Validity and attractivity of amplitude equations

Von der Fakultät Mathematik und Physik der Universität Stuttgart
zur Erlangung der Würde eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

Vorgelegt von

Kourosh Sanei Kashani

aus Teheran

Hauptberichter:	Prof. Dr. Guido Schneider
Mitberichter:	Priv.-Doz. Dr. Wolf-Patrick Düll
	Prof. Dr. Mariana Haragus
Tag der mündlichen Prüfung:	30. Oktober 2015

Institut für Analysis, Dynamik und Modellierung
der Universität Stuttgart

2016

Contents

Abstract

Whitham's equations and the Ginzburg-Landau equation belong to a set of famous amplitude equations containing the KdV equation, the NLS equation, Burgers equation, and so-called phase diffusion equations. They play an important role in the description of spatially extended dissipative or conservative physical systems. Except of Whitham's system for all other amplitude equations there exists a satisfying mathematical theory showing that the original system behaves approximately as predicted by the associated amplitude equation.

In the first part of this work we therefore derive Whitham's equations for a coupled system of equations, namely a Klein-Gordon-Boussinesq model. Subsequently we prove the validity of Whitham's equations for this system. The combination of our scaled ansatz adapted to Whitham's equations with the resonance structure of our system poses a new challenge. In order to prove the approximation results for Whitham's equations we will require some infinite series of normal transformations, for which we need to prove the convergence.

In the second part we prove the attractivity of the Ginzburg-Landau manifold for a toy problem inspired by Marangoni convection. In comparison to the previous classical situation in our case the curve of eigenvalues possesses additionally a marginally stable mode at the origin. Therefore, we will need to modify the requirements for the attractivity result and the method of proof.

Zusammenfassung

Whithams Gleichungen und die Ginzburg-Landau-Gleichung gehören zu einer Gruppe bekannter Amplitudengleichungen, zu der auch die KdV-Gleichung, die NLS-Gleichung, Burgers Gleichung und die sogenannte Phasendiffusionsgleichung zählen. Diese Gleichungen spielen eine wichtige Rolle bei der Beschreibung von räumlich ausgedehnten dissipativen oder konservativen physikalischen Systemen. Mit Ausnahme von Whithams Gleichungen existiert für alle andere Amplitudengleichungen bereits eine ausreichende mathematische Theorie, welche nachweist, dass das ursprüngliche System sich annähernd so verhält, wie es die dazugehörende Amplitudengleichung voraussagt.
Im ersten Teil der vorliegenden Arbeit leiten wir zunächst Whithams Gleichungen für ein gekoppeltes System von Gleichungen, nämlich für ein Klein-Gordon-Boussinesq-Modell, her. Im Anschluss beweisen wir die Gültigkeit von Whithams Approximation für dieses System. Die Kombination aus unserem Ansatz mit der Resonanzstruktur des verwendeten Systems stellt uns vor eine neue Herausforderung. Um jenes Aproximationsresultat für Whithams Gleichungen zu beweisen, werden wir eine unendliche Reihe von Normalformtransformationen benötigen, für welche die Konvergenz nachzuweisen ist.
Im zweiten Kapitel beweisen wir die Attraktivität der Ginzburg-Landau Mannigfaltigkeit am Beispiel eines Modellproblems, inspriert durch das Marangoni Problem. Im Vergleich zu den bisherigen klassischen Situationen haben wir in unserem Fall zusätzlich eine marginal stabile Mode im Ursprung vorliegen. Deswegen müssen hier die Anforderungen und die Beweistechniken für das genannte Attraktivitätsresultat entsprechend modifiziert werden.

Danksagung

Ich danke Dominik Zimmermann für seine technische und inhaltliche Unterstützung und die vielen nützlichen Gespräche. Ich bin James Kennedy dankbar für seine sprachliche Hilfe - er hat mir unermüdlich beim Verbessern meiner englischsprachigen Texte geholfen. Jörg Hörner hat mir immer wieder bei Latex hilfreich unter die Arme gegriffen. Wolf Patrick Düll hat mir Sachverhalte inhaltlich aufgeschlossen und mich dadurch fachlich unterstützt. Stefanie Siegert hat mich zuverlässig und sicher durch die Fahrwasser der Institution geleitet. Allen gilt mein Dank. Auch meiner Familie: Martina Volkmann danke ich für die jahrelange Wegbereitung, Rückenstärkung und die fürsorgliche Betreuung unserer Kinder, die ich bei ihr gut aufgehoben weiß. Meiner Schwester Percina danke ich, weil sie mich immer wieder ermutigt hat, meinen Weg zu gehen: Abitur zu machen, zu studieren - und sie hat mir immer die Sicherheit eines familiären Rückhaltes gegeben. Und schließlich: Professor Guido Schneider. Ihm verdanke ich sehr viel. Er hat als betreuender Professor für uns stets eine angenehme und inspirierende Arbeitsatmosphäre geschaffen. Er hat mir ermöglicht, an der Uni als wissenschaftlicher Assistent meinen Lebensunterhalt zu verdienen und meine Familie zu versorgen. Kompetent, geduldig und herzlich zugewandt hat er meine Promotionsarbeit betreut und stetig vorangebracht.

Chapter 1

Introduction

The most important objects in mathematical modeling are nonlinear partial differential equations. The description or understanding of natural or physical phenomena is often connected with complex nonlinear PDEs for which we cannot find any explicit solutions. Thus we need to apply other numerical or analytical methods in order to approximate solutions.
The analytical method that we use in this work to solve such complex mathematical problems has somehow a heuristic character. We assume the solutions that we are looking for have a certain form, which we call an "ansatz". Through a suitable scaling method with a small positive parameter ε, for example a slow modulation of time and space, together with a matching of the coeficients of different powers of ε, the original equation can be reduced to a new equation which is easy to solve and for which there is already a satisfactory mathematical theory. In this method the ansatz plays an important role and the results can be characterised through their ansatz and the associated reduced equations. In fact we approximate the exact solutions by the solutions of so called "universal" equations. In this work we will discuss several aspects of such universal equations, namely the validity of the Whitham's equations and attractivity of the Ginzburg-Landau manifold. By validity, we mean that the distance between the exact solution of the original system and the approximation based on the formally derived equation (in our case the Whitham's equations) is bounded over a long time interval ($\mathcal{O}(\varepsilon^{-1})$ in our case), see e.g. Theorem 2.1.1. In other words, we prove that the solutions of the approximating equation make a correct prediction about the behaviour of the solutions of the original system over a certain time scale.
In the first part of this work, Chapter 2, we begin by constructing some solutions for a coupled system, called a Klein-Gordon-Boussinesq model, or KGB model for

short. This model consists of two nonlinear equations, the Klein-Gordon equation and the Boussinesq equation:

$$\begin{aligned} &\partial_t^2 v = \partial_x^2 v - v + u^2 + 2uv + v^2 && \text{Klein-Gordon} \\ &\partial_t^2 u = \partial_x^2 u + \partial_t^2 \partial_x^2 u + \partial_x^2 (u^2 + 2uv + v^2) && \text{Boussinesq} \end{aligned}$$

The solution of the Klein-Gordon equation represents a quantum scalar field. The Boussinesq equation occurs in the context of the water wave problem. The question that arises here is why the coupling of such seemingly unrelated equations should be interesting for us. The reason is as follows: if we want to prove the validity of the Whitham's equations for systems with some kinds of periodicity, such as Polyatomic FPU systems or the water wave problem with a periodic bottom, we will encounter new difficulties. The method of proof used for the homogenous case cannot be applied in the periodic case due to the occurence of a certain type of resonance, which up until now has not been successfully treated by such a method. The KGB model has the same resonance structure (Figure 1.1) without sharing the periodicity of the other equations. Hence it can serve as a toy problem, for which we can more easily develop new technical tools and gain insights into the mathematical nature of the resonance, which is helpful for the other problems.

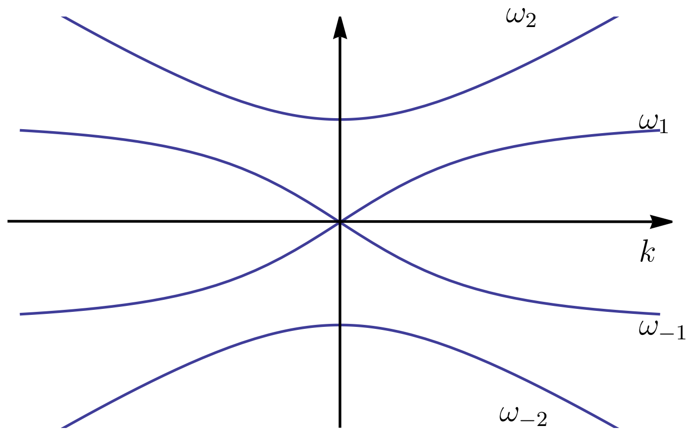

Figure 1.1: Curves of eigenvalues $\omega_{\pm 1}$ and $\omega_{\pm 2}$

If we start with an ansatz scaled with a small perturbating parameter $0 < \varepsilon \ll 1$ such as $\varepsilon^2 \psi_u(\varepsilon(x-t), \varepsilon^3 t)$ and $\psi_v = 0$, and equate the coefficient of ε^6 to zero, we will obtain the KDV equation [CS11]. Using the Ansatz $\varepsilon\psi(\varepsilon x, \varepsilon^2 t)$ yields the Burger equation. The ansatz that we use in this work is of order $\mathcal{O}(1)$ and has the form

$$\psi_u(\varepsilon x, \varepsilon t) \quad \text{and} \quad \psi_v(\varepsilon x, \varepsilon t).$$

This ansatz yields Whitham's equations in the coefficient of ε^2. Mathematically, Whitham's equations are a universal approximate equation for large classes of nonlinear PDEs of periodic wave type (see for example [DS09]).The resonance structure of our system (see Figure 1.1) is the same as in the situation in which one is really interested, namely the description of slow modulations in time and space of a periodic traveling wave in a dispersive wave system. By linearising around the periodic wave in a co-moving frame, we obtain an eigenvalue problem which is periodic in the spatial variable. Its solutions are given by Bloch modes $e^{\mathrm{i}lx+\mathrm{i}\omega_n(l)t}v_n(lx)$ with $n \in \mathbb{Z} \setminus \{0\}$, $l \in [\frac{-1}{2L}, \frac{1}{2L})$, and where the amplitude v_n possesses the same periodicity L w.r.t. x as the periodic wave. The curves $l \mapsto \omega_n(l)$ are ordered by $\omega_n(l) \leq \omega_{n+1}(l)$ and by $\omega_n(l) = -\omega_{-n}(l)$. In general we have $\omega_{\pm 1}(0) = 0$ and $\omega_{\pm 2}(0) \neq 0$ since in such systems the periodic wave is accompanied by an at least two-dimensional family of periodic waves. Whitham's equations describe the dynamics of the modes associated with the two curves $\omega_{\pm 1}$ in the long time limit, i.e. in the limit for $l \to 0$ (see Figure 1.1).

The most part of Chapter 2 is devoted to the proof of Theorem 2.1.1, which formulates the validity of Whitham's equations for the above-mentioned KGB model. Proving the validity of such approximations is a highly nontrivial task since the solutions, which are of order $\mathcal{O}(1)$, have to be shown to exist on a time scale of order $\mathcal{O}(\varepsilon^{-1})$. Our proof is based on two main analytical tools: normal form transformation, or NFT, and energy estimates. In general NFTs are of the form:

$$\tilde{R} = R + \delta M(\psi, R)$$

where R symbolises the former error function, $\tilde{R}$ the new error function, δ a scalar and M a suitably chosen bilinear mapping. In order to eliminate $\mathcal{O}(1)$ terms we set $\delta = 1$. The combination of such an ansatz with the resonance structure of our model (see Figure 1.1) necessarily leads to the application of an infinite sequence of NFTs. The convergence of the series resulting from this sequence of transformations is an important question and needs to be proved. In Section 2.3 we show that these series will converge like a geometrical series, and so with the help of the Neumann theorem we obtain the existence of an inverse. In order to complete the proof of Theorem 2.1.1, we prove the boundedness of a suitably chosen energy and then we apply Gronwall's Lemma.

In Chapter 3 we discuss the attractivity of the Ginzburg-Landau manifold for a toy problem inspired by Marangoni convection. The validity of the GL-approximation of this problem was recently proven in [SZ13]. After rescaling and renaming the variables, the Ginzburg-Landau equation looks as follows

$$\partial_T A = (1 + \mathrm{i}a)\partial_X^2 A + A - (1 + \mathrm{i}b)|A|^2|A|. \tag{1.1}$$

This equation arises in many branches of science as an approximating equation describing the evolution of patterns by instabilities and bifurcations. In order to derive the Ginzburg-Landau equation for a bifurcating solution, we apply the modulation

$$U(x,t) = \varepsilon A_1(\varepsilon x, \varepsilon^2 t)e^{\mathrm{i}x} + \varepsilon^2 A_2(\varepsilon x, \varepsilon^2 t)e^{2\mathrm{i}x} + \frac{\varepsilon^2}{2} A_0(\varepsilon x, \varepsilon^2 t) + c.c.. \tag{1.2}$$

Now let the GL-manifold be given by the following set:
$G = \{U$ as in (1.2) $\mid A_1$ satisfies GL's equation with initial condition in the set $B_\delta\}$, where $B_\delta := \{u_0 \mid \|u_0\|_i \leq \delta\}$ and $\|\cdot\|_i$ denotes a suitably chosen norm. In this context the local attractivity can be explained as follows: if we start with an initial perturbation p_0 with $\|p_0\|_i \leq \delta_0$, then the corresponding original solution u (i.e. the solution with initial condition $U(x,0) + p_0$) is well-defined, and after some transient behaviour for $0 \leq \varepsilon^2 t \leq T_0$ with an arbitrarily small T_0 can be approximated by (1.2) via

$$\|u - U\| \leq C\varepsilon^2 \quad \text{for} \quad T_0 \leq \varepsilon^2 t \leq T_1$$

with T_0, $T_1 = \mathcal{O}(1)$. Note that the norm $\|\cdot\|_i$ is not necessarily the same norm as $\|\cdot\|$. In this sense the size of the attractivity domain can be measured by δ_0. If we set $\|\cdot\|_i = \|\cdot\| = \sup_t \sup_x |\cdot|$, $\delta_0 = \mathcal{O}(\varepsilon^{3/2})$ and $(A_1, A_0) \in C([0,T_0], H^m(\mathbb{R}) \times H^{m-1}(\mathbb{R}))$ with $m > 11$ we would obtain the same approximation result as in [SZ13]. In this work we obtain estimates in L^1 for solution u for which u_0 satisfies an estimate of the form $\|u_0\rho_1\|_{L^1} \leq C$, where $1/\rho_1$ is of the form depicted in Figure 3.5.

Such an attractivity result has been established for the GL-manifold in the case of the Kuramoto-Shivashinsky equation in [BvHS95], where the associated eigenvalue curves possess two instability modes at $k = \pm 1$. In our case we have to consider and handle in addition a marginal stable mode at the origin (cf. Figure 3.1 and Figure 3.2). In order to show that the GL-manifold is an attractor for the solutions of our toy problem, we consider a comb-like family of weight functions $\rho_n(k)$ (sketched in Figure 1.2), which correspond to the bifurcating solutions of the (Fourier-transformed) GL-approximation. The so called "clustered mode-distribution" sketched in Figure 3.3 arises typically for pattern forming systems and looks more or less similar to our chosen weight functions. This clustered mode-distribution was introduced for the

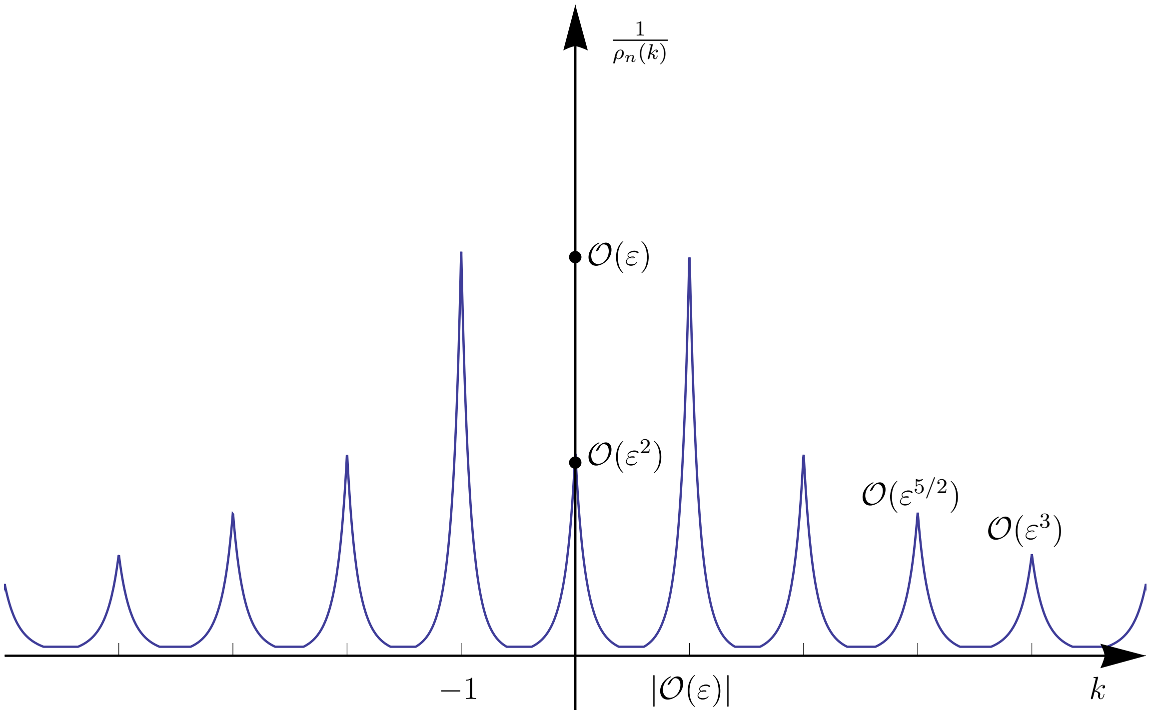

Figure 1.2: Sketch of the inverse of the weights ρ_n

first time 1971 by W. Eckhaus and R. C. DiPrima in [dRCES71]. The boundedness of the solution in Fourier space with respect to the L^1-norm with weights ρ_n for certain time will imply the above-mentioned attractivity.

Chapter 2

The validity of Whitham's approximation for a Klein-Gordon-Boussinesq model

The main application of Whitham's approximation is the description of slow modulations in time and space of periodic wave trains in general dispersive wave systems. In this chapter we prove the validity of Whitham's equations for a Boussinesq equation coupled with a Klein-Gordon equation. The proof is based on an infinite series of normal form transformation and an energy estimate. We expect that the steps pursued in this paper will be a part of a general approximation theory for Whitham's equations.

2.1 Introduction

We start with the formulation of the result followed by a longer discussion about the relevance of the result. We consider the system of partial differential equations

$$\partial_t^2 u = \partial_x^2 u + \partial_t^2 \partial_x^2 u + \partial_x^2 (u^2 + 2uv + v^2), \tag{2.1}$$

$$\partial_t^2 v = \partial_x^2 v - v + u^2 + 2uv + v^2, \tag{2.2}$$

with $u = u(x,t)$, $v = v(x,t)$, $x,\ t \in \mathbb{R}$. The solutions of linearised problem are given by $u(x,t) = e^{\mathrm{i}kx + \mathrm{i}\omega_{\pm 1}(k)t}$ and $v(x,t) = e^{\mathrm{i}kx + \mathrm{i}\omega_{\pm 2}(k)t}$ with

$$\omega_{\pm 1}(k)^2 = \frac{k^2}{k^2+1} \quad \text{and} \quad \omega_{\pm 2}(k)^2 = k^2 + 1. \tag{2.3}$$

We make the ansatz

$$\psi_u^{\mathrm{Whitham}}(x,t) = U(\varepsilon x, \varepsilon t) \quad \text{and} \quad \psi_v^{\mathrm{Whitham}}(x,t) = V(\varepsilon x, \varepsilon t) \tag{2.4}$$

with $0 < \varepsilon \ll 1$ a small perturbation parameter. Inserting this ansatz in (2.1) and (2.2) we find

$$\begin{aligned}
\mathrm{Res}_u &= -\partial_t^2 u + \partial_x^2 u + \partial_t^2 \partial_x^2 u + \partial_x^2 (u^2 + 2uv + v^2) \\
&= \varepsilon^2 (-\partial_T^2 U + \partial_X^2 U + \partial_X^2 (U^2 + 2UV + V^2)) + \varepsilon^4 \partial_T^2 \partial_X^2 U, \\
\mathrm{Res}_v &= -\partial_t^2 v + \partial_x^2 v - v + u^2 + 2uv + v^2 \\
&= -V + U^2 + 2UV + V^2 + \varepsilon^2 (-\partial_T^2 V + \partial_X^2 V).
\end{aligned}$$

Hence equating the coefficients of ε^0 in Res_v to zero yields

$$-V + U^2 + 2UV + V^2 = 0$$

and so $V = H(U) = U^2 + \mathcal{O}(U^3)$ due to the implicit function theorem for U and V of $\mathcal{O}(1)$, but sufficiently small. Equating the coefficients of ε^2 in Res_u to zero gives

$$-\partial_T^2 U + \partial_X^2 U + \partial_X^2 (U^2 + 2UV + V^2) = 0. \tag{2.5}$$

By substituting $V = H(U)$ into (2.5) we find

$$-\partial_T^2 U + \partial_X^2 U + \partial_X^2 (U^2 + 2UH(U) + H(U)^2) = 0. \tag{2.6}$$

It is the purpose of this paper to prove the following approximation result.

Theorem 2.1.1. *There exists a $C_1 > 0$ such that the following is true. Let $U \in C([0,T_0], H^6(\mathbb{R},\mathbb{R}))$ be a solution of (2.6) with $\sup_{T\in[0,T_0]} \|U(\cdot,T)\|_{H^6} \leq C_1$ and let $V = H(U)$. Then there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have solutions (u,v) of (2.1)-(2.2) such that*

$$\sup_{t\in[0,T_0/\varepsilon]} \sup_{x\in\mathbb{R}} |(u,v)(x,t) - (U,V)(\varepsilon x, \varepsilon t)| \leq C_2 \varepsilon^{3/2}.$$

Two questions emerge. Why should somebody be interested in such a result, and why is the proof of the result a real challenge? The reasons are as follows.

Remark 2.1.2. The scaling used in the ansatz (2.4) is the same scaling as it is used for the derivation of Whitham's equations. Whitham derived his equations first in [Whi65a, Whi65b], and they are still a subject of active research. They can be derived in various physical contexts in the description of modulations of periodic waves in nonlinear systems, cf. [DS09]. Very often they are derived from the Lagrangian of the underlying problem leading to a system of conservation laws, similar to (2.6) which can be rewritten in conservation law form as

$$\begin{aligned} \partial_T U &= \partial_X W, \\ \partial_T W &= \partial_X (U + U^2 + 2UH(U) + H(U)^2). \end{aligned}$$

Remark 2.1.3. There are many other amplitude equations like Whitham's equation, e.g. the Ginzburg-Landau equation, the KdV equation and the NLS equation. There exist already a series of approximation results for the Ginzburg-Landau approximation for instance in [CE90, vH91, Sch94a, Sch94b], for the KdV approximation for instance in [Cra85, SW00, SW02], and for the NLS approximation in [Kal87, Sch98, Sch05, BSTU06]. In all these cases the starting point of the multiple scaling analysis is the trivial spatially homogeneous solution of the system. If the starting point is a periodic traveling wave, the modulations of the wave in the dissipative case can be described by so-called phase diffusion equations, Burgers equation or conservation laws and in the conservative case by Whitham's equations. Approximation results in the dissipative case can be found in [MS04b, MS04a, DSSS09]. In the conservative case, i.e. for Whitham's equations, so far only one approximation result has been established, namely the validity of Whitham's equations for the NLS equation as original system which however has a much simpler resonance structure [DS09]. By the resonance structure we mean the situation that occurs for different curves of eigenvalue and their interaction, e.g. intersections (cf. 1.1).

In the following remark we explain why this resonance structure combined with the chosen scaling is a real challenge in establishing a suitable approximation result.

Remark 2.1.4. System (2.1)-(2.2) can be written as a first order system

$$\partial_t W = \Lambda W + B(W, W),$$

with Λ a linear skew symmetric operator and B a bilinear symmetric mapping. By adding higher order terms to the approximation (2.4) we construct an approximation ψ which is $\mathcal{O}(\varepsilon^2)$-close to (2.4) and satisfies formally

$$\text{Res}(\psi) = -\partial_t \psi + \Lambda \psi + B(\psi, \psi) = \mathcal{O}(\varepsilon^4).$$

The error function R defined by $W(x, t) = \psi(x, t) + \varepsilon^\beta R(x, t)$ fulfills

$$\partial_t R = \Lambda R + 2B(\psi, R) + \varepsilon^\beta B(R, R) + \varepsilon^{-\beta} \text{Res}(\psi).$$

We have to prove an $\mathcal{O}(1)$-bound for R on an $\mathcal{O}(\varepsilon^{-1})$-time scale. In order to do so we have to control the terms on the right hand side on this long time scale. The first term is skew-symmetric and will lead to oscillations without any growth rates. The last term can be $\mathcal{O}(\varepsilon)$-bounded if $\beta \leq 3$. If β is chosen larger than 1 the third terms gives a bound smaller than $\mathcal{O}(\varepsilon)$. However, the second term $2B(\psi, R)$ is only $\mathcal{O}(1)$-bounded. One approach to control this term is its elimination by a near-identity change of variables (normal form transformation abbreviated by NFT) $R = \widetilde{R} + M(\psi, \widetilde{R})$ with M being a suitably chosen bilinear mapping. The term $B(\psi, R)$ consists of a resonant and a non-resonant part, i.e.,

$$B(\psi, R) = B_r(\psi, R) + B_{nr}(\psi, R).$$

It turns out that only one part of this term can be eliminated. Since we can handle easily the resonant terms later with the help of a suitably chosen energy, our focus will be on eliminating the non-resonant terms. After an NFT the non-resonant term splits into a new resonant and non-resonant term again. Applying the transformation the relevant part of the equation for the new error function $\widetilde{R}$ is of the form

$$\partial_t \widetilde{R} = \Lambda \widetilde{R} + B_r(\psi, \widetilde{R}) + B(\psi, M(\psi, \widetilde{R})) + \mathcal{O}(\varepsilon).$$

Hence with the transformation new terms of $\mathcal{O}(1)$, namely $B(\psi, M(\psi, \widetilde{R}))$, appear. They can be split again into resonant and non-resonant terms. Another normal form transform is necessary to eliminate these non-resonant terms, but again terms of $\mathcal{O}(1)$ are created. However, they are cubic w.r.t. ψ. This goes ad infinitum and so the convergence of the composition of these infinitely many transformations has to be proven. Since the n-th transformation is of order $\mathcal{O}(\|\psi\|^n)$ the convergence finally can be established for $\|\psi\| = \mathcal{O}(1)$, but sufficiently small w.r.t. some $\|\cdot\|$-norm. After all these transformations the equation for the error takes the form

$$\partial_t R = \Lambda R + F(\psi, R) + \mathcal{O}(\varepsilon)$$

where F is a function which is linear w.r.t. R and which contains infinitely many resonant terms. Since all these terms have a long-wave character w.r.t. t (i.e. these terms depend from εt) a suitably chosen energy $E(R)$ satisfies

$$\partial_t E(R) = \mathcal{O}(\varepsilon),$$

and so an $\mathcal{O}(1)$-bound for the error R can be established on the $\mathcal{O}(\varepsilon^{-1})$-time scale. The normal form transformations can be found in Section 2.3 and the energy estimates in Section 2.5. The improved approximation is constructed in Section 2.2.

Remark 2.1.5. As explained above we think that our analysis is a necessary step for the validity of Whithams equations in the general situation. However, before applying these ideas a number of additional questions have to be answered, most essential: how to extract the wave numbers in non S^1-symmetric systems such that these satisfy equations which are suitable for existing functional analytic tools?

Remark 2.1.6. Recently Whitham's equations have been in the focus of investigations concerning modulations of periodic wave trains in dissipative systems containing conservation laws [JZ10]. The problems addressed in this work do not appear in the dissipative situation. We expect that the analysis for a justification result in the sense of Theorem 2.1.1 in the dissipative situation is very similar to the one given in [DSSS09, Section 6] where a single conservation law has been justified as an amplitude equation.

The rest of this chapter is dedicated to the proof of Theorem 2.1.1.

Notation. The many possible constants that are independent of $0 < \varepsilon \ll 1$ are denoted by C. The space $H^s(m)$ consists of s-times weakly differentiable functions for which $\|u\|_{H^s(m)} = \|u\rho^m(x)\|_{H^s} = (\sum_{j=0}^{s} \int |\partial_x^j (u\rho^m(x))|^2 dx)^{1/2}$ with $\rho(x) = \sqrt{1+x^2}$ is finite, where we do not distinguish between scalar and vector-valued functions or real- and complex-valued functions. We use H^s as an abbreviation for $H^s(0)$. From now we write $\int$ instead of $\int_{-\infty}^{\infty}$ hence the Fourier transform of a function u is denoted by

$$(\mathcal{F}u)(k) = \widehat{u}(k) = \frac{1}{2\pi} \int u(x) e^{-ikx} dx$$

and is an isomorphism between $H^s(m)$ and $H^m(s)$. The pointwise multiplication $(uv)(x) = u(x)v(x)$ in x-space corresponds to the convolution

$$(\widehat{u} * \widehat{v})(k) = \int \widehat{u}(k-l)\widehat{v}(l) dl$$

in Fourier space. The pseudo-differential operator $\omega(i\partial_x)$ in x-space is defined in Fourier space,

$$\omega(\mathrm{i}\partial_x) u(x) = \mathcal{F}^{-1}(\omega(k)\widehat{u}(k))(x),$$

where $\omega(k)$ is a piecewise analytic function.

2.2 The improved approximation and estimates for the residual

As explained in Remark 2.1.4 we need the residual to be small. With the approximation (2.4) we formally find that

$$\mathrm{Res}_u = \mathcal{O}(\varepsilon^4), \quad \text{but} \quad \mathrm{Res}_v = \mathcal{O}(\varepsilon^2).$$

In order to have $\mathrm{Res}_v = \mathcal{O}(\varepsilon^4)$ too, we improve the approximation (2.4) and make the improved ansatz

$$\psi_u(x,t) = U(\varepsilon x, \varepsilon t) \quad \text{and} \quad \psi_v(x,t) = V(\varepsilon x, \varepsilon t) + \varepsilon^2 V_2(\varepsilon x, \varepsilon t) \tag{2.7}$$

and find

$$\begin{aligned} \mathrm{Res}_v \;=\; & -V + U^2 + 2UV + V^2 + \varepsilon^2(-\partial_T^2 V + \partial_X^2 V - V_2 + 2UV_2 + 2VV_2) \\ & + \varepsilon^4(-\partial_T^2 V_2 + V_2^2 + \partial_X^2 V_2). \end{aligned}$$

We formally obtain $\mathrm{Res}_v = \mathcal{O}(\varepsilon^4)$ by choosing

$$V_2 = \frac{\partial_X^2 V - \partial_T^2 V}{1 - 2U - 2V}. \tag{2.8}$$

Since U and V could be choosen small enough V_2 is well-defined.

Remark 2.2.1. In the following we estimate the difference between a true solution of (2.1)-(2.2) and the improved ansatz (2.7). The estimate for the difference between a true solution of (2.1)-(2.2) and the original ansatz (2.4) then follows by the triangle inequality using

$$\sup_{t \in [0, T_0/\varepsilon]} \sup_{x \in \mathbb{R}} |(\psi_u, \psi_v)(x,t) - (U,V)(\varepsilon x, \varepsilon t)| \le C\varepsilon^2.$$

The difference between a true solution of (2.1)-(2.2) and the improved ansatz (2.7) defines the error functions R_u and R_v by

$$\varepsilon^\beta R_u = u - \psi_u \quad \text{and} \quad \varepsilon^\beta R_v = v - \psi_v$$

with a suitably chosen β. The error functions satisfy

$$\begin{aligned}\partial_t^2 R_u &= \partial_x^2 R_u + \partial_t^2\partial_x^2 R_u + \partial_x^2(2\psi_u R_u + 2\psi_v R_u + 2\psi_u R_v + 2\psi_v R_v) \\ &\quad + \varepsilon^{\beta}\partial_x^2(R_u^2 + 2R_uR_v + R_v^2) \\ &\qquad + \varepsilon^{-\beta}\underbrace{\left(-\partial_t^2\psi_u + \partial_x^2\psi_u + \partial_t^2\partial_x^2\psi_u + \partial_x^2(\psi_u^2 + 2\psi_u\psi_v + \psi_v^2)\right)}_{=\mathrm{Res}_u}\end{aligned} \tag{2.9}$$

$$\begin{aligned}\partial_t^2 R_v &= \partial_x^2 R_v - R_v + 2R_u\psi_u + 2R_u\psi_v + 2R_v\psi_u + 2R_v\psi_v \\ &\quad + \varepsilon^{\beta}R_u^2 + 2\varepsilon^{\beta}R_uR_v + \varepsilon^{\beta}R_v^2 \\ &\qquad + \varepsilon^{-\beta}\underbrace{(-\partial_t^2\psi_v + \partial_x^2\psi_v - \psi_v + \psi_u^2 + 2\psi_u\psi_v + \psi_v^2)}_{=\mathrm{Res}_v},\end{aligned} \tag{2.10}$$

where the residual terms are formally of order $\mathcal{O}(\varepsilon^4)$. These equations for the error functions will be solved in some Sobolev spaces. Estimating the residual terms in these Sobolev spaces will lose $\varepsilon^{-1/2}$ due to the scaling properties of the L^2-norm, namely

$$\left(\int |U(\varepsilon x)|^2 dx\right)^{1/2} = \left(\varepsilon^{-1}\int |U(X)|^2 dX\right)^{1/2}, \tag{2.11}$$

and so we have the following lemma.

Lemma 2.2.2. *Fix $s \geq 1$ and let $U \in C([0,T_0], H^{s+4}(\mathbb{R},\mathbb{R}))$ be a solution of* (2.6) *and $V = H(U)$. Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon]}(\|\mathrm{Res}_u\|_{H^s} + \|\mathrm{Res}_v\|_{H^s}) < C\varepsilon^{7/2}.$$

Proof. Combining the formal calculations from above with the scaling properties (2.11) of the L^2-norm yields the required estimates. In order to avoid losing more powers of ε in products arising in $\mathrm{Res}_{u,v}$ only one factor is estimated in H^s. All others are estimated in C_b^s. The assumption $U(.,T) \in H^{s+4}(\mathbb{R},\mathbb{R})$ is necessary to estimate $\partial_X^2 V_2 \in H^s(\mathbb{R},\mathbb{R})$ via $V_2 = \mathcal{O}(\partial_X^2 V)$ due to (2.8). $\square$

2.3 The series of normal transformations

In order to establish the validity of Theorem 2.1.1 we have to prove an $\mathcal{O}(1)$-bound for R_u and R_v on an $\mathcal{O}(\varepsilon^{-1})$ time scale. Therefore we need to control the terms on the right hand sides of (2.12) and (2.13) on this long time scale. The linear, ψ-independent terms are skew-symmetric and will lead to oscillations without any growth rates. The residual term can be $\mathcal{O}(\varepsilon)$-bounded if $\beta \leq 5/2$. Writing the error equations below as first order system decreases this number to $\beta \leq 3/2$. If β is chosen larger than 1 also the nonlinear terms gives a bound smaller than $\mathcal{O}(\varepsilon)$. Hence we will choose $\beta = 3/2$ in the following. However, the linear, ψ-dependent terms are only $\mathcal{O}(1)$-bounded. Therefore the biggest part of this work is devoted to the handling of these terms. The error equations are of the form

$$\partial_t^2 R_u = \partial_x^2 R_u + \partial_t^2\partial_x^2 R_u + \partial_x^2(2\psi_u R_u + 2\psi_v R_u + \underline{2\psi_u R_v + 2\psi_v R_v}) + p_u, \qquad (2.12)$$

$$\partial_t^2 R_v = \partial_x^2 R_v - R_v + \underline{2\psi_u R_u + 2\psi_v R_u} + 2\psi_u R_v + 2\psi_v R_v + p_v, \qquad (2.13)$$

where the terms p_v and p_u are defined by

$$p_u = \varepsilon^\beta \partial_x^2(R_u^2 + 2R_uR_v + R_v^2) + \varepsilon^{-\beta}\mathrm{Res}_u,$$
$$p_v = \varepsilon^\beta R_u^2 + 2\varepsilon^\beta R_uR_v + \varepsilon^\beta R_v^2 + \varepsilon^{-\beta}\mathrm{Res}_v$$

and they provide high enough orders w.r.t. ε such that they cause no difficulties in arriving at the $\mathcal{O}(\varepsilon^{-1})$ time scale. Their H^s-norm can be estimated with help of Lemma 2.2.2 by

$$\leq C(\varepsilon^{3/2}(\|R_u\|_{H^{s+2}} + \|R_v\|_{H^{s+2}})^2 + \varepsilon^2).$$

The terms that do cause difficulties are those in equations. (2.12)-(2.13) with no ε in front.

We start by eliminating the non-resonant terms, which are underlined. In order to do so we write (2.12)-(2.13) as a first order system, which in Fourier space has the form,

$$\begin{aligned}
\partial_t\widehat{R}_u &= \omega_1\widehat{W}_u \\
\partial_t\widehat{W}_u &= -\omega_1\widehat{R}_u - 2\omega_1(\widehat{\psi}_u * \widehat{R}_u + \widehat{\psi}_v * \widehat{R}_u + \underline{\widehat{\psi}_u * \widehat{R}_v + \widehat{\psi}_v * \widehat{R}_v}) + \varepsilon\check{p}_u \\
\partial_t\widehat{R}_v &= \omega_2\widehat{W}_v \\
\partial_t\widehat{W}_v &= -\omega_2\widehat{R}_v + 2\omega_2^{-1}(\underline{\widehat{\psi}_u * \widehat{R}_u + \widehat{\psi}_v * \widehat{R}_u} + \widehat{\psi}_u * \widehat{R}_v + \widehat{\psi}_v * \widehat{R}_v) + \varepsilon\check{p}_v.
\end{aligned}$$

where $\widehat{W}_u = \omega_1^{-1}\partial_t\widehat{R}_u$ and $\widehat{W}_v = \omega_2^{-1}\partial_t\widehat{R}_v$. The functions $\omega_1(k)$ and $\omega_2(k)$ are defined by (2.3). The $H^0(s)$-norm of the terms $\check{p}_u(k,t) = \varepsilon^{-1}\omega_1(k)^{-1}\frac{1}{k^2+1}\widehat{p}_u(k,t)$ and $\check{p}_v(k,t) = \varepsilon^{-1}\omega_2(k)^{-1}\widehat{p}_v(k,t)$, where $\widehat{p}_u$ and $\widehat{p}_v$ are the Fourier transform of p_u and p_v, can be estimated by

$$\leq C((\varepsilon^{1/2}(\|\widehat{R}_u\|_{H^0(s)} + \|\widehat{R}_v\|_{H^0(s)})^2 + 1).$$

The reasons are as follows. Since the nonlinear terms in (2.9) have two spatial derivatives in front, in Fourier space they are $\mathcal{O}(k^2)$, and so the application of $\omega_1(k)^{-1}$

is well-defined for all the terms containing $\widehat{R}_u$ and $\widehat{R}_v$ and for most terms from the residual. The terms which have to be dealt separately in the residual are time derivatives. They can be expressed via (2.6) as terms with spatial derivatives in front. Hence, in Fourier space all terms in $\widehat{p}_u(k,t)$ have at least a factor k and so the application of $\omega_1(k)^{-1}$ to these terms is well-defined. However, in the residual there is a loss of $\mathcal{O}(\varepsilon^{-1})$ since one derivative is canceled by the application of $\omega_1(k)^{-1}$. Such a loss does not occur in the linear and nonlinear terms w.r.t. $\widehat{R}_u$ and $\widehat{R}_v$ since their order w.r.t. ε purely comes from the amplitude and not from the long-wave character of the ansatz (2.4).

We diagonalize this system with

$$\begin{pmatrix} \widehat{R}_u \\ \widehat{W}_u \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \widehat{R}_1 \\ \widehat{R}_{-1} \end{pmatrix}, \begin{pmatrix} \widehat{R}_v \\ \widehat{W}_v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \widehat{R}_2 \\ \widehat{R}_{-2} \end{pmatrix}$$

and find

$$\begin{aligned} \partial_t \widehat{R}_1 &= \mathrm{i}\omega_1 \widehat{R}_1 - \mathrm{i}\omega_1 (\widehat{\psi}_u * (\widehat{R}_1 + \widehat{R}_{-1}) + \widehat{\psi}_v * (\widehat{R}_1 + \widehat{R}_{-1}) \\ &\quad + \underline{\widehat{\psi}_u * (\widehat{R}_2 + \widehat{R}_{-2})} + \underline{\widehat{\psi}_v * (\widehat{R}_2 + \widehat{R}_{-2})}) + \varepsilon \check{p}_1, \end{aligned} \tag{2.14}$$

$$\begin{aligned} \partial_t \widehat{R}_{-1} &= -\mathrm{i}\omega_1 \widehat{R}_{-1} + \mathrm{i}\omega_1 (\widehat{\psi}_u * (\widehat{R}_1 + \widehat{R}_{-1}) + \widehat{\psi}_v * (\widehat{R}_1 + \widehat{R}_{-1}) \\ &\quad + \underline{\widehat{\psi}_u * (\widehat{R}_2 + \widehat{R}_{-2})} + \underline{\widehat{\psi}_v * (\widehat{R}_2 + \widehat{R}_{-2})}) + \varepsilon \check{p}_{-1}, \end{aligned} \tag{2.15}$$

$$\begin{aligned} \partial_t \widehat{R}_2 &= \mathrm{i}\omega_2 \widehat{R}_2 - \mathrm{i}\omega_2^{-1} (\underline{\widehat{\psi}_u * (\widehat{R}_1 + \widehat{R}_{-1})} + \underline{\widehat{\psi}_v * (\widehat{R}_1 + \widehat{R}_{-1})} \\ &\quad + \widehat{\psi}_u * (\widehat{R}_2 + \widehat{R}_{-2}) + \widehat{\psi}_v * (\widehat{R}_2 + \widehat{R}_{-2})) + \varepsilon \check{p}_2, \end{aligned} \tag{2.16}$$

$$\begin{aligned} \partial_t \widehat{R}_{-2} &= -\mathrm{i}\omega_2 \widehat{R}_{-2} + \mathrm{i}\omega_2^{-1} (\underline{\widehat{\psi}_u * (\widehat{R}_1 + \widehat{R}_{-1})} + \underline{\widehat{\psi}_v * (\widehat{R}_1 + \widehat{R}_{-1})} \\ &\quad + \widehat{\psi}_u * (\widehat{R}_2 + \widehat{R}_{-2}) + \widehat{\psi}_v * (\widehat{R}_2 + \widehat{R}_{-2})) + \varepsilon \check{p}_{-2}. \end{aligned} \tag{2.17}$$

The $H^0(s)$-norm of the terms $\check{p}_{-2}, \ldots, \check{p}_2$ can be estimated by

$$\leq C((\varepsilon^{1/2}(\|\widehat{R}_{-2}\|_{H^0(s)} + \ldots + \|\widehat{R}_2\|_{H^0(s)})^2 + 1).$$

2.3.1 The iteration process

Since we have to perform infinitely many transformations in order to control the solutions of this system it is essential to extract its structure. Due to the symmetry of the system it is sufficient to consider the equations for $\widehat{R}_1$ and $\widehat{R}_2$. We write for the error functions

$$\begin{aligned}\partial_t \widehat{R}_{1j}(k,t) = \; & i\omega_1(k)\widehat{R}_{1,j}(k,t) + \int \widehat{f}^{(j)}_{1r}(k,k-m,\varepsilon t)\big(\widehat{R}_{1,j}(m,t)+\widehat{R}_{-1,j}(m,t)\big)dm \\ & + \int \widehat{f}^{(j)}_{1n}(k,k-m,\varepsilon t)\big(\widehat{R}_{2j}(m,t)+\widehat{R}_{-2,j}(m,t)\big)dm + \varepsilon\check{p}_{1,j}(k,t),\end{aligned} \tag{2.18}$$

$$\begin{aligned}\partial_t \widehat{R}_{2j}(k,t) = \; & i\omega_2(k)\widehat{R}_{2,j}(k,t) + \int \widehat{f}^{(j)}_{2n}(k,k-m,\varepsilon t)\big(\widehat{R}_{1,j}(m,t)+\widehat{R}_{-1,j}(m,t)\big)dm \\ & + \int \widehat{f}^{(j)}_{2r}(k,k-m,\varepsilon t)(\widehat{R}_{2j}(m,t)+\widehat{R}_{-2,j}(m,t)\big)dm + \varepsilon\check{p}_{2,j}(k,t).\end{aligned} \tag{2.19}$$

Remark 2.3.1. Note $\widehat{f}^{(j)}_{\cdot,\cdot}$ is proportional to the solutions $\widehat{\Psi}^j := (\widehat{\psi}_u + \widehat{\psi}_v)^j$ of Whitham's equations. And initially we have

$$\widehat{f}^{(1)}_{1r}(k,k-m,\varepsilon t) = \widehat{f}^{(1)}_{1n}(k,k-m,\varepsilon t) = -\mathrm{i}\omega_1(k)\widehat{\Psi}(k-m,\varepsilon t)$$

$$\widehat{f}^{(1)}_{2r}(k,k-m,\varepsilon t) = \widehat{f}^{(1)}_{2n}(k,k-m,\varepsilon t) = -\mathrm{i}\omega_2^{-1}(k)\widehat{\Psi}(k-m,\varepsilon t).$$

There are the same equations for $\widehat{R}_{-1,j}$ and $\widehat{R}_{-2,j}$, which are complex conjugates of the equations from above, i.e., especially $\widehat{f}^{(j)}_{-1r} = \overline{\widehat{f}^{(j)}_{1r}}$, $\widehat{f}^{(j)}_{-1n} = \overline{\widehat{f}^{(j)}_{1n}}$, $\widehat{f}^{(j)}_{-2r} = \overline{\widehat{f}^{(j)}_{2r}}$ and $\widehat{f}^{(j)}_{-2n} = \overline{\widehat{f}^{(j)}_{2n}}$ are valid.

The j-th near identity change of coordinates is given by

$$\widehat{R}_{1,j+1}(k,t) = \widehat{R}_{1,j}(k,t) + \sum_{l\in\{2,-2\}} \int \widehat{g}^{(j)}_{1l}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm, \tag{2.20}$$

$$\widehat{R}_{2,j+1}(k,t) = \widehat{R}_{2,j}(k,t) + \sum_{l\in\{1,-1\}} \int \widehat{g}^{(j)}_{2l}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm. \tag{2.21}$$

We find

$$\begin{aligned}\partial_t \widehat{R}_{1,j+1}(k,t) &= \partial_t \widehat{R}_{1,j}(k,t) + \sum_{l\in\{2,-2\}} \int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\partial_t \widehat{R}_{l,j}(m,t)dm \\ &\quad +\varepsilon \sum_{l\in\{2,-2\}} \int \partial_T \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm \\ &= \mathrm{i}\omega_1(k)\widehat{R}_{1,j}(k,t) + S_r + \widetilde{\widetilde{S}}_n + \varepsilon \check{p}_{1,j}(k,t) \\ &\quad +\varepsilon \sum_{l\in\{2,-2\}} \int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\check{p}_{l,j}(m,t)dm \\ &\quad +\varepsilon \sum_{l\in\{2,-2\}} \int \partial_T \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm\end{aligned}$$

where the resonant non-linearities S_r and non-resonant non-linearities $\widetilde{\widetilde{S}}_n$ are given by

$$\begin{aligned}S_r &= \int \widehat{f}_{1r}^{(j)}(k,k-m,\varepsilon t)\big(\widehat{R}_{1,j}(m,t)+\widehat{R}_{-1,j}(m,t)\big)dm \\ &\quad + \sum_{l\in\{2,-2\}} \int\int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{f}_{ln}^{(j)}(m,m-s,\varepsilon t)\big(\widehat{R}_{1,j}(s,t)+\widehat{R}_{-1,j}(s,t)\big)dsdm\end{aligned}$$

$$\begin{aligned}\widetilde{\widetilde{S}}_n &= \int \widehat{f}_{1n}^{(j)}(k,k-m,\varepsilon t)\big(\widehat{R}_{2,j}(m,t)+\widehat{R}_{-2,j}(m,t)\big)dm \\ &\quad + \sum_{l\in\{2,-2\}} \int \mathrm{i}\omega_l(m)\widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm \\ &\quad + \sum_{l\in\{2,-2\}} \int\int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{f}_{lr}^{(j)}(m,m-s,\varepsilon t)\big(\widehat{R}_{2,j}(s,t)+\widehat{R}_{-2,j}(s,t)\big)dsdm,\end{aligned}$$

where we have set $\omega_{-2}(m) := -\omega_2(m)$. Replacing $\widehat{R}_{1,j}$ by $\widehat{R}_{1,j+1}$ in the linear part with the help of (2.20) yields

$$\begin{aligned}\partial_t \widehat{R}_{1,j+1}(k,t) &= \mathrm{i}\omega_1(k)\widehat{R}_{1,j+1}(k,t) + S_r + \widetilde{S}_n + \varepsilon \check{p}_{1,j}(k,t) \\ &\quad +\varepsilon \sum_{l\in\{2,-2\}} \int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\check{p}_{l,j}(m,t)dm + \mathcal{O}(\varepsilon) \\ &\quad +\varepsilon \sum_{l\in\{2,-2\}} \int \partial_T \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm.\end{aligned}$$

Hence $\widetilde{S}_n$ could be written as

$$\widetilde{S}_n = \widetilde{\widetilde{S}}_n - \sum_{l\in\{2,-2\}} \int \mathrm{i}\omega_1(k)\widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm$$

In order to cancel the non-resonant terms in the non-linearities we set the first order part of $\widetilde{S}_n$ to zero as follows

$$\begin{aligned} 0 = &- \mathrm{i}\omega_1(k)\int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm \\ &+ \int \mathrm{i}\omega_l(m)\widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm \\ &+ \int \widehat{f}_{1n}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm, \end{aligned} \tag{2.22}$$

so we obtain

$$\begin{aligned} \partial_t\widehat{R}_{1,j+1}(k,t) = {} & \mathrm{i}\omega_1(k)\widehat{R}_{1,j+1}(k,t) + S_r + S_n + \varepsilon\check{p}_{1,j}(k,t) \\ &+ \varepsilon\sum_{l\in\{2,-2\}}\int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\check{p}_{l,j}(m,t)dm \\ &+ \varepsilon\sum_{l\in\{2,-2\}}\int \partial_T\widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{l,j}(m,t)dm \end{aligned} \tag{2.23}$$

and finally S_n takes the form

$$S_n = \sum_{l\in\{2,-2\}}\int\int \widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t)\widehat{f}_{lr}^{(j)}(m,m-s,\varepsilon t)\big(\widehat{R}_{2,j}(s,t)+\widehat{R}_{-2,j}(s,t)\big)dsdm.$$

Equation (2.22) leads to

$$\widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t) = (\mathrm{i}\omega_1(k)-\mathrm{i}\omega_l(m))^{-1}\widehat{f}_{1n}^{(j)}(k,k-m,\varepsilon t)$$

for $l\in\{2,-2\}$ and due to $\widehat{f}_{-1n}^{(j)} = \overline{\widehat{f}_{1n}^{(j)}}$ we get $\widehat{g}_{-1l}^{(j)} = \widehat{g}_{1-l}^{(j)}$. Analogous computations for $\widehat{g}_{2l}^{(j)}$ yield

$$\widehat{g}_{2l}^{(j)}(k,k-m,\varepsilon t) = (\mathrm{i}\omega_2(k)-\mathrm{i}\omega_l(m))^{-1}\widehat{f}_{2n}^{(j)}(k,k-m,\varepsilon t)$$

for $l\in\{1,-1\}$ and in the same way we have $\widehat{g}_{-2l}^{(j)} = \widehat{g}_{2-l}^{(j)}$. Due to the concentration of the $\widehat{\Psi}$ at zero we make only an $\mathcal{O}(\varepsilon^j)$ error if we choose

$$\widehat{g}_{1l}^{(j)}(k,k-m,\varepsilon t) = (\mathrm{i}\omega_1(k)-\mathrm{i}\omega_l(k))^{-1}\widehat{f}_{1n}^{(j)}(k,k-m,\varepsilon t) + \mathcal{O}(\varepsilon^j). \tag{2.24}$$

And the sum of all these errors will be of order $\mathcal{O}(\varepsilon)$ after infinitely many transformations since the underlying geometric series converges. The same statement is valid for the other possible indices of $\widehat{g}^{(j)}_{\cdot,\cdot}$. We assume for the moment that the transformation (2.20)- (2.21) is invertible and its inverse has the form

$$\widehat{R}_{i,j}(k,t) = \widehat{R}_{i,j+1}(k,t) + \sum_{l\in\{1,-1,2,-2\}} \int \widehat{h}^{(j)}_{il}(k,k-m,\varepsilon t)\widehat{R}_{l,j+1}(m,t)dm \tag{2.25}$$

with $i \in \{1,-1,2,-2\}$. Inserting (2.25) in (2.23) leads to

$$\begin{aligned}\partial_t \widehat{R}_{1,j+1}(k,t) =& \mathrm{i}\omega_1(k)\widehat{R}_{1,j+1}(k,t) + \int \sum_{\mu\in\{1,-1\}} \widehat{f}^{(j+1)}_{1r}(k,k-m,\varepsilon t)\widehat{R}_{\mu,j+1}(m,t)dm \\ &+ \int \sum_{\mu\in\{2,-2\}} \widehat{f}^{(j+1)}_{1n}(k,k-m,\varepsilon t)\widehat{R}_{\mu,j+1}(m,t)dm + \varepsilon\check{p}_{1,j+1}(k,t)\end{aligned} \tag{2.26}$$

where

$$\begin{aligned}\widehat{f}^{(j+1)}_{1r}(k,k-m,\varepsilon t) =& \widehat{f}^{(j)}_{1r}(k,k-m,\varepsilon t) + \sum_{\lambda\in\{2,-2\}} \int \widehat{g}^{(j)}_{1\lambda}(k,k-l,\varepsilon t)\widehat{f}^{(j)}_{\lambda n}(l,l-m,\varepsilon t)dl \\ &+ \int \widehat{f}^{(j)}_{1r}(k,k-l,\varepsilon t) \sum_{\kappa\in\{1,-1\}} \widehat{h}^{(j)}_{\kappa\mu}(l,l-m,\varepsilon t)dl \qquad (2.27) \\ &+ \sum_{\lambda\in\{2,-2\}} \int\int \widehat{g}^{(j)}_{1\lambda}(k,k-l_1,\varepsilon t)\widehat{f}^{(j)}_{\lambda n}(l_1,l_1-l_2,\varepsilon t) \\ &\times \sum_{\kappa\in\{1,-1\}} \widehat{h}^{(j)}_{\kappa\mu}(l_2,l_2-m,\varepsilon t)dl_2 dl_1 \\ &+ \sum_{\lambda\in\{2,-2\}} \int\int \widehat{g}^{(j)}_{1\lambda}(k,k-l_1,\varepsilon t)\widehat{f}^{(j)}_{\lambda r}(l_1,l_1-l_2,\varepsilon t) \\ &\times \sum_{\kappa\in\{2,-2\}} \widehat{h}^{(j)}_{\kappa\mu}(l_2,l_2-m,\varepsilon t)dl_2 dl_1,\end{aligned} \tag{2.28}$$

$$\begin{aligned}\widehat{f}_{1n}^{(j+1)}(k,k-m,\varepsilon t) &= \sum_{\lambda\in\{2,-2\}}\int \widehat{g}_{1\lambda}^{(j)}(k,k-l,\varepsilon t)\widehat{f}_{\lambda r}^{(j)}(l,l-m,\varepsilon t)dl \\ &+ \int \widehat{f}_{1r}^{(j)}(k,k-l,\varepsilon t)\sum_{\kappa\in\{1,-1\}}\widehat{h}_{\kappa\mu}^{(j)}(l,l-m,\varepsilon t)dl \\ &+ \sum_{\lambda\in\{2,-2\}}\int\int \widehat{g}_{1\lambda}^{(j)}(k,k-l_1,\varepsilon t)\widehat{f}_{\lambda n}^{(j)}(l_1,l_1-l_2,\varepsilon t) \\ &\times \sum_{\kappa\in\{1,-1\}}\widehat{h}_{\kappa\mu}^{(j)}(l_2,l_2-m,\varepsilon t)dl_2dl_1 \\ &+ \sum_{\lambda\in\{2,-2\}}\int\int \widehat{g}_{1\lambda}^{(j)}(k,k-l_1,\varepsilon t)\widehat{f}_{\lambda r}^{(j)}(l_1,l_1-l_2,\varepsilon t) \\ &\times \sum_{\kappa\in\{2,-2\}}\widehat{h}_{\kappa\mu}^{(j)}(l_2,l_2-m,\varepsilon t)dl_2dl_1,\end{aligned} \tag{2.29}$$

and

$$\begin{aligned}\check{p}_{1,j+1}(k,t) &= \sum_{\lambda\in\{2,-2\}}\int \partial_T\widehat{g}_{1\lambda}^{(j)}(k,k-m,\varepsilon t)\widehat{R}_{\lambda,j}(m,t)dm \\ &+ \check{p}_{1,j}(k,t) + \sum_{\lambda\in\{2,-2\}}\int \widehat{g}_{1\lambda}^{(j)}(k,k-m,\varepsilon t)\check{p}_{\lambda,j}(m,t)dm.\end{aligned} \tag{2.30}$$

Due to the symmetry in (2.18) and (2.19) we can obtain a similar equation for $\widehat{R}_2$ as in (2.26) but with the roles of $\widehat{R}_2$ and $\widehat{R}_1$ interchanged.

2.3.2 Functional analytic set-up and properties of the inverse transformation

In the following we use the notation

$$\|f\|_X := \int \sup_{k\in\mathbb{R}} |f(k,l)|(1+(l/\varepsilon)^2)^{s/2} dl. \tag{2.31}$$

Lemma 2.3.2. *Let $\|\cdot\|_X$ be defined as in* (2.31) *then the following estimate holds:*

$$\left\| \int_{\mathbb{R}} f(k,k-l)g(l,l-m)dl \right\|_X \leq \|f(k,l)\|_X \|g(k,l)\|_X.$$

Proof. We have

$$\begin{aligned}
&\int \sup_{k\in\mathbb{R}} \Big| \int f(k,k-l)g(l,l-m)dl \Big| (1+(m/\varepsilon)^2)^{s/2} dm \\
&\leq \int\int \sup_{k\in\mathbb{R}} |f(k,l)| \sup_{k\in\mathbb{R}} |g(k,l-m)| dl (1+(m/\varepsilon)^2)^{s/2} dm.
\end{aligned}$$

Young's inequality for convolutions in weighted L^p-spaces yields

$$\begin{aligned}
&\leq \int \sup_{k\in\mathbb{R}} |f(k,l)|(1+(l/\varepsilon)^2)^{s/2} dl \int \sup_{k\in\mathbb{R}} |g(k,m)|(1+(m/\varepsilon)^2)^{s/2} dm \\
&= \|f(k,l)\|_X \|g(k,l)\|_X.
\end{aligned}$$

□

Lemma 2.3.3. *Let $\widehat{R}_{i,j} \in L^2(s)$ with $s \geq 1$ be given and define $\widehat{R}_{i,j+1}$ by the transformation*

$$\widehat{R}_{j+1}(k,t) = (Id + T^{(j)})(\widehat{R}_j(k,t)), \tag{2.32}$$

where

$$T^{(j)}(\widehat{R}_j) = \begin{pmatrix} 0 & 0 & T_{12}^{(j)} & T_{1-2}^{(j)} \\ 0 & 0 & T_{-12}^{(j)} & T_{-1-2}^{(j)} \\ T_{21}^{(j)} & T_{2-1}^{(j)} & 0 & 0 \\ T_{-21}^{(j)} & T_{-2-1}^{(j)} & 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{R}_{1,j} \\ \widehat{R}_{-1,j} \\ \widehat{R}_{2,j} \\ \widehat{R}_{-2,j} \end{pmatrix} \tag{2.33}$$

with

$$\left(T_{il}^{(j)} \widehat{R}_{l,j}\right)(k,t) = \int \widehat{g}_{il}^{(j)}(k,k-m,\varepsilon t) \widehat{R}_{l,j}(m,t) dm.$$

Assume that there exists a $q > 0$ such that

$$\|\widehat{g}_{il}^{(j)}\|_X \leq q < \frac{1}{4} \tag{2.34}$$

holds for indices $i, l \in \{1, -1, 2, -2\}$. Then the transformation (2.32) is bijective and it has an inverse of the form

$$\widehat{R}_{i,j}(k,t) = \widehat{R}_{i,j+1}(k,t) + \sum_{l \in \{1,-1,2,-2\}} \int \widehat{h}_{il}^{(j)}(k, k-m, \varepsilon t) \widehat{R}_{l,j+1}(m,t) dm$$

with

$$\|\widehat{h}_{il}^{(j)}\|_X \leq \frac{\|\widehat{g}^{(j)}\|_X}{1 - \|\widehat{g}^{(j)}\|_X}, \tag{2.35}$$

where we have written $\|\widehat{g}^{(j)}\|_X = \max_{i,l \in \{-2,-1,1,2\}} \{\|\widehat{g}_{il}^{(j)}\|_X\}$.

Proof. It is clear that $L^2(s)$ is a Banach space and $T^{(j)} : \left(L^2(s)\right)^4 \to \left(L^2(s)\right)^4$ is a linear operator.
Let $T_1^{(j)}$ and $T_2^{(j)}$ be defined as follows

$$T_1^{(j)} = \begin{pmatrix} T_{12}^{(j)} & T_{1-2}^{(j)} \\ T_{-12}^{(j)} & T_{-1-2}^{(j)} \end{pmatrix} \qquad T_2^{(j)} = \begin{pmatrix} T_{21}^{(j)} & T_{2-1}^{(j)} \\ T_{-21}^{(j)} & T_{-2-1}^{(j)} \end{pmatrix},$$

and set $\|T^{(j)}\| := \max\{\|T_1^{(j)}\|, \|T_2^{(j)}\|\}$, where we use the norm

$$\|T_1^{(j)}\|^2 = \sup_{\|(\widehat{R}_{2,j}, \widehat{R}_{-2,j})^T\| \leq 1} \left(\sum_{l \in \{2,-2\}} \|T_{1l}^{(j)}(\widehat{R}_{l,j})\|^2_{L^2(s)} + \sum_{l \in \{2,-2\}} \|T_{-1l}^{(j)}(\widehat{R}_{l,j})\|^2_{L^2(s)} \right).$$

With the help of Young's inequality for convolutions we have

$$\|T_{il}^{(j)} \widehat{R}_{l,j}\|_{L^2(s)} \leq \|\widehat{g}_{il}^{(j)}\|_X \|\widehat{R}_{r,j}\|_{L^2(s)}.$$

Hence we have $\|T_1^{(j)}\| \leq 4q < 1$, for a $q < \frac{1}{4}$. In the same way we can show that $\|T_2^{(j)}\| \leq 4q < 1$ and hence $\|T^{(j)}\| \leq 4q < 1$ is valid. To justify the representation (2.25) we argue as follows. First we show the invertibility of the transformation from above. We use the Neumann series expansion for $\|T^{(j)}\| < 1$

$$(Id - (-T^{(j)}))^{-1} = \sum_{\lambda=0}^{\infty} \left(- T^{(j)} \right)^{\lambda}. \tag{2.36}$$

Secondly we prove that the series from (2.36) has the integral kernel form given by (2.25). By $\left(T^{(j)}\right)^{\lambda}$ we mean λ-times composition of $T^{(j)}$ or equivalently λ-times matrix product. Thus we obtain a composition of operators presented in Lemma 2.3.3. For each pair $T^{(j)}_{il}$ and $T^{(j)}_{st}$ we can write

$$\begin{aligned}(T^{(j)}_{il} \circ T^{(j)}_{st})\widehat{R}_{1,j+1} &= \int \widehat{g}^{(j)}_{il}(k, k-m) \int \widehat{g}^{(j)}_{st}(m, m-n)\widehat{R}_{1,j+1}(n)dndm \\ &= \int \underbrace{\int \widehat{g}^{(j)}_{il}(k, k-m) g^{(j)}_{st}(m, m-n)dm}_{=\widehat{h}(k,k-n)} \widehat{R}_{1,j+1}(n)dn.\end{aligned}$$

Hence we obtain inductively a series of integral kernels convolved with each error function $\widehat{R}_{i,j+1}$ as in (2.25). In this series we get for even λ non-resonant convolutions and for odd λ resonant convolutions with the error fuction $\widehat{R}_{i,j+1}$. The X-norm of $\widehat{h}^{(j)}_{ik}$ is bounded by

$$\|h^{(j)}_{ik}\|_X \leq \sum_{l=1}^{\infty} \left(\|g^{(j)}\|\right)^l = \frac{\|\widehat{g}^{(j)}\|_X}{1 - \|\widehat{g}^{(j)}\|_X}.$$

This is exactly (2.35).

□

2.3.3 The induction

In Lemma 2.3.3, we assumed that (2.34) holds. Here we wish to show that all the $\hat{f}^{(j)}_{\cdot,\cdot}$, $\hat{g}^{(j)}_{\cdot,\cdot}$ and $\hat{h}^{(j)}_{\cdot,\cdot}$ do in fact satisfy such estimates or even sharper estimates. First of all, using a simple application of Lemma 2.3.2 to (2.27) and (2.29) we get the following estimates

$$\begin{aligned}\|\widehat{f}_{1r}^{(j+1)}\|_X \le \|\widehat{f}_{1r}^{(j)}\|_X &+ \sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1\lambda}^{(j)}\|_X \|\widehat{f}_{\lambda n}^{(j)}\|_X + \|\widehat{f}_{1r}^{(j)}\|_X \sum_{\kappa\in\{1,-1\}} \|\widehat{h}_{\kappa\mu}^{(j)}\|_X \\ &+ \sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1\lambda}^{(j)}\|_X \|\widehat{f}_{\lambda n}^{(j)}\|_X \sum_{\kappa\in\{1,-1\}} \|\widehat{h}_{\kappa\mu}^{(j)}\|_X \\ &+ \sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1\lambda}^{(j)}\|_X \|\widehat{f}_{\lambda r}^{(j)}\|_X \sum_{\kappa\in\{2,-2\}} \|\widehat{h}_{\kappa\mu}^{(j)}\|_X \end{aligned} \tag{2.37}$$

and

$$\begin{aligned}\|\widehat{f}_{1n}^{(j+1)}\|_X \le &\sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1\lambda}^{(j)}\|_X \|\widehat{f}_{\lambda r}^{(j)}\|_X + \|\widehat{f}_{1r}^{(j)}\|_X \sum_{\kappa\in\{1,-1\}} \|\widehat{h}_{\kappa\mu}^{(j)}\|_X \\ &+ \sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1\lambda}^{(j)}\|_X \|\widehat{f}_{\lambda n}^{(j)}\|_X \sum_{\kappa\in\{1,-1\}} \|\widehat{h}_{\kappa\mu}^{(j)}\|_X \\ &+ \sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1\lambda}^{(j)}\|_X \|\widehat{f}_{\lambda r}^{(j)}\|_X \sum_{\kappa\in\{2,-2\}} \|\widehat{h}_{\kappa\mu}^{(j)}\|_X. \end{aligned} \tag{2.38}$$

One can obtain analogous inequalities for $\|\widehat{f}_{2r}^{(j+1)}\|_X$ and $\|\widehat{f}_{2n}^{(j+1)}\|_X$ with suitably adjusted indices. We will now give the important estimates mentioned earlier, from which our main convergence results will follow.

Lemma 2.3.4. *Let* $C_\omega = \max_{\mu\in\{1,-1\}} \max_{\lambda\in\{2,-2\}} \sup_{k\in\mathbb{R}} |\mathrm{i}\omega_\mu(k) - \mathrm{i}\omega_\lambda(k)|^{-1}$. *There exists a* $q < 1$ *with*

$$\|\widehat{f}_{\nu r}^{(1)}\|_X + \|\widehat{f}_{\nu n}^{(1)}\|_X = q, \quad \nu \in \{1,-1,2,-2\}$$

such that for all $j \in \mathbb{N}$

a) $\|\widehat{f}_{\kappa r}^{(j)}\|_X \le q\frac{1-q^{\frac{j}{2}}}{1-q^{\frac{1}{2}}}$

b) $\|\widehat{f}_{\kappa n}^{(j)}\|_X \le q^{\frac{j+1}{2}}$

c) $\|\widehat{g}_{\kappa\lambda}^{(j)}\|_X \le C_\omega q^{\frac{j+1}{2}}$

d) $\|\widehat{h}_{\kappa\lambda}^{(j)}\|_X \le 2C_\omega q^{\frac{j+1}{2}}$

with $\kappa, \lambda \in \{2,-2,1,-1\}$.

Proof. Each $\widehat{f}^{(j)}_{\cdot,\cdot}$ consists of powers of the solution $\widehat{\Psi}$ and we can choose for $s \leq 5$ the norm $\|\widehat{\Psi}\|_{L^1(s)}$ sufficiently small in particular there exists a $q > 0$ with $\|\widehat{f}^{(1)}_{ir}\|_X + \|\widehat{f}^{(1)}_{in}\|_X = q \ll 1$. For the proof of this theorem we proceed by induction.
Inductive basis: For $j = 1$ the estimates a) and b) follow from our choice of q. From (2.24) follows the assertion in c) for $j = 1$. From (2.35) we have

$$\|h^{(j)}_{ik}\|_X \leq \frac{\|\widehat{g}^{(j)}\|_X}{1 - \|\widehat{g}^{(j)}\|_X}.$$

Additionally let q be smaller than $\frac{1}{2C_\omega}$ then due to the inductive basis for c) we obtain the assertion in d) for $j = 1$.
Inductive step: We first obtain an estimate for the resonant term $\hat{f}^{(j+1)}_{1,r}$. For the ease of notation we define $\|\widehat{f}^{(j)}_n\|_X := \max_{\kappa \in \{1,-1,2,-2\}} \|\widehat{f}^{(j)}_{\kappa n}\|_X$. Using (2.24), (2.35) and (2.37) we write

$$\begin{aligned}\|\widehat{f}^{(j+1)}_{1r} - \widehat{f}^{(j)}_{1r}\|_X \leq \quad & 2C_\omega \|\widehat{f}^{(j)}_n\|^2_X + 4C_\omega \|\widehat{f}^{(j)}_{1r}\|_X \|\widehat{f}^{(j)}_n\|_X \\ & + 8C^2_\omega \|\widehat{f}^{(j)}_n\|^3_X + 8C^2_\omega \|\widehat{f}^{(j)}_n\|^2_X \|\widehat{f}^{(j)}_{2r}\|_X\end{aligned}$$

using the assumptions $\|\widehat{f}^{(j)}_{in}\|_X \leq q^{\frac{j+1}{2}}$ and $\|\widehat{f}^{(j)}_{ir}\|_X \leq q\frac{1-q^{\frac{j}{2}}}{1-q^{\frac{1}{2}}}$ we write

$$0 < \|\widehat{f}^{(j+1)}_{1r} - \widehat{f}^{(j)}_{1r}\|_X \leq 2C_\omega q^{j+1} + \frac{4C_\omega}{1 - q^{\frac{1}{2}}} q^{\frac{j+3}{2}} + 8C^2_\omega q^{\frac{3(j+1)}{2}} + \frac{8C^2_\omega}{1 - q^{\frac{1}{2}}} q^{j+2}.$$

On the right-hand side of the inequality the power of q is greater than $\frac{j}{2} + 1$, since q can be chosen sufficiently small, we can estimate the difference from above by

$$\|\widehat{f}^{(j+1)}_{1r} - \widehat{f}^{(j)}_{1r}\|_X \leq q q^{\frac{j}{2}}.$$

Summing over j, we obtain using the formula for a geometric series that

$$\|\widehat{f}^{(j+1)}_{1r}\|_X \leq q \sum_{k=0}^{j} q^{\frac{k}{2}} = q \frac{1 - q^{\frac{j+1}{2}}}{1 - q^{\frac{1}{2}}}.$$

Using (2.38) and (2.35) we can estimate $\|\widehat{f}^{(j)}_{in}\|_X$ similarly to the resonant terms as follows:

$$\|\widehat{f}^{(j+1)}_{in}\|_X \quad \leq \quad 6C_\omega q^{\frac{j+1}{2}} q \frac{1 - q^{\frac{j}{2}}}{1 - q^{\frac{1}{2}}} + 8C_\omega q^{\frac{3(j+1)}{2}} + 8C_\omega q^{j+2} \frac{1 - q^{\frac{j}{2}}}{1 - q^{\frac{1}{2}}}.$$

Notice that here in each summand we have a higher power than $q^{\frac{j+2}{2}}$, so that the estimate

$$\|\widehat{f}_{in}^{(j+1)}\|_X \leq q^{\frac{j+2}{2}}$$

is valid for small q. With the help of (2.35) we obtain the estimates from **c)** and **d)** directly.

□

Remark 2.3.5. The same approximation result from Lemma 2.3.4 is also valid for the partial derivatives of $\widehat{f}_{ir}^{(j)}(k,l,\varepsilon t)$, $\widehat{g}_{ir}^{(j)}(k,l,\varepsilon t)$ and $\widehat{h}_{ir}^{(j)}(k,l,\varepsilon t)$ w.r.t. the first component and the time component, because the functions $\omega_\lambda(k)$ with $\lambda \in \{2,-2,1,-1\}$ are continously differentiable and the underlying series for $\hat{h}_{\cdot,\cdot}^{(j)}$ converges uniformly. The proof works in an absolutely analogous way.

2.3.4 Control of the $\mathcal{O}(\varepsilon)$ terms of the equation (2.26)

In this paragraph we show that the sum of all summands involving the factor ε in (2.26) is bounded by a constant C throughout the whole transformation process. Using (2.30) and Lemma 2.3.2 we find the following estimate for the $H^0(s)$-norm of $\check{p}_{1,j+1}$.

$$\|\check{p}_{1,j+1}\|_{H^0(s)} \leq \sum_{\lambda\in\{2,-2\}} \|\partial_T \widehat{g}_{1,\lambda}^{(j)}\|_X \|\widehat{R}_{\lambda,j}\|_{H^0(s)} + \|\check{p}_{1,j}\|_{H^0(s)} + \sum_{\lambda\in\{2,-2\}} \|\widehat{g}_{1,\lambda}^{(j)}\|_X \|\check{p}_{\lambda,j}\|_{H^0(s)},$$

where $\check{p}_{1,1}(k,t)$ was introduced in (2.18), which can be traced back to the equation (2.10). Let $\|\check{p}_j\|_{H^0(s)} = \max_{\lambda\in\{1,2\}} \|\check{p}_{\lambda,j}\|_{H^0(s)}$ and $\|\widehat{R}_j\|_{H^0(s)} = \max_{\lambda\in\{1,2\}} \|\widehat{R}_{\lambda,j}\|_{H^0(s)}$. Applying Lemma 2.3.4 we obtain an additional estimate for the $H^0(s)$-norm of $\check{p}_{1,j+1}$

$$\|\check{p}_{1,j+1}\|_{H^0(s)} \leq 2C_\omega q^{\frac{j+1}{2}} \|\widehat{R}_j\|_{H^0(s)} + (1 + 2C_\omega q^{\frac{j+1}{2}})\|\check{p}_j\|_{H^0(s)}.$$

W.l.o.g. let $\|\check{p}_j\|_{H^0(s)} \geq \|\widehat{R}_j\|_{H^0(s)}$ for some j (otherwise the bound $\|\widehat{R}_j\|_{H^0(s)}$ already provides the desired control of these terms due to the factor of ε in front). Then the following inequality holds

$$\|\check{p}_{1,j+1}\|_{H^0(s)} \leq (1 + 4C_\omega q^{\frac{1}{2}}(q^{\frac{1}{2}})^j)\|\check{p}_j\|_{H^0(s)}.$$

We can choose q such that $4C_\omega q^{\frac{1}{2}} \leq 1$ and $q^{\frac{1}{2}} \leq \widetilde{q} < 1$. The $\mathcal{O}(\varepsilon)$ terms after N transformations can now be estimated by the product

$$\|\check{p}_N\|_{H^0(s)} \leq \Pi_{j=0}^N (1 + \widetilde{q}^j)\|\check{p}_1\|_{H^0(s)}.$$

We can represent the latter product using the logarithm as follows:

$$\Pi_{j=0}^{N}(1+\widetilde{q}^{j}) = e^{\sum_{j=0}^{N}\ln(1+\widetilde{q}^{j})}.$$

The exponent will converge like a geometric series since $\ln(1+\widetilde{q}^{j}) = \mathcal{O}(\widetilde{q}^{j})$ and $\widetilde{q} < 1$.

2.4 The transformed equations

After infinitely many transformations we can eliminate the non-resonant terms in (2.26) to arrive at

$$\begin{aligned}
\partial_t \widehat{R}_1(k,t) =& \mathrm{i}\omega_1(k)\widehat{R}_1(k,t) + \int \sum_{\mu\in\{1,-1\}} \widehat{f}_{1r}(k,k-m,\varepsilon t)\widehat{R}_\mu(m,t)dm + \varepsilon\check{p}_1,\\
\partial_t \widehat{R}_{-1}(k,t) =& -\mathrm{i}\omega_1(k)\widehat{R}_{-1}(k,t) - \int \sum_{\mu\in\{1,-1\}} \widehat{f}_{1r}(k,k-m,\varepsilon t)\widehat{R}_\mu(m,t)dm + \varepsilon\check{p}_1,\\
\partial_t \widehat{R}_2(k,t) =& \mathrm{i}\omega_2(k)\widehat{R}_2(k,t) + \int \sum_{\mu\in\{2,-2\}} \widehat{f}_{2r}(k,k-m,\varepsilon t)\widehat{R}_\mu(m,t)dm + \varepsilon\check{p}_2,\\
\partial_t \widehat{R}_{-2}(k,t) =& -\mathrm{i}\omega_2(k)\widehat{R}_{-2}(k,t) - \int \sum_{\mu\in\{2,-2\}} \widehat{f}_{2r}(k,k-m,\varepsilon t)\widehat{R}_\mu(m,t)dm + \varepsilon\check{p}_2,
\end{aligned}$$

where $\widehat{R}_\lambda$ and $\widehat{f}_{\lambda r}$ are limits of $\widehat{R}_{\lambda,j}$ and $\widehat{f}_{\lambda r}^{(j)}$ for $j\to\infty$, respectively. The $H^0(s)$-norm of the terms $\check{p}_1,\ldots,\check{p}_2$ is bounded by a constant C.

Remark 2.4.1. The sequence $\left(\widehat{f}_{1r}^{(j)}(k,k-m,\varepsilon t)\right)_{j\in\mathbb{N}}$ is a Cauchy sequence in X. From the completeness of $L^2(s)$ and the uniform convergence it follows that the limit $\widehat{f}_{\lambda r}$ is in $L^2(s)$ as well for $\lambda\in\{2,-2,1,-1\}$.

Undoing the diagonalisation yields

$$\begin{aligned}
\partial_t \widehat{R}_u(k,t) =& \omega_2(k)\widehat{W}_u(k,t)\\
\partial_t \widehat{W}_u(k,t) =& -\omega_2(k)\widehat{R}_u(k,t) + \int \widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)dm + \mathcal{O}(\varepsilon)\\
\partial_t \widehat{R}_v(k,t) =& \omega_1(k)\widehat{W}_v(k,t)\\
\partial_t \widehat{W}_v(k,t) =& -\omega_1(k)\widehat{R}_v(k,t) + \int \widehat{f}_v(k,k-m,\varepsilon t)\widehat{R}_v(m,t)dm + \mathcal{O}(\varepsilon),
\end{aligned} \qquad (2.39)$$

where

$$\widehat{f}_u(k,k-m,\varepsilon t) = 2\mathrm{i}\widehat{f}_{2r}(k,k-m,\varepsilon t) \in X,$$
$$\widehat{f}_v(k,k-m,\varepsilon t) = -2\mathrm{i}\widehat{f}_{1r}(k,k-m,\varepsilon t) \in X.$$

Now we want to demonstrate some useful properties in our nonlinearities.

Lemma 2.4.2. *The functions $\widehat{f}_w(k,k-m,\varepsilon t)$ with $w \in \{u,v\}$ satisfy the following properties:*

(i) $\widehat{f}_w(k,k-m,\varepsilon t) = \overline{\widehat{f}_w(k,m-k,\varepsilon t)}$,

(ii) $\|\widehat{f}_w(k,k-m,\varepsilon t) - \widehat{f}_w(m,k-m,\varepsilon t)\|_X = \mathcal{O}(\varepsilon)$.

Proof. We know that $\widehat{f}^{(1)}_{\lambda r}(k,k-m,\varepsilon t) = \mathrm{i}\omega_\lambda(k)\widehat{\Psi}(k-m,\varepsilon t)$, where $\omega_\lambda(k)$ and $\Psi(k-m,\varepsilon t)$ are real valued functions. Using the property $\widehat{\Psi}(k) = \overline{\widehat{\Psi}(-k)}$, we obtain

$$\widehat{f}^{(j)}_{\lambda r}(k,k-m,\varepsilon t) = -\overline{\widehat{f}^{(j)}_{\lambda r}(k,m-k,\varepsilon t)} \qquad (2.40)$$
$$\widehat{f}^{(j)}_{\lambda n}(k,k-m,\varepsilon t) = -\overline{\widehat{f}^{(j)}_{\lambda n}(k,m-k,\varepsilon t)}$$

for $j=1$. Using (2.24) we find that $\widehat{g}^{(1)}_{\lambda\mu}(k,k-m,\varepsilon t) = \overline{\widehat{g}^{(1)}_{\lambda\mu}(k,m-k,\varepsilon t)}$ and hence

$$\widehat{h}^{(1)}_{\lambda\mu}(k,k-m,\varepsilon t) = \overline{\widehat{h}^{(1)}_{\lambda\mu}(k,m-k,\varepsilon t)}.$$

With the help of (2.27) and a simple induction, it follows that (2.40) holds for all $j \in \mathbb{N}$. Hence, the assertion in *(i)* is valid.
We now prove (ii). We have

$$\|\widehat{f}_w(k,k-m,\varepsilon t) - \widehat{f}_w(m,k-m,\varepsilon t)\|_X \leq \| \sup_\xi \left|\partial_1 \widehat{f}_w(\xi,k-m)\right| \underbrace{(k-m)}_{=l} \|_{L^1}$$
$$= \int \sup_\xi \partial_1 \widehat{f}_w(\xi,l)(1+(l/\varepsilon)^2)^{s/2} \underbrace{\frac{l}{(1+(l/\varepsilon)^2)^{s/2}}}_{=\mathcal{O}(\varepsilon)} dl = \|\partial_1 \widehat{f}_w\|_X \mathcal{O}(\varepsilon).$$

From Lemma 2.3.4 and Remark 2.3.5 the assertion in *(ii)* follows.

□

Remark 2.4.3. The reason why the second property in Lemma 2.4.2 holds is the concentration of $\widehat{\psi}_n$ at the wave number $k=0$.

2.5 The final energy estimates

Now we use a technique from [CS11]; we will prove $\mathcal{O}(1)$-boundedness of a certain well-chosen energy, from which the desired approximation in Theorem 2.1.1 will follow. To that end let E_u and E_v be the energies defined by

$$E_u(t) = \int \omega_2(k)\widehat{R}_u(k,t)\overline{\widehat{R}}_u(k,t) + \omega_2(k)\widehat{W}_u(k,t)\overline{\widehat{W}}_u(k,t)dk \tag{2.41}$$

$$E_v(t) = \int \omega_1(k)\widehat{R}_v(k,t)\overline{\widehat{R}}_v(k,t) + \omega_1(k)\widehat{W}_v(k,t)\overline{\widehat{W}}_v(k,t)dk, \tag{2.42}$$

respectively. Since the linear terms cancel we have

$$\begin{aligned}
\frac{d}{dt}E_u(t) &= \int \omega_2(k)(\partial_t\widehat{R}_u(k,t))\overline{\widehat{R}_u}(k,t) + \omega_2(k)\widehat{R}_u(k,t)(\partial_t\overline{\widehat{R}_u}(k,t)) \\
&\quad +\omega_2(k)(\partial_t\widehat{W}_u(k,t))\overline{\widehat{W}_u}(k,t) + \omega_2(k)\widehat{W}_u(k,t)(\partial_t\overline{\widehat{W}_u}(k,t))dk \\
&= \int\int \omega_2(k)\overline{\widehat{W}_u}(k,t)\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)dmdk \\
&\quad + \int\int \omega_2(k)\widehat{W}_u(k,t)\overline{\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)}dmdk + \mathcal{O}(\varepsilon) \\
&= \int\int \partial_t\overline{\widehat{R}_u}(k,t)\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)dmdk \\
&\quad + \int\int \partial_t\widehat{R}_u(k,t)\overline{\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)}dmdk + \mathcal{O}(\varepsilon) \\
&= \int\int \partial_t\overline{\widehat{R}_u}(k,t)\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)dmdk \\
&\quad + \int\int \partial_t\widehat{R}_u(m,t)\overline{\widehat{f}_u(m,m-k,\varepsilon t)\widehat{R}_u(k,t)}dmdk + \mathcal{O}(\varepsilon) \\
&= \int\int \partial_t\overline{\widehat{R}_u}(k,t)\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)dmdk \\
&\quad + \int\int \partial_t\widehat{R}_u(m,t)\widehat{f}_u(m,k-m,\varepsilon t)\overline{\widehat{R}_u(k,t)}dmdk + \mathcal{O}(\varepsilon) \\
&= \int\int \partial_t\overline{\widehat{R}_u}(k,t)\widehat{f}_u(k,k-m,\varepsilon t)\widehat{R}_u(m,t)dmdk \\
&\quad + \int\int \partial_t\widehat{R}_u(m,t)\widehat{f}_u(k,k-m,\varepsilon t)\overline{\widehat{R}_u(k,t)}dmdk + \vartheta + \mathcal{O}(\varepsilon),
\end{aligned} \tag{2.43}$$

where ϑ is given by

$$\vartheta = \int\int \partial_t\widehat{R}_u(m,t)\big(\widehat{f}_u(k,k-m,\varepsilon t) - \widehat{f}_u(m,k-m,\varepsilon t)\big)\widehat{R}_u(k,t)dmdk.$$

Applying Hölder's inequality and Young's inequality we arrive at

$$\vartheta \leq \|\partial_t \widehat{R}_u\|_{L^2} \|\widehat{f}_u(k, k-m, \varepsilon t) - \widehat{f}_u(m, k-m, \varepsilon t)\|_{L^1} \|\widehat{R}_u\|_{L^2}.$$

With the help of Lemma 2.4.2 we obtain that $\vartheta = \mathcal{O}(\varepsilon)$ and equation (2.43) reads

$$\begin{aligned}
\frac{d}{dt} E_u(t) &= \int\int \partial_t(\overline{\widehat{R}_u}(k,t)\widehat{R}_u(m,t))\widehat{f}_u(k, k-m, \varepsilon t) dm dk + \mathcal{O}(\varepsilon) \\
&= \partial_t \left(\int\int \overline{\widehat{R}_u}(k,t)\widehat{R}_u(m,t)\widehat{f}_u(k, k-m, \varepsilon t) dm dk \right) + \mathcal{O}(\varepsilon).
\end{aligned}$$

As a consequence we have

$$\partial_t E_1 = \mathcal{O}(\varepsilon),$$

where

$$\begin{aligned}
E_1(t) = E_u(t) &- \int\int \overline{\widehat{R}_u}(k,t)\widehat{R}_u(m,t)\widehat{f}_u(k, k-m, \varepsilon t) dm dk \\
+E_v(t) &- \int\int \overline{\widehat{R}_v}(k,t)\widehat{R}_v(m,t)\widehat{f}_v(k, k-m, \varepsilon t) dm dk.
\end{aligned}$$

Finally a simple application of Gronwall's inequality yields the $\mathcal{O}(1)$-boundedness of E_1 for all $t \in [0, T_0/\varepsilon]$ for $\varepsilon > 0$ sufficiently small. Taking into account (2.41) we conclude the $\mathcal{O}(1)$-boundedness also for R_u and the boundedness for R_v can be concluded in an analogous way with the help of E_v defined in (2.42). Consequently we have the following estimate

$$\sup_{t \in [0, T_0/\varepsilon]} \| (u,v)(\cdot, t) - (U,V)(\cdot, t) \|_{H^s} = \varepsilon^{3/2} \sup_{t \in [0, T_0/\varepsilon]} \| R(\cdot, t) \|_{H^s} \leq C\varepsilon^{3/2}.$$

With the help of the embedding theorem for Sobolev spaces in Hölder spaces in [Alt06] we finally obtain the upper bound in Theorem 2.1.1. □

Chapter 3

Attractivity of the Ginzburg-Landau mode distribution for a pattern forming system with marginally stable long modes

Approximation and attractivity results are the basis of the classical Ginzburg-Landau theory which allowed to prove global existence results and upper semicontinuity of attractors towards the Ginzburg-Landau attractor for classical pattern forming systems like the Taylor-Couette problem close to the first instability. Recently, first approximation results for the Ginzburg-Landau approximation for pattern forming systems with marginally stable long modes, like the Bénard-Marangoni system, have been shown. It is the purpose of this chapter to prove the second fundamental property, namely the attractivity for such systems, too.

3.1 Introduction and result

The Ginzburg-Landau equation can be derived via multiple scaling analysis in order to describe slow modulations in time and space of the envelope of the most unstable modes of classical pattern forming systems, like the Taylor-Couette problem or Benard's problem, close to the first instability. It was derived first in [NW69, dRCES71] and first approximation results have been shown in [CE90, vH91, KSM92, Sch94b]. In [Eck93, Sch95, BvHS95] the attractivity of the set of solutions which can be

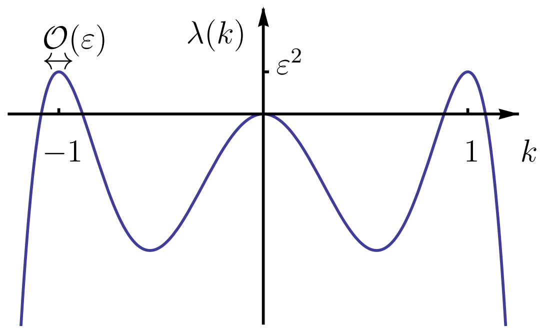

Figure 3.1: Curve of eigenvalues for the new situation. There is a touching point at wavenumber $k = 0$.

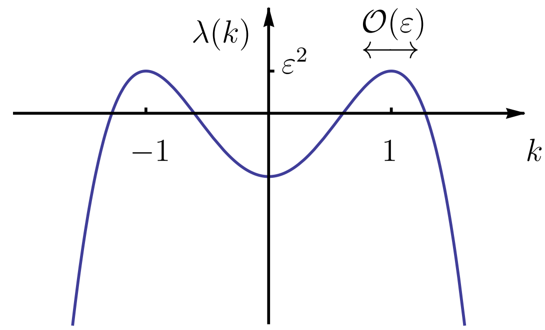

Figure 3.2: Curve of eigenvalues for the classical situation. The modes of the instabilities are of width $\mathcal{O}(\varepsilon)$.

described by the Ginzburg-Landau equation has been established. The approximation and attractivity property can be combined to establish global existence results [Sch94c, Sch99a] and the upper semicontinuity of the Ginzburg-Landau attractor [MS95, Sch99b]. Recently, first approximation results [HSZ11, SZ13] for the Ginzburg-Landau approximation for pattern forming systems with marginally stable long modes, like the Benard-Marangoni system, were shown. This situation is called in the following new situation. Here we will prove the second property, namely the attractivity for such systems. In [BvHS95] the local attractivity property has been shown in the case of the Kuramoto-Shivashinsky equation, for which the mode $k = 0$ is stable. The difference between the classical and the new situation can be seen in Figure 3.1 and Figure 3.2. In the new situation we have in addition to the critical modes a touching point at $k = 0$. As a consequence the quadratic interaction of the critical modes gives rise to modes which are no longer exponentially damped.

We refrain from considering the greatest possible degree of generality and restrict ourselves to a one-dimensional toy problem, namely

$$\partial_t u = \frac{1}{2}\partial_x^2(1-\partial_x^2)^2 u + \frac{1}{2}\varepsilon^2(\partial_x^6 - 3\partial_x^2)u + \partial_x^2(u^2), \tag{3.1}$$

with $x \in \mathbb{R}$, $t \geq 0$, $u(x,t) \in \mathbb{R}$, and $0 \leq \varepsilon \ll 1$. The solutions of the linearised problem are given by $e^{ikx+\lambda t}$, where

$$\lambda(k,\varepsilon) = -\frac{1}{2}k^2(1-k^2)^2 + \frac{\varepsilon^2}{2}(3k^2 - k^6). \tag{3.2}$$

The corresponding curves of eigenvalue, which are plotted in Figure 3.1, possess positive eigenvalues for $\varepsilon > 0$ in the intervals $|k \pm 1| \leq \mathcal{O}(\varepsilon)$. The instability occurs at a non-zero wave number, namely here for $k = \pm 1$. With the ansatz

$$U(x,t) = \varepsilon A_1(\varepsilon x, \varepsilon^2 t)e^{\mathrm{i}x} + \varepsilon^2 A_2(\varepsilon x, \varepsilon^2 t)e^{2\mathrm{i}x} + \frac{\varepsilon^2}{2}A_0(\varepsilon x, \varepsilon^2 t) + c.c. \tag{3.3}$$

we find

$$\begin{aligned} \partial_T A_1 &= 2\partial_X^2 A_1 + A_1 - 2(A_0 A_1 + A_2 A_{-1}), \\ \partial_T A_0 &= \tfrac{1}{2}\partial_X^2 A_0 + 2\partial_X^2 (A_1 A_{-1}), \\ 0 &= -18 A_2 - 4A_1^2, \end{aligned} \tag{3.4}$$

where $X = \varepsilon x$ and $T = \varepsilon^2 t$. Using the third equation of (3.4) yields the Ginzburg-Landau like system

$$\begin{aligned} \partial_T A_1 &= 2\partial_X^2 A_1 + A_1 - 2A_0 A_1 + \tfrac{4}{9}A_1|A_1|^2 \\ \partial_T A_0 &= \tfrac{1}{2}\partial_X^2 A_0 + 2\partial_X^2 (A_1 A_{-1}). \end{aligned} \tag{3.5}$$

Remark 3.1.1. The solutions of the GL-system A_1 and A_0 are chosen in (3.3) as the amplitude of the critical mode $k = 1$ and marginally stable mode $k = 0$ respectively. These critical modes give rise to the Ginzburg-Landau pattern.

Remark 3.1.2. In [SZ13] it was shown that if A_1 and A_0 with certain regularity solve the Ginzburg-Landau system (3.5) for $0 \leq \varepsilon^2 t \leq T$, then for all initial perturbations $\varepsilon^{3/2}u_0$ with $\sup_{x\in\mathbb{R}} |u_0| \leq C = \mathcal{O}(1)$, the corresponding $U(x,t)$ as in (3.3) makes correct predictions about the dynamics of our model (3.1) and satisfies following estimate:

$$\sup_{t\in[0,T/\varepsilon^2]} \sup_{x\in\mathbb{R}} |u(x,t) - U(x,t)| \leq C\varepsilon^{3/2}.$$

For the classical case an attractivity property has been shown in the sense that every small solution evolves in such a way that after a certain time it can be described by the Ginzburg-Landau approximation. We will establish such a result for (3.1), i.e., for the class of pattern forming systems with marginally stable long modes. Following the classical situation, cf. [Eck93, BvHS95, Sch95], we prove that after a time proportional to the natural scale of the Ginzburg-Landau approximation a Fourier mode distribution necessary for the derivation of the Ginzburg-Landau equation occurs. After this time one observes for the solution of (3.1) some strongly concentrated peaks at integer multiples of the critical wave number, which look more or less like in Figure 3.3. These peaks correspond to the Fourier mode distribution of the Ginzburg-Landau approximation.

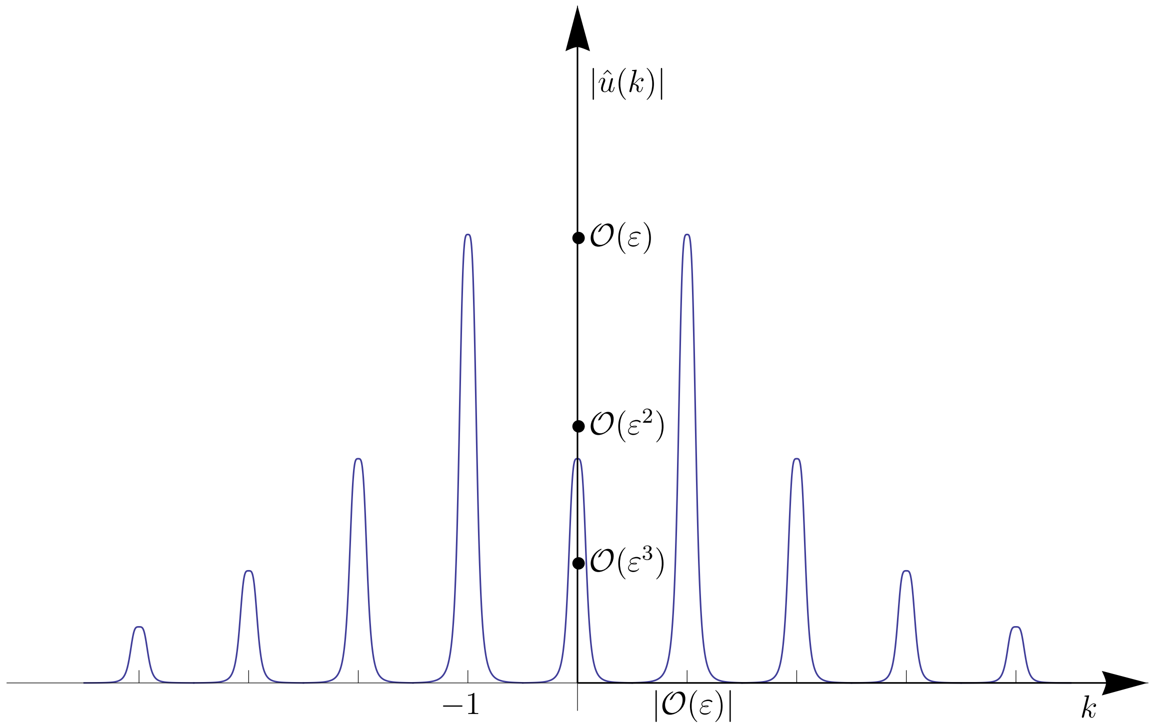

Figure 3.3: Peaks in the power spectrum corresponding to the structure of Ginzburg-Landau solutions

We begin with the Fourier transform given by

$$\widehat{u}(k) = \frac{1}{2\pi}\int_{\mathbb{R}} u(x)e^{-ikx}dx.$$

Writing (3.1) in Fourier space yields

$$\partial_t \widehat{u} = \lambda(k,\varepsilon)\widehat{u} - f(k)\widehat{u} * \widehat{u}, \tag{3.6}$$

where $f(k) = k^2$ and $*$ denotes convolution. In order to describe such a pattern as in Figure 3.3 we use a weighted L^1-norm, with the family of weights defined by

$$\begin{aligned} \frac{1}{\rho_n(k)} \quad = \quad & \varepsilon^{(n+3)/2} \max_{\substack{j=\pm1,\pm2;\\ l=\pm3,\cdots,\pm n}} \big\{2, 1/(\varepsilon+|k-j|)^{n/2-|j|+3/2}, \\ & 1/(\varepsilon+|k|)^{n/2-1/2}, 1/(\varepsilon+|k-l|)^{n/2-|l|/2+1/2}\big\}, \quad n\in\mathbb{N}. \end{aligned} \tag{3.7}$$

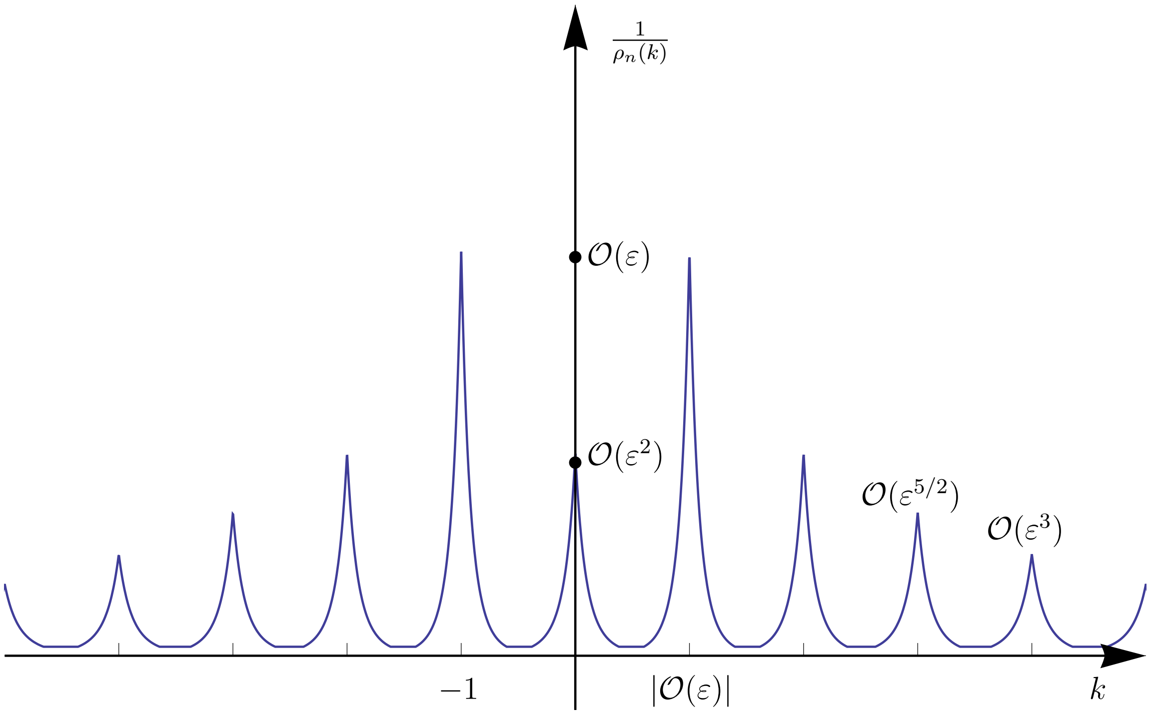

Figure 3.4: Sketch of the inverse of the weights ρ_n. In between the peaks the spectrum is of order $\mathcal{O}(\varepsilon^{(n+3)/2})$.

We see in Figure 3.4 that the plot of $k \to 1/\rho_n(k)$ looks more or less similar to Figure 3.3. Note that by construction the solution of (3.6) will feature such a pattern shown in Figure 3.3 if we show that the L^1-norm of the product $\widehat{u}(k,t)\rho_n(k)$ is bounded by a constant C. This is because with these weights peaks can only occur at integer multiples of critical mode $k_c = \pm 1$ and be of order $\mathcal{O}(\varepsilon^{\alpha})$ for $\alpha \geq 1$ as in Figure 3.4. Therefore in the next step we will prove the following theorem for the solution of (3.6):

Theorem 3.1.3. *Consider the solution of* (3.6) *with initial condition* $\widehat{u}_0(k)$ *satisfying*

$$\int_{\mathbb{R}} |\widehat{u}_0(k)|\rho_1(k)dk \leq C_0 \tag{3.8}$$

Then for all $C_0 > 0$, $n \in \mathbb{N}$ *there exist* C_n, T_n *and* $\varepsilon_0 > 0$ *such that for all* $\varepsilon \in (0, \varepsilon_0)$ *we have for the associated solution* $t \mapsto \widehat{u}(k,t)$ *with* $\widehat{u}(k,0) = \widehat{u}_0(k)$ *that*

$$\int_{\mathbb{R}} |\widehat{u}(k, T_n/\varepsilon^2)|\rho_n(k)dk \leq C_n.$$

Remark 3.1.4. Theorem 3.1.3 asserts that, beginning with the initial condition $\widehat{u}_0$ as in (3.8), the solution of (3.6) converges to the clustered mode-distribution corresponding to evolution of the Ginzburg Landau solution.

Remark 3.1.5. We follow the approach of [BvHS95] with a modification for the initial condition $\widehat{u}_0$. Due to the assumption (3.8) only two peaks of order $\mathcal{O}(\varepsilon)$ around critical modes are allowed, elsewhere the initial condition is at least of order $\mathcal{O}(\varepsilon^2)$ (cf. Figure 3.5). This restriction is neccessary because of the structure of the GL-solution at the wavenumber $k = 0$. In (3.3) we have ε^2 in front of A_0 and the curves of eigenvalue are only marginally stable at the mode $k = 0$ (see Figure 3.1), hence there, our solution should be initially $\mathcal{O}(\varepsilon^2)$ since there is no decay at $k = 0$.

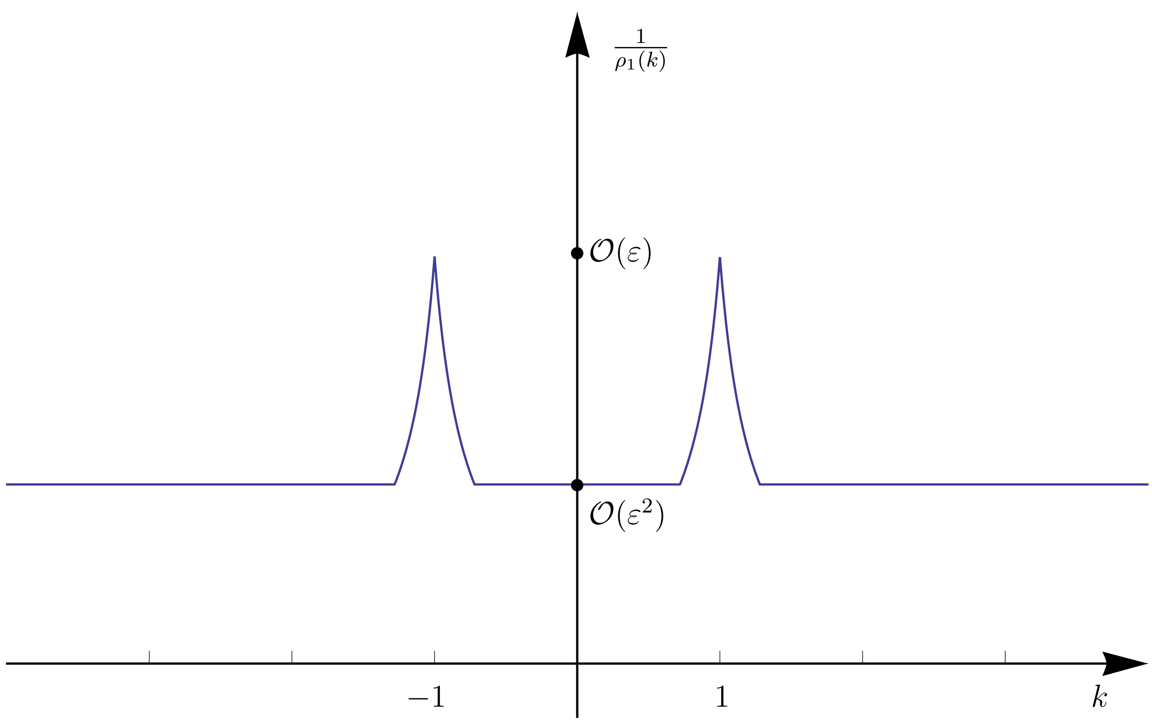

Figure 3.5: Sketch of the inverse of the weight ρ_1

3.2 Preparations

In order to control the convolution (i.e. the nonlinearity in physical space) we need to define a second family of weights ρ_n^* as follows:

$$\frac{1}{\rho_n^*(k)} = \varepsilon^{(n+5)/2} \max_{\substack{m=\pm 1;\\ l=\pm 2,\cdots,\pm n}} \{2\varepsilon^{-1/2}, 1/(\varepsilon+|k|)^{n/2+1/2}, 1/(\varepsilon+|k-m|)^{n/2-1/2}, 1/(\varepsilon+|k-l|)^{n/2-|l|/2+3/2}\}.$$

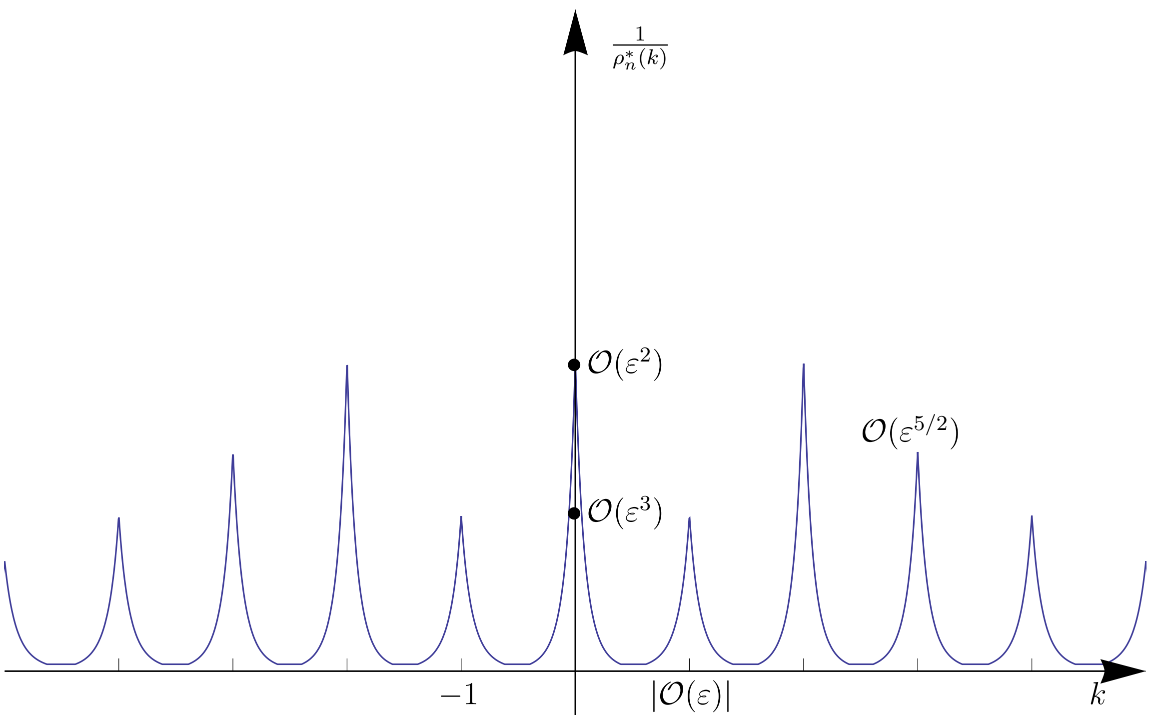

Figure 3.6: Sketch of the inverse of the weights ρ_n^*. In between the peaks is the spectrum of order $\mathcal{O}(\varepsilon^{(n+4)/2})$.

The plot of $k \to 1/\rho_n^*(k)$ can be found in Figure 3.6. We note that for an arbitrary integrable function $h \in L^1$, the mode structure and ε-powers of $\frac{h*h}{\rho_n^*(k)}$ are almost the same as those of $\frac{h}{\rho_n(k)}$ convoluted with itself. Since this is true for all such $h \in L^1$, we consider that, at least formally, $1 * 1/\rho_n^*(k)$ and $1/\rho_n(k) * 1/\rho_n(k)$ have the same mode structure and ε-powers, even though $1/\rho_n(k)$ are technically speaking not integrable. Due to the invariance of the distribution under convolution, the peaks of $\frac{1}{\rho_n^*(k)}$ again appear at integer multiples of the critical wave number. Since we have to treat the critical modes, the noncritical mode and the mode at the origin separately, we define the so-called "mode filters" E_c, E_b and E_0. These mode filters are defined as characteristic functions on the intervals $I_c = [-5/4, -3/4] \cup [3/4, 5/4]$, $I_0 = [-1/4, 1/4]$ and I_b as the complement of $I_c \cup I_0$, respectively. In addition we need to define ρ_n^c, ρ_n^b, ρ_n^0, $\rho_n^{c,*}$, $\rho_n^{b,*}$ and $\rho_n^{0,*}$. In mathematical terms they satisfy:

$$E_i(k) = \begin{cases} 1, & k \in I_i \\ 0, & k \notin I_i \end{cases}, \quad \rho_n^i = E_i \rho_n \quad \text{and} \quad \rho_n^{i,*} = E_i \rho_n^* \quad \text{with} \quad i \in \{c, b, 0\}.$$

A simple calculation shows that the weights have the following properties:

$$\begin{aligned}
&\rho_n^{c,*}(k+l) \leq C\rho_n^c(k)\rho_n^b(l), && \rho_n^{c,*}(k+l) \leq C\rho_n^c(k)\rho_n^0(l), \\
&\rho_n^{c,*}(k+l) \leq C\varepsilon\rho_n^b(k)\rho_n^b(l), && \rho_n^{c,*}(k+l) \leq C\varepsilon\rho_n^b(k)\rho_n^0(l), \\
&\rho_n^{b,*}(k+l) \leq C\varepsilon\rho_n^b(k)\rho_n^b(l), && \rho_n^{b,*}(k+l) \leq C\varepsilon^{1/2}\rho_n^c(k)\rho_n^b(l), \\
&\rho_n^{b,*}(k+l) \leq C\varepsilon\rho_n^b(k)\rho_n^0(l), && \rho_n^{b,*}(k+l) \leq C\rho_n^c(k)\rho_n^c(l), \\
&\rho_n^{b,*}(k+l) \leq C\varepsilon^{1/2}\rho_n^c(k)\rho_n^0(l), && \rho_n^{b,*}(k+l) \leq C\varepsilon^{3/2}\rho_n^0(k)\rho_n^0(l), \\
&\rho_n^{0,*}(k+l) \leq C\varepsilon\rho_n^0(k)\rho_n^0(l), && \rho_n^{0,*}(k+l) \leq C\rho_n^c(k)\rho_n^c(l), \\
&\rho_n^{0,*}(k+l) \leq C\rho_n^c(k)\rho_n^b(l), && \rho_n^{0,*}(k+l) \leq C\varepsilon\rho_n^b(k)\rho_n^b(l), \\
&\rho_n^{0,*}(k+l) \leq C\varepsilon\rho_n^0(k)\rho_n^b(l)
\end{aligned} \tag{3.9}$$

The inequalities (3.9) are valid for all k and l in the indexed intervals, e.g. the inequality $\rho_n^{c,*}(k+l) \leq C\rho_n^c(k)\rho_n^0(l)$ holds for all $k \in I_c$, $l \in I_0$ and $k+l \in I_c$. The other index combinations do not contribute to the corresponding convolutions. The convolution of the critical modes gives noncritical modes and this property is one of the essential points in the proof of Theorem 3.1.3. In the following we demonstrate in detail how the estimates in (3.9) are obtained. We do this for a few selected indices, namely for $\rho_n^{0,*}(k+l) \leq C\rho_n^c(k)\rho_n^b(l)$ and $\rho_n^{0,*}(k+l) \leq C\varepsilon\rho_n^0(k)\rho_n^0(l)$, since the other cases are similar.
Let $k+l \in I_0$ then we assume $k \in I_c$ and thus

$$l \in [-3/2, -5/4] \cup [-3/4, -1/2] \cup [1/2, 3/4] \cup [5/4, 3/2].$$

Then we have

$$\begin{aligned}
\rho_n^{0,*}(k+l) &= \varepsilon^{-\frac{n+5}{2}}\left(\varepsilon+|k+l|\right)^{\frac{n+1}{2}} = \varepsilon^{-2}\left(1+\frac{|k+l|}{\varepsilon}\right)^{\frac{n+1}{2}} \\
&\le \varepsilon^{-2}\max_{j=\pm 1}\left(1+\frac{|k-j|}{\varepsilon}\right)^{\frac{n+1}{2}}\left(1+\frac{|l+j|}{\varepsilon}\right)^{\frac{n+1}{2}} \\
&\le C\rho_n^c(k)\varepsilon^{-1}\underbrace{\left(1+\frac{|l+j|}{\varepsilon}\right)^{\frac{n+1}{2}}}_{=\mathcal{O}(\varepsilon^{-\frac{n+1}{2}})} \\
&\le C\rho_n^c(k)\rho_n^b(l),
\end{aligned}$$

where we have used the inequality $(1+a+b)^p \le (1+a)^p(1+b)^p$ for all $a, b, p \ge 0$. For the second estimate we assume $k, l \in I_0$ and write

$$\begin{aligned}
\rho_n^{0,*}(k+l) &= \varepsilon^{-\frac{n+5}{2}}\left(\varepsilon+|k+l|\right)\left(\varepsilon+|k+l|\right)^{\frac{n-1}{2}} \\
&= \varepsilon^{-3}\underbrace{\left(\varepsilon+|k+l|\right)}_{\le\mathcal{O}(1)}\left(1+\frac{|k+l|}{\varepsilon}\right)^{\frac{n-1}{2}} \\
&\le C\varepsilon\varepsilon^{-2}\left(1+\frac{|k|}{\varepsilon}\right)^{\frac{n-1}{2}}\varepsilon^{-2}\left(1+\frac{|l|}{\varepsilon}\right)^{\frac{n-1}{2}} \\
&\le C\varepsilon\rho_n^0(k)\rho_n^0(l).
\end{aligned}$$

This proves the two representative cases from (3.9) that we wished to show.

3.3 Proof of attractivity

The proof of Theorem 3.1.3 is based on induction. We begin directly with the inductive step, i.e., we show that $\|\widehat{u}(\cdot, T_n/\varepsilon^2)\rho_n\|_{L^1} < C_n$ implies the estimate

$$\|\widehat{u}(\cdot, T_{n+1}/\varepsilon^2)\rho_{n+1}\|_{L^1} \le C_{n+1}$$

for a $T_{n+1} > T_n$. In order to do so we determine a time $\widetilde{T}_n$ such that

$$\sup_{0<t<T_{n+1}/\varepsilon^2}\|\widehat{u}(\cdot, t)\rho_n\|_{L^1} < C_n$$

with $T_{n+1} = T_n + \widetilde{T}_n$, and then we show the boundedness of the solution with the new weight ρ_{n+1} i.e. $\sup_{T_n/\varepsilon^2<t<T_{n+1}/\varepsilon^2}\|\widehat{u}(\cdot, t)\rho_{n+1}\|_{L^1} \le C_{n+1}$. We consider the following lemma:

Lemma 3.3.1. *(Induction step) Let $\widehat{u}(k,t)$ be a solution of (3.6) as in Theorem 3.1.3 and $\rho_n(k)$ be defined as in (3.4). Then for all C_n there exist C_{n+1}, $T_{n+1} > T_n > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the implication*

$$\|\widehat{u}(k, T_n/\varepsilon^2)\rho_n(k)\|_{L^1} \leq C_n \Longrightarrow \|\widehat{u}(k, T_{n+1}/\varepsilon^2)\rho_{n+1}(k)\|_{L^1} \leq C_{n+1} \tag{3.10}$$

is valid.

In Lemma 3.3.1 the constants C_n are independent of the choice of ε. The base clause as we will see later follows in an analogous fashion using (3.8). The next section is completely devoted to the proof of Lemma 3.3.1 and hence the proof of Theorem 3.1.3

We will use the same letter C to denote constants which may vary from line to line, but are always independent of ε. Here the notation $\widehat{u}_n(\cdot, 0) := \widehat{u}(\cdot, T_n/\varepsilon^2)$ will be useful.

I) First, we show that there exist $\widetilde{C}_n$, $\widetilde{T}_n$ and $\varepsilon_0 > 0$ such that

$$\sup_{0<t<\widetilde{T}_n/\varepsilon^2} \|\widehat{u}_n(t)\rho_n\|_{L^1} \leq \widetilde{C}_n$$

for all $0 < \varepsilon < \varepsilon_0$. For this purpose we split the estimate with the help of E_c, E_b and E_0 into three parts. We set $M_i(t) = \sup_{0\leq s\leq t} \|E_i\widehat{u}_n(s)\rho_n\|_{L^1}$ with $i \in \{c, b, 0\}$ and $\|\widehat{u}_n(0)\rho_n\|_{L^1} =: C_n/4$. With the help of the variation of constants formula we can write for the critical part

$$\begin{aligned} M_c(t) &= \sup_{0\leq s\leq t} \|E_c\widehat{u}_n(s)\rho_n\|_{L^1} \\ &\leq e^{\varepsilon^2 t}\|E_c\widehat{u}_n(0)\rho_n\|_{L^1} \\ &\quad + \int_0^t \sup_{k\in I_c} \left| k^2 e^{\lambda(k)(t-\tau)} \frac{\rho_n}{\rho_n^*}\right| \|E_c(\widehat{u}_n * \widehat{u}_n)(\tau)\rho_n^*\|_{L^1} d\tau. \end{aligned} \tag{3.11}$$

We split the function $\widehat{u}_n$ into $\widehat{u}_n^b$, $\widehat{u}_n^0$ and $\widehat{u}_n^c$, where $\widehat{u}_n^i = E_i\widehat{u}_n$, and we write for the convolution

$$\begin{aligned} E_c\Big((\widehat{u}_n^c + \widehat{u}_n^b + \widehat{u}_n^0) * (\widehat{u}_n^c + \widehat{u}_n^b + \widehat{u}_n^0)\Big) &= \underbrace{E_c\widehat{u}_n^c * \widehat{u}_n^c}_{=0} + 2E_c\widehat{u}_n^c * \widehat{u}_n^b \\ &\quad + 2E_c\widehat{u}_n^c * \widehat{u}_n^0 + \underbrace{E_c\widehat{u}_n^0 * \widehat{u}_n^0}_{=0} \\ &\quad + 2E_c\widehat{u}_n^b * \widehat{u}_n^0 + E_c\widehat{u}_n^b * \widehat{u}_n^b, \end{aligned}$$

thus we find with the help of (3.9)

$$\sup_{0\le s\le t} \|E_c(\widehat{u}_n * \widehat{u}_n)(s)\rho_n^*\|_{L^1} \le C\Big(2M_c(t)M_b(t) + 2M_c(t)M_0(t) + \varepsilon M_b^2(t) + \varepsilon M_b(t)M_0(t)\Big).$$

For the exponential part we calculate

$$\sup_{k\in I_c} \left|k^2 e^{\lambda(k)(t-\tau)}\frac{\rho_n}{\rho_n^*}\right| \le \sup_{K\in\mathbb{R}} C\varepsilon^2 e^{-K^2T}(|K|+1).$$

Here we use the substitutions $K = \frac{k-1}{\varepsilon}$ and $T = \varepsilon^2(t-\tau)$. A further substitution given by $S^2 = K^2T$ yields the estimate

$$\sup_{k\in I_c} \left|k^2 e^{\lambda(k)(t-\tau)}\frac{\rho_n}{\rho_n^*}\right| \le C\sup_{S\in\mathbb{R}} e^{-S^2}\left(\frac{|S|}{\sqrt{T}}+1\right)\varepsilon^2.$$

Since $e^{-S^2}|S|^\alpha$ is bounded by a constant C for all $S\in\mathbb{R}$ and $\alpha\in\mathbb{N}$, we write for (3.11)

$$M_c(t) \le \frac{C_n}{2} + C\int_{\varepsilon^2 t}^0 -\left(\frac{1}{\sqrt{T}}+1\right)dT \times\left(2M_c(t)M_b(t) + 2M_c(t)M_0(t) + \varepsilon M_b^2(t) + \varepsilon M_b(t)M_0(t)\right).$$

A simple integration gives us

$$M_c(t) \le \frac{C_n}{2} + C(2\sqrt{\varepsilon^2 t} + \varepsilon^2 t)(2C_n(C_b+C_0)+1) \tag{3.12}$$

as long as $M_c(t)\le C_n,\ M_b(t)\le C_b,\ M_0(t)\le C_0$ and if we choose ε and t such that $\varepsilon C_b^2 + \varepsilon C_b C_0 < 1$ and $e^{\varepsilon^2 t} < 2$.
For the estimate in the set I_b we write

$$M_b(t) \le \|\widehat{u}_n(0)\rho_n\|_{L^1} + \int_0^t \sup_{k\in I_b}\left|k^2 e^{\lambda(k)(t-\tau)}\frac{\rho_n}{\rho_n^*}\right| \|E_b(\widehat{u}_n * \widehat{u}_n)(\tau)\rho_n^*\|_{L^1}d\tau.$$

As above we obtain here for the nonlinearities

$$\sup_{0\le s\le t} \|E_b(\widehat{u}_n * \widehat{u}_n)(s)\rho_n^*\|_{L^1} \le C\Big(M_c^2(t) + 2\varepsilon^{1/2}M_c(t)M_b(t) + 2\varepsilon^{1/2}M_c(t)M_0(t) + \varepsilon^{3/2}M_0^2(t) + 2\varepsilon M_b(t)M_0(t) + \varepsilon M_b^2(t)\Big).$$

Now since the integrand in the set I_b is exponentially damped we can write

$$M_b(t) \leq \frac{C_n}{2} + C(C_n^2 + 1), \tag{3.13}$$

where we have chosen $\varepsilon > 0$ such that

$$2\varepsilon^{1/2} C_n C_b + 2\varepsilon^{1/2} C_n C_0 + \varepsilon^{3/2} C_0^2 + 2\varepsilon C_b C_0 + \varepsilon C_b^2 \leq 1.$$

The estimate in the I_0 works slightly differently. On the one hand, the quotient $\frac{\rho_n}{\rho_n^*}$ in I_0 is only $\mathcal{O}(1)$ in I_0. On the other hand, we have to control the integrand over the time T/ε^2, which leads to a potential problem. Therefore we arrange the integral in the following way:

$$M_0(t) \leq \|\widehat{u}_n(0)\rho_n\|_{L^1} + \int_0^t \sup_{k\in I_0} |ke^{\lambda(k)(t-\tau)}| \|kE_0(\widehat{u}_n * \widehat{u}_n)(\tau)\rho_n\|_{L^1} d\tau.$$

Here, with the help of a simple substitution $k^2(t-\tau) = s^2$, we can write

$$\sup_{k\in I_0} \left|ke^{\lambda(k)(t-\tau)}\right| \leq \sup_{s\in\mathbb{R}} \left|Ce^{-s^2} \frac{s}{(t-\tau)^{1/2}}\right| \leq C(t-\tau)^{-1/2}.$$

Hence

$$M_0(t) \leq \frac{C_n}{4} + \int_0^t C(t-\tau)^{-1/2} \|kE_0(\widehat{u}_n * \widehat{u}_n)(\tau)\rho_n\|_{L^1} d\tau.$$

We write

$$\begin{aligned} \|kE_0(\widehat{u}_n * \widehat{u}_n)(\tau)\rho_n\|_{L^1} \leq\ & \|kE_0(\widehat{u}_n^c * \widehat{u}_n^c)(\tau)\rho_n\|_{L^1} + 2\|kE_0(\widehat{u}_n^c * \widehat{u}_n^b)(\tau)\rho_n\|_{L^1} \\ & + \|kE_0(\widehat{u}_n^0 * \widehat{u}_n^0)(\tau)\rho_n\|_{L^1} + 2\|kE_0(\widehat{u}_n^0 * \widehat{u}_n^b)(\tau)\rho_n\|_{L^1} \\ & + \|kE_0(\widehat{u}_n^b * \widehat{u}_n^b)(\tau)\rho_n\|_{L^1}. \end{aligned}$$

In order to gain an ε we proceed as follows. Since $|k| \leq \varepsilon + |k|$ we can write

$$\begin{aligned} kE_0\rho_n(k) =& k\varepsilon^{-\frac{n+3}{2}} (\varepsilon + |k|)^{\frac{n-1}{2}} \\ \leq& \varepsilon\varepsilon^{-\frac{n+5}{2}} (\varepsilon + |k|)(\varepsilon + |k|)^{\frac{n-1}{2}} \\ \leq& \varepsilon E_0 \rho_n^*(k). \end{aligned}$$

Using the properties in (3.9) and the fact that $\sup_{k\in I_0} |k| \leq C$ yields

$$M_0(t) \leq \frac{C_n}{4} + C\varepsilon\sqrt{t}(C_n^2 + C_nC_b + \varepsilon C_0^2 + \varepsilon C_0C_b + \varepsilon C_b^2). \tag{3.14}$$

Due to the inequalities (3.13) und (3.14) we can substitute both C_0 and C_b by C_n in (3.12). Let $\widetilde{T}_n = \varepsilon^2 t > 0$ be arbitrary small such that

$$C\left(\sqrt{\widetilde{T}_n} + \widetilde{T}_n\right)(2C_n(C_b + C_0) + 1) < \frac{C_n}{2}.$$

Hence the inequality (3.12) yields the boundedness for C_n and with the help of (3.13) and (3.14) we obtain the boundedness for C_b and C_0. Now the estimates $M_i(\widetilde{T}_n/\varepsilon^2) \leq \widetilde{C}_n$ holds for a positive time $\widetilde{T}_n$ and for all $i \in \{c, b, 0\}$ where we set $\widetilde{C}_n := C_n + C_b + C_0$ and $T_{n+1} = T_n + \widetilde{T}_n$.

II) Now we show that there exist ε_0, $\widehat{C}_n > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ the estimate $\|\widehat{u}_n(\widetilde{T}_n/\varepsilon^2)\rho_{n+1}\|_{L^1} \leq \widehat{C}_n$ holds. Using the weight ρ_{n+1} for the estimate around critical modes yields

$$\begin{aligned}
\|E_c\widehat{u}_n(\widetilde{T}_n/\varepsilon^2)\rho_{n+1}\|_{L^1} \leq\ & C \sup_{k\in[3/4,5/4]} \left| e^{-(k-1)^2\widetilde{T}_n/\varepsilon^2} \frac{\rho_{n+1}}{\rho_n} \right| \|\widehat{u}_n(0)\rho_n\|_{L^1} \\
& + \int_0^{\widetilde{T}_n/\varepsilon^2} \sup_{k\in I_c} \left| k^2 e^{\lambda(k)(\widetilde{T}_n/\varepsilon^2-\tau)} \frac{\rho_{n+1}}{\rho_n^*} \right| d\tau (2CC_n(C_b + C_0) + 1) \\
\leq\ & C\frac{C_n}{4} \sup_{|k|<1/4} \left| e^{-k^2\widetilde{T}_n/\varepsilon^2}(1 + |k|/\varepsilon)^{1/2} \right| \\
& + C\int_0^{\widetilde{T}_n/\varepsilon^2} \sup_K e^{-K^2\varepsilon^2(\widetilde{T}_n/\varepsilon^2-\tau)}\varepsilon^2(\varepsilon K + 1)^2(1 + |K|)^{3/2} d\tau \\
& \times (2CC_n(C_b + C_0) + 1) \\
\leq\ & C\frac{C_n}{4}(1 + \widetilde{T}_n^{-1/4}) + C(\widetilde{T}_n + \widetilde{T}_n^{1/4})(2CC_n(C_b + C_0) + 1) \\
=\ & \mathcal{O}(1),
\end{aligned}$$

where we applied the substitution $K = \frac{k-1}{\varepsilon}$ and used the fact that $\lambda(k)$ is only positive for k with $|k \pm 1| \leq \mathcal{O}(\varepsilon)$. For the stable modes we calculate

$$\begin{aligned}
\|E_b\widehat{u}_n(\widetilde{T}_n/\varepsilon^2)\rho_{n+1}\|_{L^1} \leq\ & C_n e^{-\sigma^2\widetilde{T}_n/\varepsilon^2}\frac{1}{\varepsilon} \\
& + \int_0^{\widetilde{T}_n/\varepsilon^2} \sup_{k\in I_b} \left| k^2 e^{\lambda(k)(\widetilde{T}_n/\varepsilon^2-\tau)} \frac{\rho_{n+1}}{\rho_n^*} \right| d\tau (CC_n^2 + \mathcal{O}(\varepsilon^{1/2})) \\
\leq\ & 1 + C(C_n^2 + \mathcal{O}(\varepsilon^{1/2})) = \mathcal{O}(1),
\end{aligned}$$

where we use $\rho^b_{n+1} \le \rho^{b,*}_n$, and have chosen ε such that $C_n e^{-\sigma^2 \widetilde{T}_n/\varepsilon^2} \frac{1}{\varepsilon} \le 1$.
For the last section I_0 we write

$$
\begin{aligned}
\|E_0 \widehat{u}_n(\widetilde{T}_n/\varepsilon^2)\rho_{n+1}\|_{L^1} \;&\le\; C \sup_{k\in I_0} \left| e^{-k^2 \widetilde{T}_n/\varepsilon^2} \frac{\rho_{n+1}}{\rho_n} \right| \|\widehat{u}(0)\rho_n\|_{L^1} \\
&\quad + \int_0^{\widetilde{T}_n/\varepsilon^2} \sup_{k\in I_0} \left| k e^{\lambda(k)(\widetilde{T}_n/\varepsilon^2 - \tau)} \frac{\rho_{n+1}}{\rho_n} \right| \|k E_0(\widehat{u}_n * \widehat{u}_n)(\tau)\rho_n\|_{L^1} d\tau \\
&\le\; C \frac{C_n}{4}(1 + \widetilde{T}_n^{-1/4}) \\
&\quad + \underbrace{\int_0^{\widetilde{T}_n/\varepsilon^2} \sup_{k\in I_0} \left| k e^{\lambda(k)(\widetilde{T}_n/\varepsilon^2 - \tau)} \left(1 + \frac{|k|}{\varepsilon}\right)^{1/2} \right| d\tau}_{\le C\widetilde{T}^{3/4}} (\mathcal{O}(\varepsilon)) \\
&\le\; C \frac{C_n}{4}(1 + \widetilde{T}_n^{-1/4}) + 1 = \mathcal{O}(1).
\end{aligned}
$$

Here we apply the same method as above.

Due to our choice of initial condition in Theorem 3.1.3, for the induction base, we only need to show that there exist a T_1 and C_1 such that $\|\hat{u}(., T_1/\varepsilon^2)\rho_1\| \le C_1$. But this will coincides with the first part of our inductive step for the choice $T_1 = \widetilde{T}_0$ and $T_0 = 0$. Hence Theorem 3.1.3 and Lemma 3.3.1 is proved.

Bibliography

[Alt06] H. W. Alt. *Lineare Funktionalanalysis*, volume 1. Springer Verlag, 5 edition, 2006.

[BSTU06] Kurt Busch, Guido Schneider, Lasha Tkeshelashvili, and Hannes Uecker. Justification of the nonlinear Schrödinger equation in spatially periodic media. *Z. Angew. Math. Phys.*, 57(6):905–939, 2006.

[BvHS95] P. Bollerman, A. van Harten, and G. Schneider. On the justification of the Ginzburg-Landau approximation. In *Nonlinear dynamics and pattern formation in the natural environment (Noordwijkerhout, 1994)*, volume 335 of *Pitman Res. Notes Math. Ser.*, pages 20–36. Longman, Harlow, 1995.

[CE90] P. Collet and J.-P. Eckmann. The time dependent amplitude equation for the Swift-Hohenberg problem. *Comm. Math. Phys.*, 132(1):139–153, 1990.

[Cra85] Walter Craig. An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits. *Comm. Partial Differential Equations*, 10(8):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.

[dRCES71] diPrima R. C., W. Eckhaus, and L. A. Segel. Non-linear wave-number interaction in near critical two dimensional flows. *J. Fluid Mech.*, 49:705–744, September 1971.

[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.

[DSSS09] Arjen Doelman, Björn Sandstede, Arnd Scheel, and Guido Schneider. The dynamics of modulated wave trains. *Mem. Amer. Math. Soc.*, 199(934):viii+105, 2009.

[Eck93] W. Eckhaus. The Ginzburg-Landau manifold is an attractor. *J. Nonlinear Sci.*, 3(3):329–348, 1993.

[HSZ11] Tobias Häcker, Guido Schneider, and Dominik Zimmermann. Justification of the Ginzburg-Landau approximation in case of marginally stable long waves. *J. Nonlinear Sci.*, 21(1):93–113, 2011.

[JZ10] M. A. Johnson and K. Zumbrun. Rigorous justification of the Whitham modulation equations for the generalized Korteweg-de Vries equation. *Stud. Appl. Math.*, 125(1):69–89, 2010.

[Kal87] L. A. Kalyakin. Asymptotic decay of a one-dimensional wave packet in a nonlinear dispersive medium. *Mat. Sb. (N.S.)*, 132(174)(4):470–495, 592, 1987.

[KSM92] Pius Kirrmann, Guido Schneider, and Alexander Mielke. The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. Roy. Soc. Edinburgh Sect. A*, 122(1-2):85–91, 1992.

[MS95] Alexander Mielke and Guido Schneider. Attractors for modulation equations on unbounded domains—existence and comparison. *Nonlinearity*, 8(5):743–768, 1995.

[MS04a] Ian Melbourne and Guido Schneider. Phase dynamics in the complex Ginzburg-Landau equation. *J. Differential Equations*, 199(1):22–46, 2004.

[MS04b] Ian Melbourne and Guido Schneider. Phase dynamics in the real Ginzburg-Landau equation. *Math. Nachr.*, 263/264:171–180, 2004.

[NW69] A. Newell and J. Whitehead. Finite bandwidth, finite amplitude convection. *J. Fluid Mech.*, (38 (1969) 279-303), 1969.

[Sch94a] G. Schneider. A new estimate for the Ginzburg-Landau approximation on the real axis. *J. Nonlinear Sci.*, 4(1):23–34, 1994.

[Sch94b] Guido Schneider. Error estimates for the Ginzburg-Landau approximation. *Z. Angew. Math. Phys.*, 45(3):433–457, 1994.

[Sch94c] Guido Schneider. Global existence via Ginzburg-Landau formalism and pseudo-orbits of Ginzburg-Landau approximations. *Comm. Math. Phys.*, 164(1):157–179, 1994.

[Sch95] Guido Schneider. Analyticity of Ginzburg-Landau modes. *J. Differential Equations*, 121(2):233–257, 1995.

[Sch98] Guido Schneider. Justification of modulation equations for hyperbolic systems via normal forms. *NoDEA Nonlinear Differential Equations Appl.*, 5(1):69–82, 1998.

[Sch99a] Guido Schneider. Global existence results for pattern forming processes in infinite cylindrical domains—applications to 3D Navier-Stokes problems. *J. Math. Pures Appl. (9)*, 78(3):265–312, 1999.

[Sch99b] Guido Schneider. Some characterizations of the Taylor-Couette attractor. *Differential Integral Equations*, 12(6):913–926, 1999.

[Sch05] Guido Schneider. Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. *J. Differential Equations*, 216(2):354–386, 2005.

[SW00] Guido Schneider and C. Eugene Wayne. The long-wave limit for the water wave problem. I. The case of zero surface tension. *Comm. Pure Appl. Math.*, 53(12):1475–1535, 2000.

[SW02] Guido Schneider and C. Eugene Wayne. The rigorous approximation of long-wavelength capillary-gravity waves. *Arch. Ration. Mech. Anal.*, 162(3):247–285, 2002.

[SZ13] Guido Schneider and Dominik Zimmermann. Justification of the Ginzburg-Landau approximation for an instability as it appears for Marangoni convection. *Math. Methods Appl. Sci.*, 36(9):1001–1013, 2013.

[vH91] A. van Harten. On the validity of the Ginzburg-Landau equation. *J. Nonlinear Sci.*, 1(4):397–422, 1991.

[Whi65a] G. B. Whitham. A general approach to linear and non-linear dispersive waves using a Lagrangian. *J. Fluid Mech.*, 22:273–283, 1965.

[Whi65b] G. B. Whitham. Non-linear dispersive waves. *Proc. Roy. Soc. Ser. A*, 283:238–261, 1965.