$ the spatially homogeneous steady state $u = 0$ becomes linearly unstable via a short wave instability. In addition to the classical Turing instability we also have a curve of eigenvalues touching the imaginary axis at the wave number $k = 0$, see Fig. 1. Hence, we have a spectral situation as considered in [3,9]. There, we derived amplitude equations for the propagation of small spatially periodic patterns using the typical Ginzburg–Landau scaling $X = \varepsilon x$, $T = \varepsilon^2 t$ for the large spatial and temporal scale, respectively, an $\mathcal{O}(\varepsilon)$ amplitude scaling for the pattern modes and an $\mathcal{O}(\varepsilon^2)$ amplitude scaling for the long modes.

In [4], Matthews and Cox pointed out that for the Nikolaevskiy equation such a scaling leads to amplitude equations that are asymptotically inconsistent in the sense that they contain $\mathcal{O}(1/\varepsilon)$ coefficients. Instead, they proposed an $\mathcal{O}(\varepsilon^{3/2})$ amplitude scaling of the pattern mode. Using the ansatz

$$\varepsilon^{3/2}\psi_{MC}(x,t) = \varepsilon^{3/2}A_1(\varepsilon x, \varepsilon^2 t)e^{ix} + \text{c.c.} + \varepsilon^2 A_0(\varepsilon x, \varepsilon^2 t),$$

where “c.c.” denotes the complex conjugate of the terms to the left, they derived the following system of amplitude equations for (1):

$$\begin{aligned} \partial_T A_1 &= 4\partial_X^2 A_1 + A_1 - iA_1A_0, \\ \partial_T A_0 &= \partial_X^2 A_0 - \partial_X(|A_1|^2). \end{aligned} \tag{2}$$

While it is reasonable to assume that $\varepsilon^{3/2}\psi_{MC}$ with A_1 and A_0 given as solutions of (2) is a good approximation to a true solution of (1), it is not obvious. In fact, there are cases where approximations based on formally correctly derived amplitude equations make wrong predictions about the original system, see, e.g., [6–8].

In case of the Nikolaevskiy equation, so far, the question of validity has been tackled by numerical investigations only. While in [4,10] the simulations seem to verify the unusual scaling by Matthews and Cox, more recent results raise doubts, see [11].

In this paper we give a rigorous proof that the Matthews–Cox approximation is indeed valid and that all the dynamics of the Matthews–Cox system (2) in the respective phase spaces can be found in the Nikolaevskiy equation as well. For the proof of validity we apply methods that have already proven useful in the context of the justification of the Ginzburg–Landau approximation.

It turns out that the justification of the Matthews–Cox approximation is even simpler than its analogues in [3,9] due to the unusual scaling.

Notation. Throughout this paper, many different constants are denoted with the same symbol C, as long as they are independent of the small parameter $0 < \varepsilon \ll 1$.

2. Preliminaries

In [3,9] we worked in Sobolev spaces, which had the disadvantage that the approximation results did not cover spatially periodic solutions or fronts, etc. In order to include such types of solutions, we choose to work in the spaces of functions that are uniformly locally Sobolev. The abstract theory for these spaces has been developed in [5] for the application to hydrodynamical problems. Our situation is much simpler, which facilitates the presentation.

2.1. Basics and notation

Let L^2 and L^2_{loc} denote the spaces of (equivalence classes of almost everywhere equal) functions $u : \mathbb{R} \to \mathbb{C}$ that are square-integrable on $\mathbb{R}$ or any compact subset of $\mathbb{R}$, respectively. Furthermore, let H^m, $m \in \mathbb{N}$, be the space of functions in L^2 whose first m weak derivatives $\partial_x^j u$, $j = 1, \dots, m$, lie in L^2.

We introduce weighted Sobolev spaces.

Definition 1. Let $n, m \in \mathbb{N}_0$ and $\rho(x) := (1 + x^2)^{1/2}$. Then we define

$$H^m(n) := \{u \in H^m \mid \|u\|_{H^m(n)} = \|u\rho^n\|_{H^m} < \infty\}.$$

We write $L^2(n)$ instead of $H^0(n)$.

Let $\mathcal{S}$ be the space of rapidly decreasing functions on $\mathbb{R}$. For any $u \in \mathcal{S}$ we define the Fourier transform $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ by

$$\mathcal{F}u(k) := \widehat{u}(k) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u(x) \mathrm{e}^{-\mathrm{i}kx}\, \mathrm{d}x.$$

It is easy to see that $\mathcal{F}$ extends to an isomorphism $\mathcal{F} : H^m(n) \to H^n(m)$ with inverse $\mathcal{F}^{-1}$, formally given by

$$\mathcal{F}^{-1}u(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u(k) \mathrm{e}^{\mathrm{i}kx}\, \mathrm{d}k.$$

Definition 2. Let

$$L^2_{u,l} := \left\{ u \in L^2_{\mathrm{loc}}(\mathbb{R}, \mathbb{C}) \mid \|u\|_{L^2_{u,l}} < \infty \right\},$$

where the norm $\|\cdot\|_{L^2_{u,l}}$ is defined by

$$\|u\|_{L^2_{u,l}} := \sup_{\xi\in\mathbb{R}} \left(\int_{\xi}^{\xi+1} |u(x)|^2 \,\mathrm{d}x \right)^{1/2} .$$

Furthermore, for $m \in \mathbb{N}_0$ we define

$$H^m_{u,l} := \left\{ u \in L^2_{u,l} \mid \|u\|_{H^m_{u,l}} < \infty \right\},$$

where the norm $\|\cdot\|_{H^m_{u,l}}$ is defined by

$$\|u\|_{H^m_{u,l}} := \max_{j=0,\dots,m} \|\partial_x^j u\|_{L^2_{u,l}}.$$

We note that the $H^m_{u,l}$ spaces are Banach algebras for $m \geq 1$ and that they are no Hilbert spaces.

2.2. *Multipliers*

Given $\widehat{M} \in L^\infty(\mathbb{R},\mathbb{C})$, we can define $\widehat{\mathcal{M}}, \mathcal{M} \in \mathscr{L}(L^2, L^2)$ by

$$\begin{aligned} \widehat{\mathcal{M}}\widehat{u} &= (k \mapsto \widehat{M}(k)\cdot \widehat{u}(k)) \\ \mathcal{M}u &= \mathcal{F}^{-1}(\widehat{\mathcal{M}}\mathcal{F}u), \end{aligned}$$

with $\|\mathcal{M}\|_{\mathscr{L}(L^2,L^2)} \leq C\|\widehat{M}\|_{L^\infty}$.

In a similar way we can define bounded operators on $H^m_{u,l}$-spaces. Since any $u \in L^2_{u,l}$ induces a tempered distribution $T_u \in \mathcal{S}'$, by

$$T_u(\phi) = \int_{\mathbb{R}} u(x)\phi(x)\,\mathrm{d}x, \quad \forall \phi \in \mathcal{S},$$

the Fourier transform $\mathcal{F}u := \mathcal{F}(T_u) \in \mathcal{S}'$ is well defined for any $u \in L^2_{u,l}$ in the sense of distributions. Since for any $u \in H^m_{u,l}$ we have

$$\begin{aligned} |(\mathcal{F}u)\phi| &= |T_u(\mathcal{F}\phi)| = |T_{u\rho^{-2}}\big((\mathcal{F}\phi)\rho^2\big)| \\ &\leq \|u\rho^{-2}\|_{L^2}\cdot\|\mathcal{F}\phi\|_{L^2(2)} \leq C\Big(\sum_{j=1}^{\infty}\frac{1}{j^2}\Big)\|u\|_{L^2_{u,l}}\|\phi\|_{H^2}, \end{aligned}$$

the multiplication $\widehat{M}\cdot(\mathcal{F}u)$, where $\widehat{M}\cdot(\mathcal{F}u)(\phi) = (\mathcal{F}u)(\widehat{M}\phi)$, is well defined for all $\widehat{M} \in C^2_b$. More generally, the following result holds.

Lemma 3. *Let $n \in \mathbb{Z}$ and $(\rho^n \widehat{M}) \in C^2_b$. Then $M_{u,l} := \mathcal{F}^{-1}(\widehat{M}\cdot(\mathcal{F}u))$ is a bounded operator from $H^q_{u,l}$ to $H^{q+n}_{u,l}$ for all $q \in \mathbb{N}_0$ with $q+n \geq 0$ and it holds*

$$\|M_{u,l}\|_{\mathscr{L}(H^q_{u,l}, H^{q+n}_{u,l})} \leq C(q,n)\|\rho^n \widehat{M}\|_{C^2_b}$$

with a constant $C(q,n)$ independent of $\widehat{M}$. We call $M_{u,l}$ the multiplier *corresponding to $\widehat{M}$.*

Proof. See [5, Lemma 5]. □

Remark 4. With the help of Lemma 3 the local well-posedness of the Nikolaevskiy equation in $H^m_{u,l}$ follows. To see this, we consider the associated stationary problem

$$(\lambda - L_\varepsilon)u = f.$$

For $\operatorname{Re}\lambda \geq 1$, we obtain in Fourier space the solution

$$\widehat{u} = \frac{1}{\lambda - \lambda_\varepsilon}\widehat{f}, \quad \text{where} \quad \lambda_\varepsilon(k) = -k^2(1-k^2)^2 + \varepsilon^2 k^2.$$

By direct calculations it follows that there exist constants $C, C_1 > 0$ such that

$$\left|\frac{1}{\lambda - \lambda_\varepsilon(k)}\right| \leq \frac{C}{|\lambda + C_1 k^6|}, \qquad \sup_{\substack{k\in\mathbb{R},\\ \operatorname{Re}\lambda \geq 1}} \left|\frac{\lambda'_\varepsilon(k)}{\lambda - \lambda_\varepsilon(k)}\right| \leq C, \qquad \sup_{\substack{k\in\mathbb{R},\\ \operatorname{Re}\lambda \geq 1}} \left|\frac{\lambda''_\varepsilon(k)}{\lambda - \lambda_\varepsilon(k)}\right| \leq C.$$

Hence, we have due to Lemma 3 that for all λ in the half-plane $\{\operatorname{Re}\lambda \geq 1\}$ the operator $(\lambda - L_\varepsilon)$ has a bounded inverse from $H^m_{u,l}$ to $H^m_{u,l}$ and satisfies

$$\left\|(\lambda - L_\varepsilon)^{-1}\right\|_{\mathscr{L}(H^m_{u,l}, H^m_{u,l})} \leq C\left\|(\lambda - \lambda_\varepsilon(\cdot))^{-1}\right\|_{C^2_b} \leq \frac{C}{|\lambda|}.$$

Hence, it follows that $L_\varepsilon : H^m_{u,l} \to H^m_{u,l}$ is sectorial for any $m \geq 0$, such that the analytic semigroup $\mathrm{e}^{L_\varepsilon t}$ is well defined. Then standard semigroup theory yields the local well-posedness of (1) in $C([0, T_1], H^m_{u,l})$ for $m \geq 6$ and some $T_1 > 0$, see, e.g., [2]. An analogous argument gives the local well-posedness of the Matthews–Cox system (2) in $C([0, T_0], (H^m_{u,l})^2)$ for all $m \geq 2$. □

In order to estimate terms coming from the approximation, we investigate how multipliers act on scaled functions with unscaled modulation. To this end, we introduce the scaling operator $\mathbf{S}_\varepsilon$ by $(\mathbf{S}_\varepsilon u)(x) := u(\varepsilon x)$ and the modulation operator $\mathbf{m}_\kappa$ by $(\mathbf{m}_\kappa u)(x) := u(x) \cdot \mathrm{e}^{\mathrm{i}\kappa x}$.

Lemma 5. *Let $n \in \mathbb{N}$ and $\rho^{-n}\widehat{M} \in C^2_b$ and $M_{u,l}$ be the corresponding multiplier as defined in Lemma 3. Then $(M_{u,l}\mathbf{m}_\kappa\mathbf{S}_\varepsilon) : H^q_{u,l} \to H^{q-r}_{u,l}$ is a bounded operator for all $q \geq r \geq n$ with*

$$\|M_{u,l}\mathbf{m}_\kappa\mathbf{S}_\varepsilon\|_{\mathscr{L}(H^q_{u,l}, H^{q-r}_{u,l})} \leq C(q,n,r)\|\rho^{-n}\widehat{M}(\varepsilon\cdot + \kappa)\|_{C^2_b} \cdot \|\mathbf{S}_\varepsilon\|_{\mathscr{L}(H^{q-n}_{u,l}, H^{q-r}_{u,l})},$$

where the constant $C(q,n,r)$ is independent of $\widehat{M}$, ε and κ.

If additionally, $\widehat{M}(k+\kappa) = \mathcal{O}(k^s)$ for $k \to 0$ and $s \leq n$, then

$$\|\rho^{-n}\widehat{M}(\varepsilon\cdot + \kappa)\|_{C^2_b} = \mathcal{O}(\varepsilon^s).$$

Proof. See [5, Section 3.3]. □

Remark 6 *(Scaling properties).* Note that the operator norm of the scaling operator $\mathbf{S}_\varepsilon$ depends qualitatively on the exponents of the considered $H^m_{u,l}$-spaces. For example, the best estimate we can get for $\mathbf{S}_\varepsilon : L^2_{u,l} \to L^2_{u,l}$ is

$$\|\mathbf{S}_\varepsilon\|_{\mathscr{L}(L^2_{u,l}, L^2_{u,l})} \leq C\varepsilon^{-1/2},$$

since $\|\mathbf{S}_\varepsilon u\|_{L^2_{u,l}} = \varepsilon^{-\alpha}\|u\|_{L^2_{u,l}}$ for $u(x) = |x|^{-\alpha}$, $\alpha \in [0, 1/2)$.

However, we have for $n \geq 1$

$$\|\mathbf{S}_\varepsilon u\|_{L^2_{u,l}} \leq \|u\|_{L^\infty} \leq C\|u\|_{H^n_{u,l}},$$

such that for $q \geq n$ we have $\|\mathbf{S}_\varepsilon\|_{\mathscr{L}(H^q_{u,l}, H^{q-n}_{u,l})} \leq C$. Together with the second statement of Lemma 5 this implies that for $s \leq q$ we have

$$\|\partial_x^s \mathbf{S}_\varepsilon\|_{\mathscr{L}(H^q_{u,l}, H^{q-s}_{u,l})} = \mathcal{O}(\varepsilon^s).$$

This makes the formal calculation $\partial_x[u(\varepsilon\cdot)] = \varepsilon(\partial_X u)(\varepsilon\cdot)$ rigorous, i.e., derivatives w.r.t. x gain one order in ε. Thus, we conclude that for $q \in \mathbb{N}$ we have

$$\|\mathbf{S}_\varepsilon u\|_{H^q_{u,l}} \leq C\big(\|\mathbf{S}_\varepsilon u\|_{H^{q-1}_{u,l}} + \|\partial_x^q \mathbf{S}_\varepsilon u\|_{L^2_{u,l}}\big) \leq C\big(\|u\|_{H^q_{u,l}} + \varepsilon^q\|u\|_{H^q_{u,l}}\big).$$

Hence, for $q \in \mathbb{N}$ the scaling operator $\mathbf{S}_\varepsilon : H^q_{u,l} \to H^q_{u,l}$ has norm bounded independently of $\varepsilon > 0$. □

3. The approximation results

We are now able to state our main result.

Theorem 7. *Let $m_A \geq 12$, $m_A - 6 \geq m \geq 6$ and $(A_1, A_0) \in C([0, T_0], (H^{m_A}_{u,l})^2)$ be a solution of the Matthews–Cox system (2). Then there exist constants $\varepsilon_0 > 0$ and $C > 0$, such that for all $\varepsilon \in (0, \varepsilon_0)$ there exist solutions $u \in C([0, T_0/\varepsilon^2], H^m_{u,l})$ of the Nikolaevskiy equation (1) satisfying*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \big\|u(\cdot, t) - \big(\varepsilon^{3/2}A_1(\varepsilon\cdot, \varepsilon^2 t)e^{i\cdot} + \text{c.c.}\big)\big\|_{H^m_{u,l}} \leq C\varepsilon^2.$$

We give the proof of this result in the next section. There, it will become clear that the approximation result can be refined in order to include the approximate evolution of the long modes associated to wave numbers around $k \approx 0$.

Corollary 8. *With the same assumptions as in Theorem 7 the following holds.*

There exist constants $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ there exist solutions $u \in C([0, T_0/\varepsilon^2], H^m_{u,l})$ of the Nikolaevskiy equation (1) satisfying

$$\sup_{t\in[0,T_0/\varepsilon^2]} \left\| u(\cdot,t) - \left(\varepsilon^{3/2}A(\varepsilon\cdot,\varepsilon^2 t)e^{i\cdot} + \varepsilon^2 A_0(\varepsilon\cdot,\varepsilon^2 t) + \text{c.c.}\right)\right\|_{H^m_{u,l}} \le C\varepsilon^{5/2}.$$

Due to Sobolev's embeddings, the $H^m_{u,l}$-norm in the estimates of Theorem 7 and Corollary 8 can be replaced by the more common C_b^{m-1}-Norm.

Remark 9. With the same method of proof – even simpler in many respects – it is possible to obtain analogous approximation results for the usual Sobolev spaces H^m. Due to the scaling properties of H^m, we have to make the formal error smaller than in the case of $H^m_{u,l}$, which leads to the stronger regularity conditions $m_A \ge 14$, $m_A - 8 \ge m \ge 6$. Otherwise, the assertion of the approximation results remain almost the same. In fact, we only have to replace $H^m_{u,l}$ by H^m and decrease the order of ε in the respective estimates by $1/2$ due to the scaling properties of H^m. □

4. Controlling the error

The assertions of Theorem 7 and Corollary 8 follow if we can show that there exist solutions u of (1) of the form

$$u = \varepsilon^{3/2}\psi + \varepsilon^{5/2}R$$

with

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|R(t)\|_{H^m_{u,l}} = \mathcal{O}(1)$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon^{3/2}\psi - (\varepsilon^{3/2}A_1(\varepsilon\cdot,\varepsilon^2 t)e^{ix} + \text{c.c.} + \varepsilon^2 A_0(\varepsilon\cdot,\varepsilon^2 t))\|_{H^m_{u,l}} = \mathcal{O}(\varepsilon^{5/2}).$$

It will turn out that we even have $\mathcal{O}(\varepsilon^3)$ in the last equation.

As a necessary condition, the approximation $\varepsilon^{3/2}\psi$ has to be chosen in such a way that the formal error, the so-called residual, $\text{Res}(\varepsilon^{3/2}\psi)$, defined by

$$\text{Res}\, v := -\partial_t v + L_\varepsilon v + \frac{1}{2}\partial_x(u^2)$$

is sufficiently small.

As in [5,9] we use mode filters to split the error R and the approximation ψ into different parts corresponding to modes in Fourier space with different growth rates. Finally, an argument based on Gronwall's inequality will give the desired result.

4.1. Mode filters and splitting

Let $\delta = 0.05$ and $\widehat{\chi}_0 \in C_0^\infty(\mathbb{R},[0,1])$ with $\text{supp}\,\widehat{\chi}_0 \in [-2\delta, 2\delta]$ and $\widehat{\chi}_0(k) = 1$ for all $k \in [-\delta,\delta]$. We define the following multipliers E_0, $E_{\pm 1}$, E_c, E_s by

$$\begin{aligned} E_0 u &:= \mathcal{F}^{-1}(\widehat{\chi}_0 \widehat{u}), \\ E_{\pm 1} u &:= \mathcal{F}^{-1}(\widehat{\chi}_0(\cdot \mp 1)\widehat{u}), \\ E_c u &:= E_{-1} + E_1, \\ E_s u &:= 1 - E_0 - E_c. \end{aligned}$$

Due to the compact support of χ_0 it follows from Lemma 3 that E_0 and E_c are bounded linear mappings from $H^q_{u,l}$ to $H^{q+m}_{u,l}$ for any $m \geq 0$.

The mode filters defined above are no projections. Hence, we proceed as in [5] and introduce auxiliary mode filters E_0^h, $E_{\pm 1}^h$, E_c^h, E_s^h. The mode filters E_0^h and E_c^h are defined in the same way as above, with $\widehat{\chi}_0$ replaced by $\widehat{\chi}_0^h := \mathbf{S}_{1/2}\widehat{\chi}_0$. The mode filter E_s^h is defined by

$$E_s^h u = \mathcal{F}^{-1}\big((1 - \mathbf{S}_2\widehat{\chi}_0 - (\mathbf{S}_2\widehat{\chi}_0)(\cdot - 1) - (\mathbf{S}_2\widehat{\chi}_0)(\cdot + 1))\widehat{u}\big)$$

Thus, we have

$$E_j = E_j E_j^h = E_j^h E_j, \quad j = 0, \pm 1, c, s.$$

Furthermore, we have $\partial_t E_j = E_j \partial_t$, $L_\varepsilon E_j = E_j L_\varepsilon$ as well as $\partial_t E_j^h = E_j^h \partial_t$, $L_\varepsilon E_j^h = E_j^h L_\varepsilon$ for $j = 0, \pm 1, c, s$.

Now we split the approximation $\varepsilon^{3/2}\psi$ and the error $\varepsilon^2 R$ in the following way,

$$\begin{aligned} \varepsilon^{3/2}\psi &= \varepsilon^{3/2}\psi_c + \varepsilon^2\psi_0 + \varepsilon^3\psi_s, \\ \varepsilon^{5/2}R &= \varepsilon^{5/2}R_c + \varepsilon^3 R_0 + \varepsilon^4 R_s, \end{aligned}$$

where $\psi_c = E_c\psi$, $\varepsilon^{1/2}\psi_0 = E_0\psi$, $\varepsilon^{3/2}\psi_s = E_s\psi$, $R_c = E_c R$, $\varepsilon^{1/2}R_0 = E_0 R$, $\varepsilon^{3/2}R_s = E_s R$. Since ψ_0, ψ_c, R_0, R_c have compact support in Fourier space (in the sense of distributions), the convolution $\widehat{\psi}_i * \widehat{R}_j$ is well defined for $i, j \in \{0, c\}$. Hence, it is easy to prove that

$$\begin{aligned} &E_0(\psi_0 R_c) = E_0(\psi_c R_0) = E_0(R_c R_0) = E_0(\psi_c\psi_0) = 0, \\ &E_c(\psi_0 R_0) = E_c(\psi_c R_c) = E_c(R_c^2) = 0, \end{aligned}$$

due to disjoint supports in Fourier space. Thus, we obtain the following system for the different parts of the error

$$\begin{aligned} \partial_t R_c &= L_\varepsilon R_c + \varepsilon^2 N_c(\psi, R) + \varepsilon^3 g_c(\psi, R) + \varepsilon^{-5/2}\mathrm{Res}_c, \\ \partial_t R_s &= L_\varepsilon R_s + N_s(\psi_c, R_c) + \varepsilon^{1/2} g_s(\psi, R) + \varepsilon^{-4}\mathrm{Res}_s, \\ \partial_t R_0 &= L_\varepsilon R_0 + \varepsilon\partial_x N_0(\psi_c, R_c) + \varepsilon^{3/2}\partial_x g_0(\psi, R) + \varepsilon^{-3}\mathrm{Res}_0, \end{aligned} \tag{3}$$

where

$$
\begin{aligned}
N_c(\psi, R) &= E_c\Big(\partial_x(\psi_c R_0 + \psi_0 R_c)\Big),\\
g_c(\psi, R) &= \frac{1}{2} E_c\Big(\partial_x\big[2(\psi_c R_s + \psi_s R_c + R_c R_0) + \varepsilon^{1/2}(\psi_0 R_s + \psi_s R_0)\Big)\\
&\qquad + 2\varepsilon R_c R_s + 2\varepsilon^{3/2}(\psi_s R_s + R_0 R_s) + \varepsilon^{5/2} R_s^2\big],\\
N_s(\psi_c, R_c) &= E_s(\partial_x(\psi_c R_c)),\\
g_s(\psi, R) &= \frac{1}{2} E_s\Big(\partial_x\big[2(\psi_c R_0 - \psi_0 R_c) + \varepsilon^{1/2}(R_c^2 + 2\psi_0 R_0) + 2\varepsilon(\psi_c R_s + \psi_s R_c)\\
&\qquad + \varepsilon^{3/2}(R_0^2 + 2\psi_0 R_s + 2\psi_s R_0) + 2\varepsilon^2 R_c R_s\\
&\qquad + \varepsilon^{5/2}(\psi_s R_s + R_0 R_s) + \varepsilon^{7/2} R_s^2\big]\Big),\\
N_0(\psi_c, R_c) &= E_0(\psi_c R_c),\\
g_0(\psi, R) &= \frac{1}{2} E_0\Big(\big[2(\psi_c R_0 - \psi_0 R_c) + \varepsilon^{1/2}(R_c^2 + 2\psi_0 R_0) + 2\varepsilon(\psi_c R_s + \psi_s R_c)\\
&\qquad + \varepsilon^{3/2}(R_0^2 + 2\psi_0 R_s + 2\psi_s R_0) + 2\varepsilon^2 R_c R_s\\
&\qquad + \varepsilon^{5/2}(\psi_s R_s + R_0 R_s) + \varepsilon^{7/2} R_s^2\big]\Big).
\end{aligned}
$$

4.2. The long-time estimates

For the long-time estimates, we first construct an approximation that lies close to $\varepsilon^{3/2}\psi_{MC}$ and makes the residual sufficiently small.

Lemma 10. *Let $m_A \geq 12$, $m_A - 6 \geq m \geq 6$. There exists an approximation $\varepsilon^{3/2}\psi$ such that*

$$
\begin{aligned}
\sup_{t\in[0,T_0/\varepsilon^2]} \|\psi_j\|_{H^m_{u,l}} &= \mathcal{O}(1), \text{ for } j = 0, c, s,\\
\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon^{3/2}\psi_{MC} - \varepsilon^{3/2}\psi\|_{H^m_{u,l}} &= \mathcal{O}(\varepsilon^3),\\
\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon^{-5/2}\mathrm{Res}_c\|_{H^m_{u,l}} &= \mathcal{O}(\varepsilon^2),\\
\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon^{-3}\mathrm{Res}_0\|_{H^m_{u,l}} &= \mathcal{O}(\varepsilon^2),\\
\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon^{-4}\mathrm{Res}_s\|_{H^m_{u,l}} &= \mathcal{O}(1).
\end{aligned}
$$

Proof. We refine the ansatz $\varepsilon^{3/2}\psi_{MC}$ by adding a higher order correction term:

$$
\begin{aligned}
\varepsilon\psi(x,t) &= \varepsilon^{3/2} A_1(X,T)\mathrm{e}^{\mathrm{i}x} + \text{c.c.}\\
&\quad + \varepsilon^2 A_0(X,T)\\
&\quad + \varepsilon^3 A_2(X,T)\mathrm{e}^{2\mathrm{i}x} + \text{c.c.}
\end{aligned}
$$

If we choose A_1, A_0 to satisfy the Matthews–Cox system (2) and

$$A_2 = -\frac{i}{36} A_1^2, \tag{4}$$

all terms in $\mathrm{Res}(\varepsilon^{3/2}\psi)$ have at least a prefactor ε^4. Furthermore, we have that terms proportional to $\mathrm{e}^{0\mathrm{i}x}$ have at least a prefactor ε^5 and terms proportional to $\mathrm{e}^{\pm\mathrm{i}x}$ have at least a prefactor $\varepsilon^{9/2}$.

Using (4) and (2), we can replace the terms A_2 and $\partial_T A_2$ in the residual, such that $\mathrm{Res}(\varepsilon^{3/2}\psi)$ can be written as the sum of products of A_1, A_0 and the respective derivatives w.r.t. X. Since the highest spatial derivative is ∂_X^6, we have that $\mathrm{Res}(\varepsilon^{3/2}\psi) \in C([0, T_0/\varepsilon^2], H^{m_A-6})$. Hence, in order to estimate the residual in the $H^m_{u,l}$-norm with $m \geq 6$, we need $m_A \geq 12$, $m_A - 6 \geq m \geq 6$, which gives the regularity conditions of Theorem 7 and Corollary 8.

Since for $m \geq 1$ the $H^m_{u,l}$ spaces are algebras and the scaling operator $\mathbf{S}_\varepsilon : H^m_{u,l} \to H^m_{u,l}$ has norm bounded independently of ε due to Remark 6, any terms in the residual with a prefactor ε^4 are also $\mathcal{O}(\varepsilon^4)$ w.r.t. $\|\cdot\|_{C([0,T_0/\varepsilon^2],H^m_{u,l})}$. This immediately gives the estimate for Res_s as well as the estimate

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon^{3/2}\psi_{MC} - \varepsilon^{3/2}\psi\|_{H^m_{u,l}} = \mathcal{O}(\varepsilon^3).$$

That the order of Res_0 is not influenced by the lower order terms proportional to $\mathrm{e}^{\pm\kappa\mathrm{i}x}$ with $\kappa = 1, 2, 3$, can be seen with the help of Lemma 5. Since $\widehat{\chi}_0$ vanishes outside a neighbourhood of $k = 0$, we have for $\kappa = 1, 2, 3$, that $\widehat{\chi}_0(k \pm \kappa) = \mathcal{O}(k^s)$ for $k \to 0$ and any $s \in \mathbb{N}$. Let $w \in H^{m_A-\ell}_{u,l}$ stand for the terms proportional to $\mathrm{e}^{\mathrm{i}\kappa x}$. Then we have

$$\begin{aligned}\|E_0 \mathbf{m}_\kappa \mathbf{S}_\varepsilon w\|_{H^m_{u,l}} &\leq C\|\rho^{-n}\widehat{\chi}_0(\varepsilon\cdot + j)\|_{C_b^2} \cdot \|\mathbf{S}_\varepsilon\|_{\mathscr{L}(H^{m_A-n}_{u,l}, H^m_{u,l})} \|w\|_{H^{m_A-\ell-n}_{u,l}} \\ &\leq C\varepsilon^n,\end{aligned}$$

if $m_A - \ell - n \geq m$ due to Lemma 5 and Remark 6. Since the lowest order terms in the residual do not contain highest order derivatives, we have that $w \in H^{m_A-\ell}_{u,l}$ with $m_A - \ell \geq m + 1$, such that n can be chosen greater than or equal to 1.

Similarly, we see that we only have to consider the terms proportional to $\mathrm{e}^{\pm\mathrm{i}x}$ for the order of Res_c. The same argument also implies that $\sup_{t\in[0,T_0/\varepsilon^2]} \|\psi_j\|_{H^m_{u,l}} = \mathcal{O}(1)$ for $j = 0, c, s$. $\square$

Using that the $H^m_{u,l}$ spaces are Banach algebras for $m \geq 1$ and the fact that E_0, E_c are bounded mappings from $H^m_{u,l}$ to $H^{m+n}_{u,l}$ for any $n \geq 0$, we obtain that there exists a constant $C > 0$ depending on ψ but neither on R nor $\tau \in [0, T_0/\varepsilon^2]$, such that

$$\begin{aligned}\|N_c(\psi, R)(\tau)\|_{H^m_{u,l}} &\leq C\mathcal{R}(\tau), \\ \|g_c(\psi, R)(\tau)\|_{H^m_{u,l}} &\leq C(\mathcal{R} + \mathcal{R}^2)(\tau), \\ \|N_s(\psi_c, R_c)(\tau)\|_{H^{m-1}_{u,l}} &\leq C\|R_c(\tau)\|_{H^m_{u,l}}, \\ \|g_s(\psi, R)(\tau)\|_{H^{m-1}_{u,l}} &\leq C(\mathcal{R} + \varepsilon^{1/2}\mathcal{R}^2)(\tau), \\ \|N_0(\psi_c, R_c)(\tau)\|_{H^m_{u,l}} &\leq C\|R_c(\tau)\|_{H^m_{u,l}}, \\ \|g_0(\psi, R)(\tau)\|_{H^m_{u,l}} &\leq C(\mathcal{R} + \varepsilon^{1/2}\mathcal{R}^2)(\tau),\end{aligned} \tag{5}$$

for all $\tau \in [0, T_0/\varepsilon^2]$, where

$$\mathcal{R} := \|R_c\|_{H^m_{u,l}} + \|R_s\|_{H^m_{u,l}} + \|R_0\|_{H^m_{u,l}}.$$

In order to control the error on the $\mathcal{O}(1/\varepsilon^2)$ time scale, we need the following estimates on the growth or decay rates for the semigroup acting on critical, neutral and stable Fourier modes.

Lemma 11. *Let $S_j(t) := \mathrm{e}^{L_\varepsilon t} E^h_j$ for $j = 0, c, s$. Then there exist constants $C, \sigma > 0$ such that for any $m, \ell \in \mathbb{N}_0$ the following estimates hold*

$$\begin{aligned}
\|S_0(t)\|_{\mathscr{L}(H^m_{u,l}, H^m_{u,l})} &\le C, \\
\|S_0(t)\partial_x\|_{\mathscr{L}(H^m_{u,l}, H^m_{u,l})} &\le C t^{-1/2}, \\
\|S_c(t)\|_{\mathscr{L}(H^m_{u,l}, H^m_{u,l})} &\le C \mathrm{e}^{\varepsilon^2 t}, \\
\|S_s(t)\|_{\mathscr{L}(H^m_{u,l}, H^{m+\ell}_{u,l})} &\le C \min\{1, t^{-\ell/6}\} \mathrm{e}^{-\sigma t}.
\end{aligned}$$

Remark 12. We see that the semigroups behave in the same way as they would for the usual Sobolev spaces H^m. For example, we have

$$\begin{aligned}
\|S_c(t)u\|_{H^m} &\le C \|k \mapsto \mathrm{e}^{\lambda_\varepsilon(k)t} \chi^h_c(k) \widehat{u}(k)\|_{L^2(m)} \\
&\le C \sup_{k \in \mathbb{R}} \mathrm{e}^{\lambda_\varepsilon(k)t} \|\widehat{u}\|_{L^2(m)} \le C \mathrm{e}^{\varepsilon^2 t} \|u\|_{H^m}.
\end{aligned}$$

We cannot transfer the above method to $u \in H^m_{u,l}$, however, since then in general $\widehat{u}$ is a tempered distribution.

Since the semigroups $S_j(t)$ can be seen as t-dependent multipliers, it is tempting to use Lemma 3 in order to obtain the desired estimates. However, it turns out that the estimate with Lemma 3 is too rough. For example, we would obtain

$$\|S_0(t)\|_{\mathscr{L}(H^m_{u,l}, H^m_{u,l})} \le C(m) \|k \mapsto \mathrm{e}^{\lambda_\varepsilon(k)t} \chi^h_0(k)\|_{C^2_b} = \mathcal{O}(t).$$

Hence, a more careful study is required. □

Proof of Lemma 11. The multipliers S_j can be expressed with the help of the convolution, since

$$S_j u = \mathcal{F}^{-1}(\mathcal{F}(S_j u)) = \mathcal{F}^{-1}\big(\mathrm{e}^{\lambda_\varepsilon(\cdot)t} \chi^h_j \widehat{u}\big) = \big(\mathcal{F}^{-1}(\mathrm{e}^{\lambda_\varepsilon(\cdot)t} \chi^h_j) * u\big). \tag{6}$$

We note that $S_j E_j = E_j S_j = S_j$ for $j = 0, c, s$. Since E_0, E^h_0, E_c and E^h_c have compact support in Fourier space, we have that $\|E_0 u\|_{H^m_{u,l}} \le C(m)\|u\|_{L^2_{u,l}}$ for some constant $C(m) > 0$ independent of u. Therefore, we have

$$\|S_j(t)u\|_{H^m_{u,l}} = \|S_j(t)E_j u\|_{H^m_{u,l}} = \|E_j(S_j(t)u)\|_{H^m_{u,l}} \le C(m)\|S_j(t)u\|_{L^2_{u,l}}$$

for $j = 0, c$. In order to use representation (6) for an estimate for the S_j, we note that an analogue of Young's inequality holds in $L^2_{u,l}$ as well. More precisely, we have for any $u \in L^1$ and any $v \in L^p_{u,l}$, $p \in [1, \infty]$ that

$$\|u * v\|_{L^p_{u,l}} \leq \|u\|_{L^1} \cdot \|v\|_{L^p_{u,l}}.$$

The proof of this fact goes exactly along the lines of the classical case. Thus, we have for $\ell \in \{0, 1\}$

$$\|S_j(t)\partial_x^\ell u\|_{H^m_{u,l}} \leq C(m)\|k \mapsto \mathcal{F}^{-1}(\mathrm{e}^{\lambda_\varepsilon(k)t}\chi_j^h(k)(\mathrm{i}k)^\ell)\|_{L^1} \cdot \|u\|_{L^2_{u,l}}.$$

From now on, we consider the case $j = 0$. Due to the form of the eigenvalue curve $\lambda_\varepsilon(k)$ near $k = 0$, we expect diffusive behaviour. First, we note that diffusion acts in $H^m_{u,l}$ spaces in an analogous fashion as in the usual Sobolev spaces. It follows by direct computations that

$$\|\mathcal{F}^{-1}(k \mapsto \mathrm{e}^{-k^2t})\|_{L^1} = \mathcal{O}(1), \quad \|\mathcal{F}^{-1}(k \mapsto \mathrm{e}^{-k^2t} \cdot (\mathrm{i}k))\|_{L^1} = \mathcal{O}(t^{-1/2}).$$

Next, we show that applying the cut-off function χ_0^h does not change the qualitative behaviour. For simplicity we only consider

$$\begin{aligned}
&\|\mathcal{F}^{-1}(k \mapsto \mathrm{e}^{-k^2t}\chi_0^h(k))\|_{L^1} \\
&= \underbrace{\int_{-\infty}^{-1} |\mathcal{F}^{-1}(k \mapsto \mathrm{e}^{-k^2t}\chi_0^h(k))(x)|\,\mathrm{d}x}_{=:I_-(t)} + \underbrace{\int_{-1}^{1} |\ldots|\,\mathrm{d}x}_{=:I_0(t)} + \underbrace{\int_{1}^{\infty} |\ldots|\,\mathrm{d}x}_{=:I_+(t)}.
\end{aligned}$$

Since $\operatorname{supp}\chi_0^h \subset [-4\delta, 4\delta]$, we have

$$\begin{aligned}
I_0(t) &= \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} \left| \int_{-\infty}^{\infty} \mathrm{e}^{-k^2t}\chi_0^h(k)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k \right| \mathrm{d}x \\
&\leq \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} \int_{-4\delta}^{4\delta} \mathrm{d}k\,\mathrm{d}x = \mathcal{O}(1).
\end{aligned}$$

For $|x| \geq 1$ we apply partial integration twice, in order to find

$$\begin{aligned}
&\sqrt{2\pi}\mathcal{F}^{-1}(k \mapsto \mathrm{e}^{-k^2t}(1 - \chi_0^h(k)))(x) = \int_{-\infty}^{\infty} \mathrm{e}^{-k^2t}(1 - \chi_0^h(k))\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k \\
&= -\frac{2}{x^2}\int_{2\delta}^{\infty} \Big[(4k^2t^2 - 2t)(1 - \chi_0^h(k)) + 4kt(\chi_0^h)'(k) - (\chi_0^h)''(k)\Big]\mathrm{e}^{-k^2t}\cos(kx)\,\mathrm{d}k \\
&=: \Phi(x, t).
\end{aligned}$$

Thus, we have

$$
\begin{aligned}
|\Phi(x,t)| &\le \frac{C}{x^2}(1+\|\chi_0^h\|_{C_b^2}) \sup_{k>0}\Big((4k^2t+2+4\delta)\mathrm{e}^{-\frac{k^2}{2}t}\Big)\int_{2\delta}^{\infty} t\mathrm{e}^{-\frac{k^2}{2}t}\,\mathrm{d}k \\
&\le \frac{C}{x^2}\int_{2\delta}^{\infty} t\mathrm{e}^{-\delta k t}\,\mathrm{d}k \le \frac{C}{x^2}\cdot \mathrm{e}^{-2\delta^2 t}.
\end{aligned}
$$

This gives the estimate

$$
\begin{aligned}
I_{\pm}(t) &\le \int_1^{\infty} |\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{-k^2t})(x)|\,\mathrm{d}x + C\int_1^{\infty}|\Phi(x,t)|\,\mathrm{d}x \\
&\le \|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{-k^2t})\|_{L^1} + C\mathrm{e}^{-2\delta^2 t}\int_1^{\infty}\frac{1}{x^2}\,\mathrm{d}x \\
&= \mathcal{O}(1)+\mathcal{O}(\mathrm{e}^{-2\delta^2 t}).
\end{aligned}
$$

It is obvious that the above estimates also hold for $\|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{-\mu k^2 t}\chi_0^h(k))\|_{L^1}$ for any arbitrary but fixed $\mu>0$.

Now we turn to the estimate of $\|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{\lambda_\varepsilon(k)t})\|_{L^1}$. We make the decomposition

$$
\begin{aligned}
\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{\lambda_\varepsilon(k)t}\chi_c^h(k)) = \mathcal{F}^{-1}(k\mapsto \mathrm{e}^{-\mu k^2 t}\chi_c^h(k)) \\
+\mathcal{F}^{-1}(k\mapsto (\mathrm{e}^{\lambda_\varepsilon(k)t}-\mathrm{e}^{-\mu k^2 t})\chi_c^h(k)),
\end{aligned}
$$

with $\mu=-\lambda_0''(0)/2>0$. We already showed that the first term is $\mathcal{O}(1)$ in the L^1 norm. For the second term we split the norm into three parts similar to above,

$$
\begin{aligned}
&\|\mathcal{F}^{-1}(k\mapsto (\mathrm{e}^{\lambda_\varepsilon(k)t}-\mathrm{e}^{-\mu k^2 t})\chi_0^h(k))\|_{L^1} \\
&= \underbrace{\int_{-\infty}^{-1} |\mathcal{F}^{-1}(k\mapsto (\mathrm{e}^{\lambda_\varepsilon(k)t}-\mathrm{e}^{-\mu k^2 t})\chi_0^h(k))(x)|\,\mathrm{d}x}_{=:\widetilde{I}_-(t)} + \underbrace{\int_{-1}^{1}|\dots|\,\mathrm{d}x}_{=:\widetilde{I}_0(t)} + \underbrace{\int_1^{\infty}|\dots|\,\mathrm{d}x}_{=:\widetilde{I}_+(t)}.
\end{aligned}
$$

The estimate for $\widetilde{I}_0(t)$ works as for $I_0(t)$ using the compact support of χ_0^h.

Then we consider the case $|x|\ge 1$. Using partial integration, we get

$$
\begin{aligned}
&\mathcal{F}^{-1}\big(k\mapsto \big(\mathrm{e}^{\lambda_\varepsilon(k)t}-\mathrm{e}^{-\mu k^2 t}\big)\chi_0^h(k)\big)(x) \\
&= -\frac{1}{x^2}\int_{-4\delta}^{4\delta}\big[\big(\mathrm{e}^{\lambda_\varepsilon(k)t}(\lambda_\varepsilon'(k))^2t^2-\mathrm{e}^{-\mu k^2 t}(4\mu^2k^2t^2)\big)\chi_0^h(k)
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\mathrm{e}^{\lambda_\varepsilon(k)t}\lambda_\varepsilon''(k)t-\mathrm{e}^{-\mu k^2 t}(-2\mu t)\right)\chi_0^h(k)\\
&+2\left(\mathrm{e}^{\lambda_\varepsilon(k)t}\lambda_\varepsilon'(k)t-\mathrm{e}^{-\mu k^2 t}(-2\mu k t)\right)(\chi_0^h)'(k)\\
&+\left(\mathrm{e}^{\lambda_\varepsilon(k)t}-\mathrm{e}^{-\mu k^2 t}\right)(\chi_0^h)''(k)\Big]\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k.
\end{aligned}
$$

The terms in the integral that are proportional to $(\chi_0^h)'$ or $(\chi_0^h)''$ have compact support bounded away from $k=0$. Hence, these terms can be bounded uniformly by $C\mathrm{e}^{-\widetilde{\mu}t}$ with some $\widetilde{\mu}>0$.

As an example, we show how to estimate the first term in the integral. The estimate for the second term works similar. Let $k\in[-4\delta,4\delta]$. Then there exists $\alpha\in(0,\mu)$ with $\lambda_\varepsilon(k)<-\alpha k^2$. Furthermore we have

$$
\begin{aligned}
\lambda_\varepsilon(k)&=-\mu k^2+k^4\varphi_\varepsilon(k),\\
\frac{(\lambda_\varepsilon'(k))^2}{4\mu^2k^2}&=1+k^2\widetilde{\varphi}_\varepsilon(k),
\end{aligned}
$$

where φ_ε, $\widetilde{\varphi}_\varepsilon$ are smooth mappings, defined on an open neighbourhood of $[-4\delta,4\delta]$ with $\|\varphi_\varepsilon\|_{C_b^0([-4\delta,4\delta])}$, $\|\widetilde{\varphi}_\varepsilon\|_{C_b^0([-4\delta,4\delta])}=\mathcal{O}(1)$. Hence, we have

$$
\begin{aligned}
&\left|\left(\mathrm{e}^{\lambda_\varepsilon(k)t}(\lambda_\varepsilon'(k))^2t^2-\mathrm{e}^{-\mu k^2 t}(4\mu^2k^2t^2)\right)\chi_0^h(k)\right|\\
&\le 4\mu^2k^2t^2\|\chi_0^h\|_{C_b^0}\left|\mathrm{e}^{\lambda_\varepsilon(k)t}\frac{(\lambda_\varepsilon'(k))^2}{4\mu^2k^2}-\mathrm{e}^{-\mu k^2 t}\right|\\
&\le Ck^2t^2\mathrm{e}^{-(\mu-\alpha)k^2t}\left(|\mathrm{e}^{(\alpha k^2+\lambda_\varepsilon(k))t}-\mathrm{e}^{-(\mu-\alpha)k^2t}|+k^2\|\widetilde{\varphi}_\varepsilon\|_{C_b^0([-4\delta,4\delta])}\mathrm{e}^{(\alpha k^2+\lambda_\varepsilon(k))t}\right)\\
&\le Ck^2t^2\mathrm{e}^{-(\mu-\alpha)k^2t}\left(\Big(\sup_{\xi\le 0}\mathrm{e}^{\xi}\Big)\cdot k^4\|\varphi_\varepsilon\|_{C_b^0([-4\delta,4\delta])}t+k^2\right)\\
&\le C(k^6t^3+k^4t^2)\mathrm{e}^{-(\mu-\alpha)k^2t}\le \sup_{K\in[0,\infty)}C(K^3+K^2)\mathrm{e}^{-K^2}=\mathcal{O}(1).
\end{aligned}
$$

From this we conclude that $\widetilde{I}_\pm(t)=\mathcal{O}(1)$ and thus $\|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{\lambda_\varepsilon(k)t}\chi_c^h(k))\|_{L^1}=\mathcal{O}(1)$ for $t\to\infty$. Similar computations show that $\|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{\lambda_\varepsilon(k)t}\chi_0^h(k)(\mathrm{i}k))\|_{L^1}=\mathcal{O}(t^{-1/2})$.

Since $\lambda_\varepsilon(k)-\varepsilon^2=-\frac{\lambda_0''(1)}{2}(k\mp 1)^2+\mathcal{O}(|k\mp 1|^4)$ for $k\to\pm 1$, and χ_c^h has compact support around ± 1, we can proceed as for $S_0(t)$ and show that

$$
\|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{\lambda_\varepsilon(k)t}\chi_c^h(k))\|_{L^1}=\mathrm{e}^{\varepsilon^2 t}\|\mathcal{F}^{-1}(k\mapsto \mathrm{e}^{(\lambda_\varepsilon(k)-\varepsilon^2)t}\chi_c^h(k))\|_{L^1}\le C\mathrm{e}^{\varepsilon^2 t}.
$$

The estimate for $S_s(t)$ follows in the same way as in [5, Lemma 10]. □

Now we are ready for the long-time control of the error. We start by reformulating the error system (3) with the help of the variation of constants formula. Note that $\mathrm{e}^{tL_\varepsilon}E_j=\mathrm{e}^{tL_\varepsilon}E_j^hE_j=S_j(t)E_j$. We obtain

$$R_c(t) = S_c(t)R_c(0) + \int_0^t S_c(t-\tau)\big(\varepsilon^2 N_c(\psi, R) + \varepsilon^3 g_c(\psi, R) + \varepsilon^{-5/2}\mathrm{Res}_c\big)(\tau)\,\mathrm{d}\tau,$$

$$R_s(t) = S_s(t)R_s(0) + \int_0^t S_s(t-\tau)\big(N_s(\psi_c, R_c) + \varepsilon^{1/2} g_s(\psi, R) + \varepsilon^{-4}\mathrm{Res}_c\big)(\tau)\,\mathrm{d}\tau,$$

$$R_0(t) = S_0(t)R_0(0) + \int_0^t S_0(t-\tau)\big(\varepsilon\partial_x N_0(\psi, R) + \varepsilon^{3/2}\partial_x g_0(\psi, R) + \varepsilon^{-3}\mathrm{Res}_0\big)(\tau)\,\mathrm{d}\tau.$$

Using the estimates from Lemma 10, Lemma 11, and (5), we get

$$\|R_c(t)\|_{H^m_{u,l}} \le C_{\mathrm{Res}} + \int_0^t C\varepsilon^2 \mathrm{e}^{\varepsilon^2(t-\tau)}\big(\mathcal{R} + \varepsilon(\mathcal{R} + \varepsilon^{1/2}\mathcal{R}^2)\big)(\tau)\,\mathrm{d}\tau, \tag{7}$$

$$\begin{aligned}\|R_s(t)\|_{H^m_{u,l}} \le C_{\mathrm{Res}} + \int_0^t C\mathrm{e}^{-\sigma(t-\tau)}(1 + (t-\tau)^{-1/6})\\ \times \big(\|R_c\|_{H^m_{u,l}} + \varepsilon^{1/2}(\mathcal{R} + \varepsilon^{1/2}\mathcal{R}^2)\big)(\tau)\,\mathrm{d}\tau,\end{aligned} \tag{8}$$

$$\|R_0(t)\|_{H^m_{u,l}} \le C_{\mathrm{Res}} + \int_0^t C\varepsilon(t-\tau)^{-1/2}\big(\|R_c\|_{H^m_{u,l}} + \varepsilon^{1/2}(\mathcal{R} + \varepsilon^{1/2}\mathcal{R}^2)\big)(\tau)\,\mathrm{d}\tau, \tag{9}$$

provided that $R_c(0)$, $R_s(0)$ and $R_c(0)$ are $\mathcal{O}(1)$.

Now, we set

$$q_j(t) = \sup_{\tau\in[0,t]} \|R_j(\tau)\|_{H^m_{u,l}}, \qquad j = 0, c, s,$$
$$q = q_0 + q_s.$$

We immediately obtain from (9) and (8) that

$$q(t) \le 2C_{\mathrm{Res}} + Cq_c(t) + C\varepsilon^{1/2}\big[q_c(t) + q(t) + \varepsilon^{1/2}(q_c(t) + q(t))^2\big].$$

If $\varepsilon^{1/2}C < 1/2$ and $\varepsilon^{1/2}(q_c(t) + q(t)) < 1/2$, then there exists a constant $C_q > 0$ with

$$q(t) < C_q + C_q q_c(t).$$

Plugging this into (7) yields

$$q_c(t) \le C + \int_0^t C\varepsilon^2 \mathrm{e}^{\varepsilon^2(t-\tau)}(q_c(\tau) + \varepsilon(q_c(\tau) + q_c(\tau)^2)\,\mathrm{d}\tau.$$

If $\varepsilon(1 + q_c(t)) \leq 1$, we get

$$q_c(t) \leq C + \int_0^t C\varepsilon^2 \mathrm{e}^{T_0} q_c(\tau)\,\mathrm{d}\tau.$$

Gronwall's inequality then gives

$$q_c(t) \leq C\mathrm{e}^{CT_0\mathrm{e}^{T_0}} =: M_q.$$

Hence, we see that the above estimates are all valid for every $t \in [0, T_0/\varepsilon^2]$, if we choose $\varepsilon_0 > 0$ so small that

$$\varepsilon_0(1 + 2M_q) < 1, \quad C\varepsilon_0^{1/2} < \frac{1}{2}, \quad \varepsilon_0^{1/2}(C_q + 2M_q(C_q + 1)) < \frac{1}{2}.$$

This concludes the proof of Theorem 7.

Remark 13. The fact that the Nikolaevskiy equation (1) possesses solutions that exist on the whole time interval $[0, T_0/\varepsilon^2]$, follows from the local well-posedness of (1) combined with the long-time existence of the approximation and the long-time estimates for the error we just proved. □

Acknowledgments

The author is grateful to Arjen Doelman and Vivi Rottschäfer for bringing the Nikolaevskiy equation to his attention during a stay at the Mathematisch Instituut in Leiden.

References

[1] I.A. Beresnev, V.N. Nikolaevskiy, A model for nonlinear seismic waves in a medium with instability, Phys. D 66 (1993) 1–6.
[2] D. Henry, Geometric Theory of Semilinear Parabolic Equations, Lecture Notes in Math., vol. 840, Springer-Verlag, Berlin–New York, 1981.
[3] T. Häcker, G. Schneider, D. Zimmermann, Justification of the Ginzburg–Landau approximation in case of marginally stable long waves, J. Nonlinear Sci. 21 (1) (2011) 93–113.
[4] P.C. Matthews, S.M. Cox, One-dimensional pattern formation with Galilean invariance near a stationary bifurcation, Phys. Rev. E 62 (2) (2000) R1473–R1476.
[5] G. Schneider, Error estimates for the Ginzburg–Landau approximation, Z. Angew. Math. Phys. 45 (3) (1994) 433–457.
[6] G. Schneider, Validity and limitation of the Newell–Whitehead equation, Math. Nachr. 176 (1995) 249–263.
[7] G. Schneider, Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances, J. Differential Equations 216 (2) (2005) 354–386.
[8] G. Schneider, D.A. Sunny, D. Zimmermann, The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions, J. Dynam. Differential Equations 27 (3) (2015) 1077–1099.
[9] G. Schneider, D. Zimmermann, Justification of the Ginzburg–Landau approximation for an instability as it appears for Marangoni convection, Math. Methods Appl. Sci. 36 (9) (2013) 1003–1013.
[10] D. Tanaka, Critical exponents of Nikolaevskii turbulence, Phys. Rev. E 71 (2005) 025203(R).
[11] R.W. Wittenberg, K.-F. Poon, Anomalous scaling on a spatiotemporally chaotic attractor, Phys. Rev. E 79 (2009) 056225.