Interaction of modulated pulses in nonlinear oscillator chains

Guido Schneider, Hannes Uecker, Michael Wand

January 9, 2009

Abstract

We formally derive and rigorously justify the modulation equations of lowest order for the interaction of two modulated pulses on a one-dimensional nonlinear oscillator chain. We show that solutions with the initial form of the assumed ansatz preserve this form over time intervals with positive macroscopic length, and we show a bound on the possible shift of the envelope caused by the interaction. Thus we rigorously justify and quantify the statement that under the given conditions there is almost no interaction of the modulated pulses.

Keywords: nonlinear oscillator chains, NLS, pulse interaction

1 Introduction

Finding continuum models for discrete, atomistic systems of ODEs is a major problem in contemporary multiscale analysis. In this paper we consider the macroscopic limit of a one-dimensional nonlinear oscillator chain (NOC) with interaction potentials between neighboring oscillators and a stabilizing background potential. We choose the initial data in a specific class of functions and obtain an evolution of the data in the given function class. This will be called the macroscopic limit problem.

For instance, in different scaling regimes the famous FPU model [FPU55] can be reduced to different macroscopic limit problems. In the long wave limit one obtains the Korteveg–deVries equation (KdV), see, e.g., [ZK65, SW00, McM05], while macroscopic envelopes of modulated microscopic carrier waves fulfill the Nonlinear Schrödinger equation (NLS), see [GM04] and the references therein. Both macroscopic equations, KdV and NLS, possess soliton solutions, i.e. localised pulses which interact in a particle–like fashion. In the NLS case these formally yield approximate modulating pulse solutions of the NOC. Moreover, in [GM04] the NLS approximation of NOCs was rigorously justified for the case of a single carrier wave, which also yields rigorous results for approximate modulating pulses (with one wave number) in NOCs.

Here we investigate the interaction of two modulating pulses with different carrier waves on a NOC. These two pulses travel with different group speeds c_A and c_B, and, loosely speaking, our main result is as follows, see Section 3 for the precise result and Fig. 1 for an illustration. If we assume that initially the pulses are separated, with the slower one ahead of the faster one, then the faster one will overtake the slower one on some macroscopic time interval. Under some (generic) non–resonance conditions we show that there is almost no interaction of the two pulses and quantify this statement appropriately. In detail, we show that after interaction the pulses retain their shape and the main effect is an envelope shift, i.e., a shift in position experienced by each pulse due to the interaction. However, we also show that this shift is at most $\mathcal{O}(\varepsilon)$, where ε is the order of the amplitude (and of the inverse width) of each pulse.

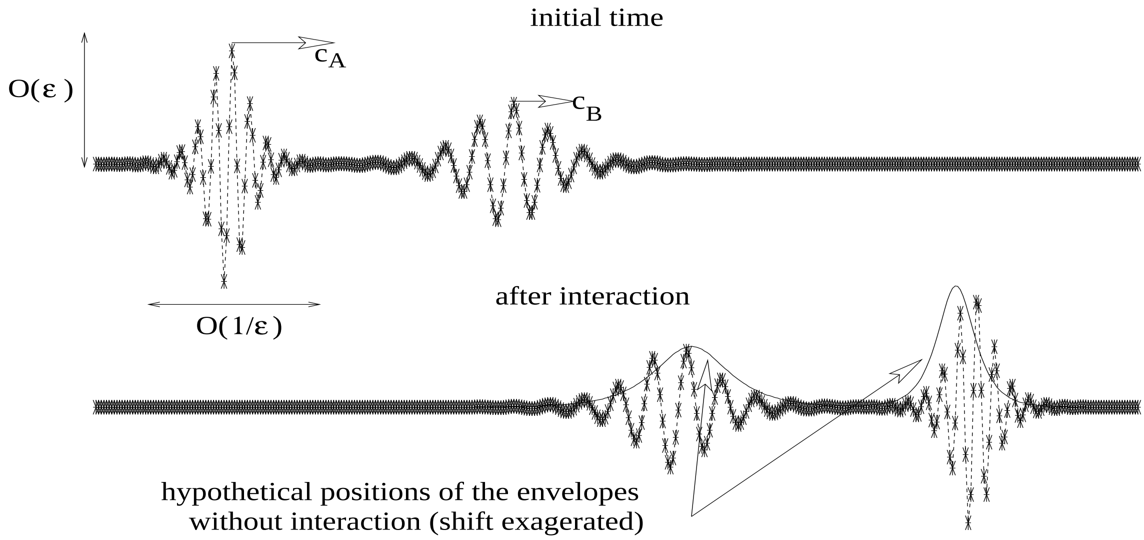

Figure 1: Sketch of pulse interaction.

This result should be contrasted to the position shift due to interaction experienced by N-solitons of a single NLS, which is proved by inverse scattering transform, see, e.g. [AS81]. Our analysis involves a system of two NLS, each one corresponding to one pulse, with different underlying wave numbers, and the proof relies on estimates for the approximation of the NOC by the system of NLS and not on genuine properties of the NLS itself.

Analogous results were obtained in [CBSU07, CBCSU08] for a similar problem on a continuous one-dimensional string, while in [BF06, see in particular §2.1] almost linear (interaction) behavior was shown in a more general setting including nonlinear lattice equations, but with less detailed asymptotics, in particular only leading to an $\mathcal{O}(1)$ upper bound for the envelope shift. Compared to the continuous case, the asymptotic expansions in the discrete case become slightly more involved due to the mixing of the discrete lattice site variable and the continuous macroscopic space variable.

Physically, the shifts caused by the interaction of pulses may for instance be used to probe the presence of a standing (or slow) pulse by sending a traveling pulse with a different wave number through the material. For the (continuous) case of photonic crystals this has been proposed in [TPB04]. In lattices, such shifts have for instance been experimentally observed for (non-modulated) pulses which are modeled by solitons in Toda lattices, see, e.g. [Tsu89]. Finally, the shifts may be used to distinguish nonlinear from linear material which shows no interaction at all.

The plan of this paper is as follows. In Section 2 we present a formal derivation of the equations for the approximation of the pulse interaction. This is called formal derivation since a priori the existence of solutions of the required form is not clear at all. In Section 3, we prove the validity of the approximation by introducing a suitable phase space for our system and by estimating the norm of the error on an $\mathcal{O}(1/\varepsilon^2)$ time scale. This finally allows us to show the desired estimate for the envelope shift of the pulses after the interaction.

2 Formal Description of the Pulse Interaction

2.1 The NLS Pulse

We consider equations of the form

$$\ddot{x}_j = V'(\partial_j^+ x) - V'(\partial_j^- x) - W'(x_j), \quad j \in \mathbb{Z}, \tag{2.1}$$

where $x_j(t)$ is the deviation at time $t \geq 0$ of an oscillator from its rest position in an infinite chain $(x_j)_{j\in\mathbb{Z}}$, and where $\partial_j^\pm x = \pm(x_{j\pm1} - x_j)$. Thus we assume that each oscillator interacts with its nearest neighbor via the potential V, and additionally the existence of a stabilizing background potential W. More general chain equations are possible and lead to similar results (e.g.[GM06]).

To simplify the presentation we make the basic assumption that the potentials do not contain cubic terms, i.e., that there are no quadratic nonlinearities. However see Remark 3.8. Prototypical potentials which yield cubic nonlinearities in (2.1) are

$$V(d) = \frac{v_1}{2}d^2 + \widetilde{V}(d) \quad \text{with} \quad \widetilde{V}(d) = \frac{v_3}{4}d^4,$$
$$W(y) = \frac{w_1}{2}y^2 + \widetilde{W}(y) \quad \text{with} \quad \widetilde{W}(y) = \frac{w_3}{4}y^4.$$

We split (2.1) into a linear and a nonlinear part, i.e. $\ddot{x}_j = L_j x + N_j(x)$. The linearized system

$$\ddot{x}_j = L_j x := v_1(\partial_j^+ x - \partial_j^- x) - w_1 x_j = v_1(x_{j+1} - 2x_j + x_{j-1}) - w_1 x_j \tag{2.2}$$

has exact spatiotemporally periodic basic solutions $\boldsymbol{E}(t,j) = \mathrm{e}^{i(\hat{\omega}t+\vartheta j)}$, where the frequency $\hat{\omega} = \hat{\omega}(\vartheta)$ depends on the wave number ϑ via the dispersion relation $\hat{\omega}^2(\vartheta) = 2v_1(1-\cos\vartheta)+w_1$. Throughout the paper, we require the stability condition $\hat{\omega}^2(\vartheta) > 0$ for all ϑ, and choose $\hat{\omega}(\vartheta) > 0$. The stability condition is equivalent to $\min\{w_1, w_1 + 4v_1\} > 0$.

We fix a wave number ϑ and write $\hat{\omega}$, $\hat{\omega}'$ and $\hat{\omega}''$ for $\hat{\omega}(\vartheta)$, $\hat{\omega}'(\vartheta)$ and $\hat{\omega}''(\vartheta)$, where ϑ needs to be chosen in such a way that $\hat{\omega}'' \neq 0$; this is always satisfied in a neighborhood of 0. Following [GM04], (2.1) has solutions which are slow modulations of the periodic basic pattern, in the form

$$X_j(t) = \varepsilon A(\varepsilon^2 t, \varepsilon(j - ct))\boldsymbol{E}(t,j) + \mathrm{cc} + \mathcal{O}(\varepsilon^2). \tag{2.3}$$

where $c = c(\vartheta)$ is the group velocity of wave packets with wave number ϑ, and cc denotes complex conjugate. We let $\tau = \varepsilon^2 t$ and $\xi = \varepsilon(j - ct)$ for the macroscopic time- and space-variable, respectively. Since the solutions given by (2.3) have small amplitude and narrow spectral content only the long time scale allows to see the effects of nonlinearity and dispersion. Inserting the ansatz (2.3) into (2.1) yields that the group velocity must satisfy $c = -\hat{\omega}'$, and that A must satisfy a Nonlinear Schrödinger Equation (NLS)

$$i\partial_\tau A = \frac{1}{2}\hat{\omega}''\partial_\xi^2 A + \rho|A|^2 A \quad \text{with} \quad \rho = -\frac{12v_3(1-\cos\vartheta)^2 + 3w_3}{2\hat{\omega}}. \tag{2.4}$$

A general time–dependent spatially localized solution of equation (2.4) describes a modulating pulse of (2.1). However, to describe the interaction of pulses, we shall restrict to functions which are time-independent in lowest order. See Remark 3.7 for the general time dependent case. It turns out that a time-independent solution of the modulation equation (2.4) can only assume the form of a pulse if we add a correction term of order $\mathcal{O}(\varepsilon^2)$ to the frequency, i.e., we define

$$\omega^2(\vartheta) = 2v_1(1-\cos\vartheta) + w_1 + \varepsilon^2\gamma \tag{2.5}$$

with a constant γ and determine which values this constant may take such that the modulation equation has modulating pulse solutions. This means that we slightly change the ansatz (2.3), but not the equation (2.1), and thus obtain a modulation equation different from (2.4). Note that $\omega'(\vartheta) = \hat{\omega}'(\vartheta) + \mathcal{O}(\varepsilon^2)$ and $\omega''(\vartheta) = \hat{\omega}''(\vartheta) + \mathcal{O}(\varepsilon^2)$.

In detail, fixing ϑ, choosing $\omega = \omega(\vartheta)$, and inserting the lowest-order time-independent ansatz

$$X_j(t) = \varepsilon A(\varepsilon(j-ct))\boldsymbol{E}(t,j) + \mathrm{cc} + \mathcal{O}(\varepsilon^2), \quad \boldsymbol{E}(t,j) = \mathrm{e}^{\mathrm{i}(\omega t + \vartheta j)}, \tag{2.6}$$

into (2.1) now yields the modulation equation

$$0 = \gamma A + \omega\omega''\partial_\xi^2 A - \sigma|A|^2 A, \tag{2.7}$$

with $\sigma = 12v_3(1-\cos\vartheta)^2 + 3w_3$. For $\gamma = 0$ this is the stationary case of (2.4) due to $\rho = \frac{\sigma}{2\omega}$. For (2.7) we seek solutions in the form

$$A(\xi;\xi^{(0)},\phi^{(0)}) = A_{\text{pulse}}(\xi - \xi^{(0)})\mathrm{e}^{i\phi^{(0)}},$$

where the envelope shift $\xi^{(0)}$ and phase shift $\phi^{(0)}$ are parameters, and where A_{pulse} satisfies the equation

$$\partial_\xi^2 \tilde{A} = C_1\tilde{A} - C_2\tilde{A}^3 \tag{2.8}$$

with $\tilde{A}(\xi)$ real-valued, $C_1 = -\gamma/\omega\omega''$ and $C_2 = -\sigma/\omega\omega''$. This equation has pulse solutions if and only if $C_1 > 0$ and $C_2 > 0$. These pulses may be calculated explicitely to

$$A^{\pm}_{\text{pulse}}(\xi) = \pm\left(\frac{2C_1}{C_2}\right)^{1/2} \operatorname{sech}(C_1^{1/2}\xi), \tag{2.9}$$

and in the $(A, \partial_X A)$-plane correspond to homoclinic connections of the origin with itself. We summarize the conditions for C_1 and C_2, which in particular imply that γ and v_1 must have the same sign, as follows,

Lemma 2.1. *(2.8) admits pulse solutions if and only if*

$$C_1 = \frac{-\gamma}{\omega\omega''} > 0 \quad \textit{and} \quad C_2 = \frac{-\left[12v_3(1-\cos\vartheta)^2 + 3w_3\right]}{\omega\omega''} > 0.$$

2.2 The pulse interaction

We are looking for approximate solutions Z of amplitude $\mathcal{O}(\varepsilon)$ of (2.1) which represent the interaction of two pulses. The two pulses, which we will symbolically denote by $\mathcal{P}_A$ and $\mathcal{P}_B$, are required to have different wave numbers. Therefore we introduce a double set of variables: For each pulse, we define wave number, frequency, and group velocity ϑ_A, $\omega_A = \omega(\vartheta_A)$, and c_A resp. ϑ_B, $\omega_B = \omega(\vartheta_B)$, and c_B, and the two different macroscopic space variables $\xi_A = \varepsilon(j - c_A t)$ and $\xi_B = \varepsilon(j - c_B t)$. The values for ϑ_A and ϑ_B must be chosen according to Lemma 2.1.

In order to quantify how well Z approximates a true solution, we shall use the residual defined componentwise by

$$\rho(Z)_j = L_j Z + N_j(Z) - \ddot{Z}_j, \tag{2.10}$$

and the goal is to compute a formal approximation Z so that the residual is of high order in ε. Thus we propose the multiscale ansatz

$$Z_j(t) = Z_j^A(t) + Z_j^B(t) + \varepsilon^3 M_j(t) \tag{2.11}$$

with

$$\begin{aligned} Z_j^A(t) &= \varepsilon A_1 \boldsymbol{E}_A + \varepsilon^3 A_3 \boldsymbol{E}_A^3 + \varepsilon^3 Y_A \boldsymbol{E}_A + \varepsilon^4 A_{4,3} \boldsymbol{E}_A^3 + \varepsilon^5 A_{5,5} \boldsymbol{E}_A^5 + \varepsilon^5 A_{5,3} \boldsymbol{E}_A^3 + \mathrm{cc}, \\ Z_j^B(t) &= \varepsilon B_1 \boldsymbol{E}_B + \varepsilon^3 B_3 \boldsymbol{E}_B^3 + \varepsilon^3 Y_B \boldsymbol{E}_B + \varepsilon^4 B_{4,3} \boldsymbol{E}_B^3 + \varepsilon^5 B_{5,5} \boldsymbol{E}_B^5 + \varepsilon^5 B_{5,3} \boldsymbol{E}_B^3 + \mathrm{cc}, \end{aligned}$$

explained in the following. $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ denote the basic pattern

$$\boldsymbol{E}_A(t,j) = \mathrm{e}^{i(\omega_A t + \vartheta_A j + \varepsilon \Omega_A(\xi_B))}, \quad \boldsymbol{E}_B(t,j) = \mathrm{e}^{i(\omega_B t + \vartheta_B j + \varepsilon \Omega_B(\xi_A))},$$

where the functions Ω_A resp. Ω_B represent the phase shifts of the pulses during interaction. Since the phase shift of a pulse should depend on the other pulse, we set $\Omega_A = \Omega_A(\xi_B)$ and $\Omega_B = \Omega_B(\xi_A)$. Furthermore, as already said, we require A_1 and B_1 to be time-independent, i.e. $A_1 = A_1(\xi_A)$ and $B_1 = B_1(\xi_B)$, and similarly $A_3 = A_3(\xi_A)$ and $B_3 = B_3(\xi_B)$, while the remaining macroscopic functions depend on τ and ξ_A, ξ_B. The idea of the ansatz is as follows:

- A_1, A_3 and B_1, B_3 determine the shape of the modulating pulse; in particular, they represent internal dynamics of each pulse since for each pulse they will be chosen independently of the other pulse.

- Y_A and Y_B are correction functions which evolve in slow time τ; they are zero at time $\tau = 0$.

- $A_{4,3}, A_{5,3}, \ldots$, and M are higher order corrections. M contains mixed terms, i.e. terms which contain mixed frequencies $\boldsymbol{E}_A^m \boldsymbol{E}_B^n$ with $mn \neq 0$.

We start with the computation of the macroscopic functions which generally works as follows: During the expansion of the multiscale ansatz (2.11) we obtain e.g. a term at $\varepsilon^k \boldsymbol{E}_A^m \boldsymbol{E}_B^n$, where for now we assume $(|m|,|n|) \neq (1,0)$ and $(|m|,|n|) \neq (0,1)$. To compensate this, we extend the ansatz by a term $\varepsilon^k H_{k,m,n} \boldsymbol{E}_A^m \boldsymbol{E}_B^n$ with a macroscopic function $H_{k,m,n} = H_{k,m,n}(\tau, \xi_A, \xi_B)$ and obtain an equation for $H_{k,m,n}$ at $\varepsilon^k \boldsymbol{E}_A^m \boldsymbol{E}_B^n$, namely

$$\left[(\omega(m\vartheta_A + n\vartheta_B))^2 - (m\omega_A + n\omega_B)^2\right] H_{k,m,n} = \mathcal{G}_{k,m,n}, \tag{2.12}$$

where the term $\mathcal{G}_{k,m,n}$ only contains functions $H_{\tilde{k},\tilde{m},\tilde{n}}$ with $\tilde{k} < k$. This allows us to compute the macroscopic functions $H_{k,m,n}$ step-by-step if the terms $[(\omega(m\vartheta_A + n\vartheta_B))^2 - (m\omega_A + n\omega_B)^2]$ in (2.12) are nonzero. We formalize this nonresonance condition:

(**FORM**) *In order to compute all macroscopic functions of an approximation to order k, we must have, for all $m, n \in \mathbb{Z}$, $|m| + |n| \leq k$ with $\mathcal{G}_{k,m,n} \neq 0$,*

$$(\omega(m\vartheta_A + n\vartheta_B))^2 - (m\omega_A + n\omega_B)^2 \neq 0. \tag{FORM}_{m,n}$$

Even for single frequencies, this condition may not be satisfied for all ϑ, cf. [GM06, Proposition 2.2]. For (2.11) it is rather straightforward to find ϑ_A, ϑ_B such that $(\text{FORM})_{m,n}$ is violated for small m, n, see Fig.2 for an example. However, if $(\text{FORM})_{m,n}$ holds to sufficient order, then we may calculate all macroscopic functions in the ansatz (2.11).

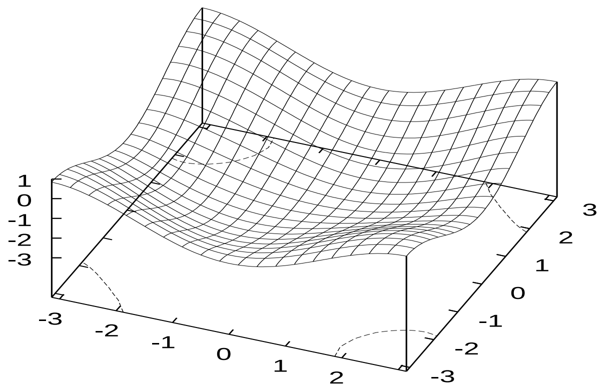

Figure 2: Violation of $(\text{FORM})_{1,1}$ for $v_1 = -0.24$, $w_1 = 1$ at zero contour lines.

Lemma 2.2. *Let* (FORM) *hold for all $k \leq 5$. Then we may explicitly calculate all functions in the ansatz* (2.11) *such that formally $\rho(Z) = \mathcal{O}(\varepsilon^6)$. The equations for A_1 and B_1 have the form given by* (2.7)*, and A_1 and B_1 may be chosen to have the form (2.9) of a pulse.*

Proof. See Section 2.3.

2.3 Proofs

Proof of Lemma 2.2. We insert (2.11) into (2.1) and then group terms according to their order and frequency, i.e. according to their factor $\varepsilon^k \boldsymbol{E}_A^m \boldsymbol{E}_B^n$. We first outline the results of this expansion. Therefore, let $H = H(\tau, \xi_A, \xi_B) : [0, \tau_0] \times \mathbb{R}^2 \to \mathbb{C}$

($\tau_0 > 0$) be any sufficiently smooth function. Then, on the left hand side of (2.1),

$$\begin{aligned}
\partial_t^2(H\boldsymbol{E}_A^n) = \Big\{ & -n^2\omega_A^2 \\
& + \varepsilon\,[2ni\omega_A(-c_A\partial_{\xi_A}H - c_B\partial_{\xi_B}H)] \\
& + \varepsilon^2\big[c_A^2\partial_{\xi_A}^2 H + c_B^2\partial_{\xi_B}^2 H + 2c_Ac_B\partial_{\xi_A}\partial_{\xi_B}H + 2ni\omega_A\partial_\tau H + 2n^2c_BH\partial_{\xi_B}\omega_A\big] \\
& + \varepsilon^3\big[-2c_A\partial_\tau\partial_{\xi_A}H - 2c_B\partial_\tau\partial_{\xi_B}H - 2nic_B\partial_{\xi_B}\omega_A(-c_A\partial_{\xi_A}H - c_B\partial_{\xi_B}H)\big] \\
& + \varepsilon^4\big[\partial_\tau^2 H + 2nic_B\partial_{\xi_B}\omega_A\partial_\tau H + nic_B^2\partial_{\xi_B}^2\omega_A H - n^2(c_B^2\partial_{\xi_B}\omega_A)^2H\big]\Big\}\boldsymbol{E}_A^n.
\end{aligned} \tag{2.13}$$

Analogous formulas hold for terms with other frequencies.

To deal with the difference operators on the right hand side of (2.1) we first note the formula

$$\boldsymbol{E}_A(t, j \pm 1) = \mathrm{e}^{i\gamma_A^\pm}\boldsymbol{E}_A(t,j), \quad \text{where} \quad \mathrm{e}^{i\gamma_A^\pm} = \mathrm{e}^{\pm i\vartheta_A + i\varepsilon(\omega_A(\xi_B\pm\varepsilon)-\omega_A(\xi_B))},$$

and then use a Taylor expansion of the macroscopic functions to obtain L_jZ^A and $N_j(Z)$ as an expansion in powers of ε and $\boldsymbol{E}_A$. In detail, for

$$Z = \varepsilon a_1 + \varepsilon^3 a_3 + \varepsilon^4 r_1,$$

with

$$a_1 = A_1\boldsymbol{E}_A + B_1\boldsymbol{E}_B + \mathrm{cc}, \quad a_3 = A_3\boldsymbol{E}_A^3 + B_3\boldsymbol{E}_B^3 + Y_A\boldsymbol{E}_A + Y_B\boldsymbol{E}_B + \mathrm{cc} + M,$$

we find

$$\partial_j^\pm Z = \varepsilon a_1^\pm + \varepsilon^2 a_2^\pm + \varepsilon^3 a_3^\pm + \varepsilon^4 r_1^\pm,$$

where

$$\begin{aligned}
a_1^\pm &= \pm(\mathrm{e}^{\pm i\vartheta_A} - 1)A_1\boldsymbol{E}_A \pm (\mathrm{e}^{\pm i\vartheta_B} - 1)B_1\boldsymbol{E}_B + \mathrm{cc}, \\
a_2^\pm &= \mathrm{e}^{\pm i\vartheta_A}\partial_{\xi_A}A_1\boldsymbol{E}_A + \mathrm{e}^{\pm i\vartheta_B}\partial_{\xi_B}B_1\boldsymbol{E}_B + \mathrm{cc}, \\
a_3^\pm &= \pm\left[\mathrm{e}^{\pm i\vartheta_A}\left(\frac{1}{2}\partial_{\xi_A}^2 A_1 + i\partial_{\xi_B}\omega_A A_1\right)\right]\boldsymbol{E}_A \\
&\quad \pm \left[(\mathrm{e}^{\pm 3i\vartheta_A} - 1)A_3\right]\boldsymbol{E}_A^3 \pm \left[(\mathrm{e}^{\pm i\vartheta_A} - 1)Y_A\right]\boldsymbol{E}_A \\
&\quad \pm \left[\mathrm{e}^{\pm i\vartheta_B}\left(\frac{1}{2}\partial_{\xi_B}^2 B_1 + i\partial_{\xi_B}\Omega_B B_1\right)\right]\boldsymbol{E}_B \\
&\quad \pm \left[(\mathrm{e}^{\pm 3i\vartheta_B} - 1)B_3\right]\boldsymbol{E}_B^3 \pm \left[(\mathrm{e}^{\pm i\vartheta_B} - 1)Y_B\right]\boldsymbol{E}_B + \mathrm{cc} + \partial_j^\pm M.
\end{aligned}$$

Inserting this into the nonlinear part yields

$$\begin{aligned}
N_j(Z) = \varepsilon^3 & \Big\{v_3\big[(a_1^+)^3 - (a_1^-)^3\big] - w_1 a_1^3\Big\} \\
& + \varepsilon^4\Big\{3v_3\big[(a_2^+)(a_1^+)^2 - (a_2^-)(a_1^-)^2\big]\Big\} \\
& + \varepsilon^5\Big\{3v_3\big[(a_3^+)(a_1^+)^2 + (a_2^+)^2(a_1^+) - (a_3^-)(a_1^-)^2 - (a_2^-)^2(a_1^-)\big] + 3w_3a_3a_1^2\Big\} \\
& + \varepsilon^6 R_N(j).
\end{aligned}$$

Due to (FORM), we know that we can calculate all macroscopic functions at orders $\varepsilon^k \boldsymbol{E}_A^m \boldsymbol{E}_B^n$, where $(|m|,|n|) \neq (1,0)$ and $(|m|,|n|) \neq (0,1)$. Therefore we only need to perform the calculations at orders $\varepsilon^k \boldsymbol{E}_A$. The case $\varepsilon^k \boldsymbol{E}_B$ gives the same results with exchanged variables.

At order ε^1 and ε^2 we observe only linear effects. At $\varepsilon \boldsymbol{E}_A$, we choose $\omega_A > 0$ to satisfy

$$\left[v_1((\mathrm{e}^{i\vartheta_A}-1)-(1-\mathrm{e}^{-i\vartheta_A}))-w_1-\varepsilon^2\gamma_A+\omega_A^2\right]A_1=0,$$

which is the dispersion relation with the $\mathcal{O}(\varepsilon^2)$ correction. At $\varepsilon^2 \boldsymbol{E}_A$, we get

$$2i\omega_A\omega_A' + 2ic_A\omega_A = 0,$$

which implies $c_A = -\omega_A'$. At order ε^3, nonlinearity and dispersion appear. At $\varepsilon^3 \boldsymbol{E}_A$, we get

$$\begin{aligned}
0 =\ & \gamma_A A_1 + 2v_1 \sin\vartheta_A \partial_{\xi_B}\Omega_A A_1 + v_1\cos\vartheta_A \partial_{\xi_A}^2 A_1 - \hat{\omega}_A^2 Y_A \\
& + v_3\big[\, \big(12(1-\cos\vartheta_B)|B_1|^2(\mathrm{e}^{i\vartheta_A}-1)A_1 \\
& \quad + 6(1-\cos\vartheta_A)|A_1|^2(\mathrm{e}^{i\vartheta_A}-1)A_1\big) \\
& \quad - \big(12(1-\cos\vartheta_B)|B_1|^2(1-\mathrm{e}^{-i\vartheta_A})A_1 \\
& \quad + 6(1-\cos\vartheta_A)|A_1|^2(1-\mathrm{e}^{-i\vartheta_A})A_1\big)\,\big] \\
& - 3w_3|A_1|^2A_1 + 6w_3|B_1|^2A_1 - \big(2c_B\omega_A\partial_{\xi_B}\Omega_A A_1 + c_A^2\partial_{\xi_A}^2 A_1\big) + \omega_A^2 Y_A.
\end{aligned} \tag{2.14}$$

At $\varepsilon^3 \boldsymbol{E}_B$, we get a similar equation. Since these equations are coupled, we cannot extract modulation equations for A_1 and B_1 directly. However, there is a simple way to uncouple these equations, namely by associating the coupling terms $|B_1|^2A_1$ resp. $|A_1|^2B_1$ with the phase shifts Ω_A resp. Ω_B. Thus, by combining these terms, we get

$$(c_A-c_B)\partial_{\xi_B}\Omega_A = -s_A^{(1)}|B_1|^2 \quad \text{and} \quad (c_B-c_A)\partial_{\xi_A}\Omega_B = -s_B^{(1)}|A_1|^2$$

with coefficients

$$s_A^{(1)} = \frac{v_3(\cos\vartheta_A-1)(1-\cos\vartheta_B)}{\omega_A} \quad \text{and} \quad s_B^{(1)} = \frac{v_3(\cos\vartheta_B-1)(1-\cos\vartheta_A)}{\omega_B},$$

and we choose Ω_A and Ω_B as

$$\Omega_A(\xi_B) = \Omega_A^0 + \int_{-\infty}^{\xi_B} \frac{-s_A^{(1)}}{c_A-c_B}|B_1|^2\mathrm{d}\tilde{\xi}_B, \quad \Omega_B(\xi_A) = \Omega_B^0 + \int_{-\infty}^{\xi_A} \frac{-s_B^{(1)}}{c_B-c_A}|A_1|^2\mathrm{d}\tilde{\xi}_A,$$

with real constants Ω_A^0 and Ω_B^0, which are determined by initial phase shifts. This means that the phase shift of a pulse travels with the other pulse.

At $\varepsilon^3 \boldsymbol{E}_A$, we now have

$$\begin{aligned}
0 =\ & \gamma_A A_1 + v_1\cos\vartheta_A\partial_{\xi_A}^2 A_1 - c_A^2\partial_{\xi_A}^2 A_1 - \hat{\omega}_A^2 Y_A \\
& - [12v_3(1-\cos\vartheta_A)^2 + 3w_3]|A_1|^2A_1 + \omega_A^2 Y_A. \\
& + \varepsilon 3v_3\Big\{ \left[(\mathrm{e}^{i\vartheta_A}-1)^2\mathrm{e}^{-i\vartheta_A} - (1-\mathrm{e}^{-i\vartheta_A})^2\mathrm{e}^{i\vartheta_A}\right] A_1^2\partial_{\xi_A}\bar{A}_1 \\
& \quad + 2(1-\cos\vartheta_A)|A_1|^2 2i\sin\vartheta_A\partial_{\xi_A}A_1\Big\},
\end{aligned} \tag{2.15}$$

where the $\mathcal{O}(\varepsilon)$ terms come from terms at order $\varepsilon^4 \boldsymbol{E}_A$ resp. $\varepsilon^4 \boldsymbol{E}_B$.

Since $\omega_A - \hat{\omega}_A = \mathcal{O}(\varepsilon^2)$, the terms with Y_A disappear from the equation at order ε^3. We similarly deal with the equation at $\varepsilon^3 \boldsymbol{E}_B$ and finally obtain two uncoupled equations for A_1 and B_1, namely

$$\begin{aligned} 0 &= \gamma_A A_1 + \omega_A \omega_A'' \partial_{\xi_A}^2 A_1 - [12 v_3 (1-\cos\vartheta_A)^2 + 3w_3]|A_1|^2 A_1 + \mathcal{O}(\varepsilon), \\ 0 &= \gamma_B B_1 + \omega_B \omega_B'' \partial_{\xi_B}^2 B_1 - [12 v_3 (1-\cos\vartheta_B)^2 + 3w_3]|B_1|^2 B_1 + \mathcal{O}(\varepsilon). \end{aligned} \tag{2.16}$$

Choosing our parameters according to Lemma 2.1, this system has pulse solutions for A_1 and B_1 which are $\mathcal{O}(\varepsilon)$ perturbations of (2.9).

Next we consider the correction functions Y_A and Y_B. At $\varepsilon^5 \boldsymbol{E}_A$ and $\varepsilon^5 \boldsymbol{E}_B$ we get a system of evolution equations for Y_A and Y_B, namely

$$\begin{aligned} 2i\omega_A \partial_\tau Y_A = {} & (v_1 \cos(3\vartheta_A) - c_A^2)\partial_{\xi_A}^2 Y_A + G_A(Y_A, Y_B) \\ & + \varepsilon^{-1} \Big\{ 2v_1 \cos\vartheta_A \left(A_1 \frac{i}{2} \partial_{\xi_B}^2 \Omega_A + \partial_{\xi_A} A_1 i \partial_{\xi_B} \Omega_A \right) \\ & \qquad + 6iv_3 \left(2(1-\cos\vartheta_B)|B_1|^2 \sin \partial_{\xi_A} A_1 - c_B \partial_{\xi_B} \Omega_A c_A \partial_{\xi_A} A_1 \right) \Big\} \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} 2i\omega_B \partial_\tau Y_B = {} & (v_1 \cos(3\vartheta_B) - c_B^2)\partial_{\xi_B}^2 Y_B + G_B(Y_A, Y_B) \\ & + \varepsilon^{-1} \Big\{ 2v_1 \cos\vartheta_B \left(B_1 \frac{i}{2} \partial_{\xi_A}^2 \Omega_B + \partial_{\xi_B} B_1 i \partial_{\xi_A} \Omega_B \right) \\ & \qquad + 6iv_3 \left(2(1-\cos\vartheta_A)|A_1|^2 \sin \partial_{\xi_B} B_1 - c_A \partial_{\xi_A} \Omega_B c_B \partial_{\xi_B} B_1 \right) \Big\}. \end{aligned} \tag{2.18}$$

Here, G_A and G_B contain only algebraic terms in Y_A and Y_B and terms which only depend on functions which we have already calculated. Lemma 2.3 below states that the system (2.17),(2.18) with initial conditions $Y_A|_{\tau=0} = 0$ and $Y_B|_{\tau=0} = 0$ has a unique solution, bounded independent of ε.

By these calculations all terms of orders $\varepsilon^k \boldsymbol{E}_A$ and $\varepsilon^k \boldsymbol{E}_B$, $k \leq 5$, cancel from ansatz (2.11). We could now proceed with calculating the modulation functions at the remaining orders, but here we restrict to A_3 and B_3 since these determine the next order shape of the single pulses. We obtain

$$A_3 = \nu_A A_1^3 \quad \text{and} \quad B_3 = \nu_B B_1^3 \tag{2.19}$$

with

$$\nu_A = \frac{v_3 s_A^2 (3 - s_A) - w_3}{\omega^2(3\vartheta_A) - 9\omega_A^2}, \qquad \text{where} \qquad s_A = 2(1-\cos\vartheta_A),$$

and analogously for ν_B, provided that $(\mathrm{FORM})_{(3,0)}$ and $(\mathrm{FORM})_{(0,3)}$ are satisfied. In particular, A_3 and B_3 also only depend on ξ_A resp. ξ_B. □

Lemma 2.3. *For all $s \geq 2$ and $\tau_0 > 0$ there is a constant $C_Y = C_Y(s, \tau_0)$ such that for all $\varepsilon \in (0,1]$ the following holds:* (2.17)–(2.18) *with the initial conditions $Y_A|_{\tau=0} = Y_B|_{\tau=0} = 0$ has a unique solution $Y_A, Y_B \in C([0,\tau_0], H^s)$. This solution satisfies*

$$\sup_{\tau \in [0,\tau_0]} \left\| \begin{pmatrix} Y_A \\ Y_B \end{pmatrix} \right\|_{H^s} \leq C_Y.$$

Proof. We rewrite (2.17),(2.18) as

$$\partial_\tau \begin{pmatrix} Y_A \\ Y_B \end{pmatrix} = \mathcal{L} \begin{pmatrix} Y_A \\ Y_B \end{pmatrix} + \frac{i}{2} \begin{pmatrix} \omega_A^{-1} G_A \\ \omega_B^{-1} G_B \end{pmatrix}, \qquad \mathcal{L} = \frac{1}{2i} \begin{pmatrix} \frac{v_1 \cos(3\vartheta_A) - c_A^2}{\omega_A} \partial_{\xi_A}^2 & 0 \\ 0 & \frac{v_1 \cos(3\vartheta_B) - c_B^2}{\omega_B} \partial_{\xi_B}^2 \end{pmatrix}.$$

By Fourier transform we obtain $\left\| \mathrm{e}^{\tau \mathcal{L}} \right\|_{L(H^s, H^s)} \leq 1$, and the remainder of the proof now works exactly as in [CBSU07, Lemma 4.2]. $\square$

Remark 2.4. The only terms which might create problems are those which appear in equations (2.17) or (2.18) at the order $\mathcal{O}(\varepsilon^{-1})$. The main statement of this lemma is that despite those terms, Y_A and Y_B are of order $\mathcal{O}(1)$.

2.4 Summary

The ansatz used to describe the pulse interaction is

$$\begin{aligned} Z_j(t) &= Z_j^A(t) + Z_j^B(t) + \varepsilon^3 M_j(t), \\ Z_j^A(t) &= \varepsilon A_1 \boldsymbol{E}_A + \varepsilon^3 A_3 \boldsymbol{E}_A^3 + \varepsilon^3 Y_A \boldsymbol{E}_A + \varepsilon^4 A_{4,3} \boldsymbol{E}_A^3 + \varepsilon^5 A_{5,5} \boldsymbol{E}_A^5 + \varepsilon^5 A_{5,3} \boldsymbol{E}_A^3 + \mathrm{cc}, \\ Z_j^B(t) &= \varepsilon B_1 \boldsymbol{E}_B + \varepsilon^3 B_3 \boldsymbol{E}_B^3 + \varepsilon^3 Y_B \boldsymbol{E}_B + \varepsilon^4 B_{4,3} \boldsymbol{E}_B^3 + \varepsilon^5 B_{5,5} \boldsymbol{E}_B^5 + \varepsilon^5 B_{5,3} \boldsymbol{E}_B^3 + \mathrm{cc}, \\ \boldsymbol{E}_A(t,j) &= \mathrm{e}^{i(\omega_A t + \vartheta_A j + \varepsilon \Omega_A(\xi_B))}, \quad \boldsymbol{E}_B(t,j) = \mathrm{e}^{i(\omega_B t + \vartheta_B j + \varepsilon \Omega_B(\xi_A))}, \end{aligned}$$

and yields the following hierarchy of equations to minimize the residual:

- the dispersion relation $\omega_A^2(\vartheta_A) = 2v_1(1 - \cos\vartheta_A) + w_1 + \varepsilon^2 \gamma_A$ and the group speed $c_A = -\omega_A'$, and similar for ω_B, c_B;
- at $\mathcal{O}(\varepsilon^3 \boldsymbol{E}_A)$ resp. $\mathcal{O}(\varepsilon^3 \boldsymbol{E}_B)$ the uncoupled NLS equations (2.16), i.e.,

  $$\begin{aligned} 0 &= \gamma_A A_1 + \omega_A \omega_A'' \partial_{\xi_A}^2 A_1 - [12 v_3 (1 - \cos\vartheta_A)^2 + 3w_3] |A_1|^2 A_1 + \mathcal{O}(\varepsilon), \\ 0 &= \gamma_B B_1 + \omega_B \omega_B'' \partial_{\xi_B}^2 B_1 - [12 v_3 (1 - \cos\vartheta_B)^2 + 3w_3] |B_1|^2 B_1 + \mathcal{O}(\varepsilon). \end{aligned}$$

  together with explicit expressions

  $$\Omega_A(\xi_B) = \Omega_A^0 + \int_{-\infty}^{\xi_B} \frac{-s_A^{(1)}}{c_A - c_B} |B_1|^2 \mathrm{d}\tilde{\xi}_B, \quad \Omega_B(\xi_A) = \Omega_B^0 + \int_{-\infty}^{\xi_A} \frac{-s_B^{(1)}}{c_B - c_A} |A_1|^2 \mathrm{d}\tilde{\xi}_A,$$

  for the phase corrections;
- two evolution equations (2.17), (2.17) for the lowest order corrections Y_A, Y_B, with $\mathcal{O}(1)$ solutions, and a number of algebraic equations like (2.19) for higher order terms as A_3, B_3 and the mixed terms contained in M; these can all be solved if (FORM) holds for all $k \leq 5$, which henceforth we *always assume.*

3 Justification of the Approximation

3.1 The Phase Space

In the previous section, we formally derived the macroscopic limit equations to be satisfied by the modulating functions A_1, B_1,etc. from the ansatz (2.11). It remains to prove that approximate solutions of (2.1) in the form (2.11) actually exist, i.e. to prove that the error stays small on a sufficiently long time scale. Therefore we need a phase space for (2.1).

Following [GM04], we rewrite (2.1) in a phase space Y which is isomorphic to $\ell^2 \times \ell^2$ as

$$\dot{\tilde{x}} = \mathcal{L}\tilde{x} + \mathcal{N}(\tilde{x}) \qquad \text{with} \qquad \tilde{x} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}, \tag{3.1}$$

where $\mathcal{L}$ is the linear part and $\mathcal{N}$ is the nonlinear part of the system. This yields

$$\begin{aligned} (\mathcal{L}\tilde{x})_j &:= (\dot{x}_j, L_j x) \quad \text{with} \quad L_j x = v_1(\partial_j^+ x - \partial_j^- x) - w_1 x_j, \\ (\mathcal{N}(\tilde{x}))_j &:= (0, N_j(x)) \quad \text{with} \quad N_j(x) = \widetilde{V}'(\partial_j^+ x) - \widetilde{V}'(\partial_j^- x) - \widetilde{W}'(x_j). \end{aligned} \tag{3.2}$$

The space Y is equipped with the norm $\|\cdot\|_Y$ defined by $\|(x,y)\|_Y^2 = \|x\|_E^2 + \|y\|_{\ell^2}^2$, where the energy norm $\|\cdot\|_E$ is defined by

$$\|x\|_E^2 := \sum_{j\in\mathbb{Z}} \left(v_1|\partial_j^+ x|^2 + w_1|x_j|^2\right) = v_1 \sum_{j\in\mathbb{Z}} |\partial_j^+ x|^2 + w_1 \|x\|_{\ell^2}^2 .$$

Due to the stability condition $\min\{w_1, w_1 + 4v_1\} > 0$, the energy norm $\|\cdot\|_E$ and the standard norm on ℓ^2 are equivalent with

$$\min\{w_1, w_1 + 4v_1\} \|x\|_{\ell^2}^2 \leq \|x\|_E^2 \leq \max\{w_1, w_1 + 4v_1\} \|x\|_{\ell^2}^2 .$$

Additionally we have the embedding

$$\|x\|_\infty \leq C_{\text{Norm}} \|\tilde{x}\|_Y . \tag{3.3}$$

The oscillator chain is a standard Hamiltonian system on $\ell^2 \times \ell^2$ with Hamiltonian

$$\mathcal{H}(x, \dot{x}) = \frac{1}{2} \|\dot{x}\|_{\ell^2}^2 + \sum_{j\in\mathbb{Z}} \left[V(\partial_j^+ x) + W(x_j)\right] .$$

The squared norm $\|\cdot\|_Y^2$ is twice the quadratic part of $\mathcal{H}$. This yields the following result [GM04].

Lemma 3.1. *The solutions $\tilde{x} : t \mapsto \tilde{x}(t) = \mathrm{e}^{t\mathcal{L}}\tilde{x}(0)$ of the linearized system* (2.2) *satisfy $\|\tilde{x}(t)\|_Y = \|\tilde{x}(0)\|_Y$ for all $t \in \mathbb{R}$.*

Proof. Since the linearized system reads $\ddot{x}_j - v_1(x_{j+1} - 2x_j + x_{j-1}) + w_1 x_j = 0$ we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \|\tilde{x}(t)\|_Y^2 &= \frac{\mathrm{d}}{\mathrm{d}t} \sum_{j\in\mathbb{Z}} \left[\dot{x}_j^2 + v_1(x_{j+1} - x_j)^2 + w_1 x_j^2\right] \\ &= 2 \sum_{j\in\mathbb{Z}} \dot{x}_j \left[\ddot{x}_j - v_1(x_{j+1} - 2x_j + x_{j-1}) + w_1 x_j\right] = 0. \end{aligned}$$

□

We can now estimate the residual of the formal approximation Z in the space Y, where compared to the formally obtained result $\rho(Z) = \mathcal{O}(\varepsilon^6)$ we lose the order $\mathcal{O}(\varepsilon^{1/2})$ due to the scaling properties of the Y-Norm.

Lemma 3.2. *There exists a C_{Res} such that for all $\varepsilon \leq \varepsilon_0$, ε_0 sufficiently small, the following holds. For the ansatz Z according to* (2.11) *we have*

$$\|\tilde{\rho}\|_Y = \left\| \begin{pmatrix} 0 \\ \rho \end{pmatrix} \right\|_Y = \|\rho\|_{\ell^2} \leq C_{\mathrm{Res}} \varepsilon^{11/2}.$$

Proof. See section 3.3.

3.2 Approximation of a Solution

Recall that we always assume (FORM) for all $k \leq 5$. We claim that initial data for (2.1) which may be approximated well by the ansatz Z at time $t = 0$, cf. (2.11), gives a solution with this property over macroscopic time scales. We start with

Lemma 3.3. *Let Z be the approximation (2.11), cf. the summary on page 10, with parameters $\vartheta_A, \vartheta_B \in (-\pi, \pi]$, $\vartheta_A \neq \vartheta_B$, $\gamma_A, \gamma_B \neq 0$ according to Lemma 2.1. Let $\tilde{Z} = (Z, \dot{Z})$, and let $\tau_0 > 0$. Then for every $C_A > 0$, there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following holds: If*

$$\left\| \tilde{U}(0) - \tilde{Z}(0) \right\|_Y \leq C_A \varepsilon^{7/2}, \tag{3.4}$$

then there exists a unique solution $\tilde{U} \in C([0, \tau_0/\varepsilon^2), Y)$ of (2.1)*, and*

$$\left\| \tilde{U}(t) - \tilde{Z}(t) \right\|_Y \leq C\varepsilon^{7/2} \text{ for all } t \in [0, \tau_0/\varepsilon^2]. \tag{3.5}$$

Proof. See Section 3.3.

In order to state the main theorem, we need the following definition:

Definition 3.4. *Let*

$$\begin{aligned} X^A(t,j) &:= \varepsilon A_1(\varepsilon(j - c_A t)) \boldsymbol{E}_A(t,j) + \varepsilon^3 A_3(\varepsilon(j - c_A t)) \boldsymbol{E}_A(t,j)^3 + \mathrm{cc}, \\ X^B(t,j) &:= \varepsilon B_1(\varepsilon(j - c_B t)) \boldsymbol{E}_B(t,j) + \varepsilon^3 B_3(\varepsilon(j - c_B t)) \boldsymbol{E}_B(t,j)^3 + \mathrm{cc}, \end{aligned}$$

with A_1 and B_1 chosen as homoclinic solutions of (2.7)*, and $\boldsymbol{E}_A, \boldsymbol{E}_B$, A_3, B_3 and Ω_A, Ω_B chosen as above. Then*

$$\Xi_j(t) = \Xi(t,j) = X^A(t,j) + X^B(t,j)$$

is called a well-formed approximation of the pulse interaction.

The purpose of this definition is to have an approximation of the pulse interaction without the correction functions from the ansatz (2.11). The following theorem states that we may approximate a pulse interaction even without those higher order corrections.

Theorem 3.5. *For every $C_1, C_{\mathrm{Dist}}, \tau_0, \delta > 0$, there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following holds. Assume that the initial condition for $\tilde{U}$ is chosen in such a way that it may be approximated by a well-formed approximation Ξ of the pulse interaction with the parameters $\vartheta_A, \vartheta_B \in (-\pi, \pi]$, $\vartheta_A \neq \vartheta_B$, $\gamma_A, \gamma_B \neq 0$ according to Lemma 2.1, and where, moreover, the initial distance fulfills, for some $\delta > 0$,*

$$|\xi_A^{(0)} - \xi_B^{(0)}| \geq C_{\mathrm{Dist}} \varepsilon^{-(1+\delta)}.$$

Then there exists a unique solution $\tilde{U} \in C([0, \tau_0/\varepsilon^2), Y)$ of (2.1), and can be well approximated by Ξ on the long time scale. In detail, if

$$\left\| \tilde{U}(0) - \tilde{\Xi}(0) \right\|_Y \leq C_1 \varepsilon^{7/2}, \tag{3.6}$$

then

$$\|U(t) - \Xi(t)\|_\infty \leq C_2 \varepsilon^3 \text{ for } t \in [0, \tau_0/\varepsilon^2]. \tag{3.7}$$

From Theorem 3.5 we also obtain:

Theorem 3.6. *Under the conditions of Theorem 3.5 the envelope shifts between U and Ξ are at most of order $\mathcal{O}(\varepsilon)$.*

Proof. Note that $\Xi = \mathcal{O}(\varepsilon)$ and therefore (due to the initial approximation) $U = \mathcal{O}(\varepsilon)$.

Since we have already dealt with the microscopic phase shift of the approximation, we only have to deal with the macroscopic functions. In order to simplify the notation, we write $U = U(X)$ and $\Xi = \Xi(X)$ for $X = \varepsilon j$ and leave out the time-dependence.

We assume that the "vertical" error (i. e. the error in the supremum norm), which is of order $\mathcal{O}(\varepsilon^3)$ (see equation (3.7)), is caused by an amplitude shift σ, i.e.we have a shift function $j \mapsto j + \sigma$. In the macroscopic space scaling, this shift becomes $X \mapsto X + \varepsilon\sigma$, and we obtain the estimate

$$U(X) = \Xi(X + \varepsilon\sigma) + \mathcal{O}(\varepsilon^3).$$

From Theorem 3.5 we obtain

$$U(X) - \Xi(X) = \varepsilon\sigma\Xi' + (\varepsilon\sigma)^2\Xi''(X + \theta\varepsilon\sigma) \leq \tilde{C}_4\varepsilon^3$$

for some $\theta \in (0, 1)$, and since $\Xi' \sim \varepsilon$ (i.e.$\lim_{\varepsilon\to 0} \frac{\Xi'}{\varepsilon} \in (0, \infty)$), we immediately get $\sigma = \mathcal{O}(\varepsilon)$. □

Remark 3.7. Similar to [CBCSU08] for the PDE case, the above analysis can be generalized to general modulating pulses, i.e., to A_1, B_1 not necessary in the form of NLS solitons but rather general spatially localized solutions of the (time-dependent) NLS. For this, the ansatz (2.11) has to be modified to

$$Z_j(t) = (\varepsilon A_1 + \varepsilon^2 A_2 + \varepsilon^3 A_3)\boldsymbol{E}_A + (\varepsilon B_1 + \varepsilon^2 B_2 + \varepsilon^3 B_3)\boldsymbol{E}_B + \mathrm{cc} + \varepsilon^3 M_{\mathrm{Mixed}},$$

where $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ are also modified such that the phase shifts Ω_A and Ω_B depend not only on the macroscopic space variables, but also on the macroscopic time. As can be seen, an additional term of order $\mathcal{O}(\varepsilon^2)$ is added at the basic frequencies $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$. Since the terms at mixed frequencies occur due to the nonlinear interaction, their lowest order remains $\mathcal{O}(\varepsilon^3)$.

Remark 3.8. Our analysis can also be extended to quadratic nonlinearities. The idea is that under (additional) non–resonance conditions and near the basic waves $\boldsymbol{E}_A(t,j)$ and $\boldsymbol{E}_B(t,j)$ the quadratic terms in (2.1) can be eliminated by normal form transforms see, e.g. [Sch98, GM06].

3.3 Proofs

Proof of Lemma 3.2. There is a constant C_{MF} such that for $\tau \in [0,\tau_0]$ and $\kappa+\lambda+2\mu \le 6$, all macroscopic functions H in (2.11) satisfy $\left\|\partial_{\xi_A}^{\kappa}\partial_{\xi_B}^{\lambda}\partial_{\tau}^{\mu}H(\tau,\cdot)\right\|_{L^2} \le C_{\mathrm{MF}}$, respectively

$$\left\|\partial_{\xi_A}^{\kappa}\partial_{\xi_B}^{\lambda}\partial_{\tau}^{\mu}H(\tau,\xi_A,\cdot)\right\|_{L^2} \le C_{\mathrm{MF}}, \quad \left\|\partial_{\xi_A}^{\kappa}\partial_{\xi_B}^{\lambda}\partial_{\tau}^{\mu}H(\tau,\cdot,\xi_B)\right\|_{L^2} \le C_{\mathrm{MF}}.$$

if H depends on both ξ_A and ξ_B. This follows first for $H = A_1$ or $H = B_1$ from equation (2.9) and subsequently for all further macroscopic functions from their definitions. Thus, for $\tau \in [0,\tau_0]$ and $\kappa+\lambda+2\mu \le 5$ we have

$$\left\|\partial_{\xi_A}^{\kappa}\partial_{\xi_B}^{\lambda}\partial_{\tau}^{\mu}H(\tau,\xi_A,\xi_B)\right\|_{H^1(\mathbb{R})} \le 2C_{\mathrm{MF}}. \tag{3.8}$$

Due to the Sobolev embedding theorem

$$\|u(\cdot)\|_{\infty} \le C_{\mathrm{Sob}}\,\|u\|_{H^1(\mathbb{R})}$$

we obtain for all types of modulation functions

$$\sup_{\tau\in[0,\tau_0]}\left\|\partial_{\xi_A}^{\kappa}\partial_{\xi_B}^{\lambda}\partial_{\tau}^{\mu}H(\tau,\cdots)\right\|_{\infty} < C_{\mathrm{MF}}C_{\mathrm{Sob}} < \infty \tag{3.9}$$

for $\kappa+\lambda+2\mu \le 5$ and a constant C_{Sob}.

Now let $T(t)$ an arbitrary term of the residual ρ. We first assume that T stems from the linear part of (2.1). Then T must have the form

$$T(t) = \varepsilon^k \widetilde{C} T^{\Omega}(t)\tilde{H}\begin{Bmatrix}\boldsymbol{E}_A^n\\ \boldsymbol{E}_B^n\end{Bmatrix},$$

where $k \ge 6$, $|n| \le k$, $\widetilde{C} > 0$ and

$$\tilde{H} = H(\tau,\varepsilon(j-c_At+\theta_1),\varepsilon(j-c_Bt+\theta_2)) \tag{3.10}$$

for a macroscopic function H and $|\theta_{1,2}| \le 1$, and where $T^{\Omega}(t,x)$ either equals 1 or is a term which is caused by the series expansion of $\mathrm{e}^{i\varepsilon(\Omega_{A,B}(\xi_{B,A}\pm\varepsilon)-\Omega_{A,B}(\xi_{B,A}))}$. Since e.g.

$$\Omega_A(\xi_B+\varepsilon)-\Omega_A(\xi_B) = \int_{\xi_B}^{\xi_B+\varepsilon}\frac{-s_A^{(1)}}{c_A-c_B}|B_1(\tilde{\xi}_B)|^2\mathrm{d}\tilde{\xi}_B \le \frac{-s_A^{(1)}}{c_A-c_B}\,\|B_1\|_{L_2}^2\,,$$

and due to equation (3.8), there is a constant C_{Ω} such that

$$\sup_{\tau\in[0,\tau_0]}\left\|T^{\Omega}(\tau,\cdot)\right\|_{\infty} \le C_{\Omega}$$

for any such term T^Ω.

Now we can estimate T. We restrict to $H = H(\tau, \xi_A)$ and obtain

$$\begin{aligned} \|T(t)\|_{\ell^2} &\leq \varepsilon^k \widetilde{C} C_\Omega \left(\sum_{j\in\mathbb{Z}} \max_{H\in\mathbb{M}} \max_{\kappa+2\mu\leq 5} \sup_{|\theta|<1} \partial_{\xi_A}^\kappa \partial_\tau^\mu H(\tau, \varepsilon(j - c_A t + \theta)) \right)^{1/2} \\ &\overset{\star}{\leq} \varepsilon^k \widetilde{C} C_\Omega \left(\frac{8}{\varepsilon} \max_{H\in\mathbb{M}} \max_{\kappa+2\mu\leq 5} \left\| \partial_{\xi_A}^\kappa \partial_\tau^\mu H \right\|_{H^1}^2 \right)^{1/2} \leq \varepsilon^k \widetilde{C} C_\Omega \left(\frac{8}{\varepsilon} (2C_{\mathrm{MF}}^2) \right)^{1/2}. \end{aligned}$$

The estimate $\star$ is based on Lemma 3.9 below. Since we assumed $k \geq 6$, we obtain, for all $\varepsilon \in (0, 1)$,

$$\|T(t)\|_{\ell^2} \leq \varepsilon^{11/2} \widetilde{C} C_\Omega 4 C_{\mathrm{MF}}.$$

A similar estimate holds if $H = H(\tau, \xi_B)$ or $H = H(\tau, \xi_A, \xi_B)$. Finally, we get $\|T(t)\|_{\ell^2} \leq \varepsilon^{11/2} \hat{C}$, which is what we wanted to show.

We estimate the terms coming from other parts of (2.1) in a similar way: A term from the nonlinearity has the form

$$T(t) = \varepsilon^k \widetilde{C} T_1^\Omega(t) T_2^\Omega(t) T_3^\Omega(t) \tilde{H}_1 \tilde{H}_2 \tilde{H}_3 \boldsymbol{E}_A^m \boldsymbol{E}_B^n,$$

where the $\tilde{H}_i$ are defined as in (3.10). We estimate $\tilde{H}_1$ and $\tilde{H}_2$ by their supremum and treat $\tilde{H}_3$ as above, obtaining

$$\|T(t)\|_{\ell^2} \leq \varepsilon^k \widetilde{C} C_\Omega^3 \left(\frac{8}{\varepsilon} \right)^{1/2} \left(4 C_{\mathrm{Sob}}^2 C_{\mathrm{MF}}^2 2 C_{\mathrm{MF}} \right).$$

Since $k \geq 6$, we obtain for all $\varepsilon \in (0, 1)$ that $\|T(t)\|_{\ell^2} \leq \varepsilon^{11/2} \hat{C}$ as above.

Terms from the left-hand side of (2.1) are estimated similarly, and we finally get

$$\|\tilde{\rho}\|_Y = \left\| \begin{pmatrix} 0 \\ \rho \end{pmatrix} \right\|_Y = \|\rho\|_{\ell^2} \leq \varepsilon^{11/2} C_{\mathrm{Res}}. \tag{3.11}$$

for the residual and for sufficiently small ε. □

We used the following lemma, which describes the scaling of the ℓ^2-norm under the transformation $x \mapsto \varepsilon x$. This is an extended version of Prop. 3.3 from [GM04].

Lemma 3.9. *For $\phi \in H^1(\mathbb{R})$, $\varepsilon \in (0, 1)$ and $c \in \mathbb{R}$ we have*

$$\sum_{j\in\mathbb{Z}} \sup_{|\theta|\leq 1} |\phi(\varepsilon(j + c + \theta))|^2 \leq \frac{8}{\varepsilon} \|\phi\|_{H^1(\mathbb{R})}^2.$$

For $\phi \in H^2(\mathbb{R}^2)$, $\varepsilon \in (0, 1)$ and $c \in \mathbb{R}$ we have

$$\sum_{j\in\mathbb{Z}} \sup_{|\theta|\leq 1} |\phi(\varepsilon j, \varepsilon(j + c + \theta))|^2 \leq \frac{32}{\varepsilon} \|\phi\|_{H^2(\mathbb{R}^2)}^2.$$

Proof. The proof of the first part is repeated for convenience from [GM04]. Let $\phi = \phi(x) \in H^1(\mathbb{R})$, $j \in \mathbb{Z}$ and $x, \tilde{x} \in (j + c - 1, j + c + 1)$. From the fundamental theorem of calculus we obtain

$$|\phi(x)| \leq |\phi(\tilde{x})| + \int_{j+c-1}^{j+c+1} |\partial_x \phi(\xi)| \mathrm{d}\xi.$$

Integration over $\tilde{x}$, the estimate $(a+b)^2 \le 2(a^2+b^2)$, and Cauchy-Schwarz inequality yield

$$|\phi(x)| \le \int_{j+c-1}^{j+c+1} (|\phi(\xi)| + |\partial_x \phi(\xi)|)\mathrm{d}\xi \le \sqrt{2}\int_{j+c-1}^{j+c+1} (|\phi(\xi)|^2 + |\partial_x\phi(\xi)|^2)^{1/2}\mathrm{d}\xi$$
$$\le 2\left(\int_{j+c-1}^{j+c+1} (|\phi(\xi)|^2 + |\partial_x\phi(\xi)|^2)\mathrm{d}\xi\right)^{1/2}$$

and thus $\sup_{|\theta|\le 1} |\phi(\varepsilon(j+c+\theta))|^2 \le 4\,\|\phi(\varepsilon\cdot)\|^2_{H^1((j+c-1,j+c+1))}$. Summing over $j \in \mathbb{Z}$, we obtain

$$\sum_{j\in\mathbb{Z}} \sup_{|s|\le 1} |\phi(\varepsilon(j+c+\theta))|^2 \le 8\,\|\phi(\varepsilon\cdot)\|_{H^1(\mathbb{R})}\,.$$

The substitution $\xi = \varepsilon x$ yields

$$\|\phi(\varepsilon\cdot)\|_{H^1(\mathbb{R})} = \int_{x\in\mathbb{R}} (|\phi(\varepsilon x)|^2 + |\partial_x\phi(\varepsilon x)|^2)\mathrm{d}x$$
$$= \frac{1}{\varepsilon}\int_{\xi\in\mathbb{R}} (|\phi(\xi)|^2 + \varepsilon^2|\partial_x\phi(\xi)|^2)\mathrm{d}\xi \le \frac{1}{\varepsilon}\,\|\phi\|^2_{H^1(\mathbb{R})}$$

for $\varepsilon \in (0,1)$, which is the desired estimate.

In order to prove the second part, we assume that $\phi \in H^2(\mathbb{R}^2)$, $x \in \mathbb{R}$, $y \in [x+\varepsilon c-\varepsilon, x+\varepsilon c+\varepsilon]$. Then as above,

$$|\phi(x,y)| \le 2\left(\int_{x+\varepsilon c-\varepsilon}^{x+\varepsilon c+\varepsilon} |\phi(x,\eta)|^2 + |\partial_y\phi(x,\eta)|^2\mathrm{d}\eta\right)^{1/2} =: \psi_\varepsilon(x).$$

We have

$$\|\psi_\varepsilon(\cdot)\|^2_{H^1} = 4\int_{-\infty}^{\infty}\int_{x+\varepsilon c-\varepsilon}^{x+\varepsilon c+\varepsilon} |\phi|^2 + |\partial_y\phi|^2 + |\partial_x\phi|^2 + |\partial_{xy}\phi|^2\mathrm{d}y\mathrm{d}x$$
$$\le 4\,\|\phi\|^2_{H^2(\mathbb{R}^2)} = \mathcal{O}(1).$$

Application of the first part yields

$$\sum_{j\in\mathbb{Z}} |\psi_\varepsilon(\varepsilon j)|^2 \le \frac{8}{\varepsilon}\,\|\psi_\varepsilon\|^2_{H^1} \le \frac{32}{\varepsilon}\,\|\phi\|^2_{H^2(\mathbb{R}^2)}\,.$$

Therefore we get the desired estimate

$$\sum_{j\in\mathbb{Z}} \sup_{|\theta|\le 1} |\phi(\varepsilon j, \varepsilon(j+c+\theta))|^2 \le \frac{32}{\varepsilon}\,\|\phi\|^2_{H^2(\mathbb{R}^2)}\,.$$

□

Proof of Lemma 3.3. We define $\tilde{R} = \begin{pmatrix} R \\ \dot{R} \end{pmatrix} = \varepsilon^{-7/2}(\tilde{U} - \tilde{Z})$. Our goal is to show that $\tilde{R}$ remains uniformly bounded for $\tau \le \tau_0$. Using the definition of the residual (2.10) we obtain the error equation

$$\dot{\tilde{R}} = \mathcal{L}\tilde{R} + \begin{pmatrix} 0 \\ M \end{pmatrix} + \varepsilon^{-7/2}\tilde{\rho} \tag{3.12}$$

with

$$\begin{pmatrix} 0 \\ M \end{pmatrix} = \varepsilon^{-7/2} \left[\mathcal{N}(\varepsilon^{7/2}\tilde{R} + \tilde{Z}) - \mathcal{N}(\tilde{Z}) \right],$$

where $(\mathcal{L}\tilde{x})_j = (\dot{x}_j, L_j x)$, $(\mathcal{N}(\tilde{x}))_j = (0, N_j(x))$, cf. (3.2), and

$$\tilde{\rho} = \begin{pmatrix} 0 \\ \rho \end{pmatrix} = \begin{pmatrix} 0 \\ L_j Z + N_j(Z) - \ddot{Z}_j \end{pmatrix}.$$

First we estimate M with

$$\begin{aligned} \varepsilon^{7/2} M_j = {} & \widetilde{V}'(\varepsilon^{7/2}\partial_j^+ R + \partial_j^+ Z) - \widetilde{V}'(\partial_j^+ Z) - \widetilde{V}'(\varepsilon^{7/2}\partial_j^- R + \partial_j^- Z) \\ & + \widetilde{V}'(\partial_j^- Z) - \widetilde{W}'(\varepsilon^{7/2} R_j + Z_j) + \widetilde{W}'(Z_j). \end{aligned}$$

The mean value theorem gives

$$M_j = \widetilde{V}''(\varepsilon d_j^+)\partial_j^+ R - \widetilde{V}''(\varepsilon d_j^-)\partial_j^- R - \widetilde{W}''(\varepsilon d_j) R_j \tag{3.13}$$

with $d_j^\pm = \theta_j^\pm \varepsilon^{5/2} \partial_j^\pm R + \frac{1}{\varepsilon}\partial_j^\pm Z$ and $d_j = \theta_j \varepsilon^{5/2} R_j + \frac{1}{\varepsilon} Z_j$, where $\theta_j^\pm, \theta_j \in (0,1)$. From (3.9) and Sobolev's embedding theorem we get the existence of a constant C_Z such that $\|Z\|_\infty \le \varepsilon C_Z$ for $\varepsilon \in (0,1)$. Thus

$$|d_j^\pm|, |d_j| \le \varepsilon^{5/2}(|R_{j+1}| + |R_j| + |R_{j-1}|) + 2C_Z \le 3\varepsilon^{5/2} \left\| \tilde{R} \right\|_Y + 2C_Z$$

for all $j \in \mathbb{Z}$, $\varepsilon < \varepsilon_0$ and $\varepsilon^2 t \le \tau_0$.

We assume for now that for a given $D > 0$ we have $\left\| \tilde{R} \right\|_Y \le D$ such that $|d_j^\pm|, |d_j| \le 2C_Z + 3\varepsilon^{5/2} D$ for all $j \in \mathbb{Z}$, $\varepsilon < \varepsilon_0$, $\varepsilon^2 t \le \tau_0$. Next we assume that, by choice of $\varepsilon_0 > 0$,

$$2C_Z + 3\varepsilon^{5/2} D \le 2C_Z + 1 = C_d. \tag{3.14}$$

By the cubic form of the nonlinearity, i.e. $\widetilde{V}''(d) = 3v_3 d^2 + \mathcal{O}(d^3)$ and $\widetilde{W}''(y) = 3w_3 y^2 + \mathcal{O}(y^3)$, (3.13) implies

$$M_j = 3\varepsilon^2 \left[v_3 (d_j^+)^2 \partial_j^+ R - v_3 (d_j^-)^2 \partial_j^- R - w_3 (d_j)^2 R_j \right] + \mathcal{O}(\varepsilon^3)$$

and thus, for sufficiently small $\varepsilon_0 > 0$,

$$|M_j| \le \varepsilon^2 \hat{C}(D)(|R_{j+1}| + |R_j| + |R_{j-1}|)$$

with $\hat{C}(D) := 6(2|v_3| + |w_3|)C_d^2$. For $\left\| \tilde{R} \right\|_Y \le D$, $\varepsilon < \varepsilon_0$ and $\varepsilon^2 t \le \tau_0$ this implies

$$\left\| \begin{pmatrix} 0 \\ M \end{pmatrix} \right\|_Y = \|M\|_{\ell^2} \le \varepsilon^2 \hat{C}(D) \left\| \tilde{R} \right\|_Y. \tag{3.15}$$

The error equation (3.12) for $\tilde{R}$ may be rewritten with the variation of constant formula as

$$\tilde{R}(t) = G(t)\tilde{R}(0) + \int_0^t G(t-s) \left(\begin{pmatrix} 0 \\ M(s) \end{pmatrix} + \varepsilon^{-7/2} \tilde{\rho}(s) \right) \mathrm{d}s,$$

where $G(t) = \mathrm{e}^{t\mathcal{L}}$ is the evolution operator of the linearized system (2.2). By (3.4) we have $\left\|\tilde{R}(0)\right\|_Y \leq C_A$. This implies, together with Lemma 3.2, (3.15) and Lemma 3.1 that

$$\begin{aligned}\left\|\tilde{R}(t)\right\|_Y \leq \|G(t)\|_{Y\to Y} \left\|\tilde{R}(0)\right\|_Y \\ + \int_0^t \|G(t-s)\|_{Y\to Y}\, \varepsilon^2 \left(\hat{C}\left\|\tilde{R}(s)\right\|_Y + \varepsilon^{-11/2}\tilde{\rho}(s)\right) \mathrm{d}s \\ \leq C_A + \varepsilon^2 \left(\int_0^t \hat{C}\left\|\tilde{R}(s)\right\|_Y \mathrm{d}s + tC_{\mathrm{Res}}\right)\end{aligned}$$

for $0 < \varepsilon \leq \varepsilon_0,\ 0 \leq \varepsilon^2 t \leq \tau_0$ and $\left\|\tilde{R}\right\|_Y \leq D$. Gronwall's inequality now yields

$$\left\|\tilde{R}(t)\right\|_Y \leq (2d + \varepsilon^2 tC_\rho)\mathrm{e}^{\varepsilon^2 t\hat{C}} \leq (2d + \tau_0 C_\rho)\mathrm{e}^{\tau_0\hat{C}} =: D \tag{3.16}$$

for $\varepsilon^2 t \leq \tau_0$, $\varepsilon \leq \varepsilon_0$. By choice of ε_0 we may fulfill (3.14), and this implies that (3.16) holds for all $t \in [0, \tau_0/\varepsilon^2]$, which proves (3.5), i.e. Lemma 3.3. □

Proof of Theorem 3.5. The idea is to rewrite the approximation (2.11) as

$$Z_j(t) = \Xi_j(t) + \varepsilon^3\Upsilon_j(t) + \varepsilon^4\Lambda_j(t),$$

with the following grouping of the macroscopic functions:

$$\begin{aligned}\Xi_j(t) &= X^A(t,j) + X^B(t,j) = \varepsilon A\boldsymbol{E}_A + \varepsilon^3 A_3\boldsymbol{E}_A^3 + \varepsilon B\boldsymbol{E}_B + \varepsilon^3 B_3\boldsymbol{E}_B^3 + \mathrm{cc} \\ \varepsilon^3\Upsilon_j(t) &= \varepsilon^3 Y_A\boldsymbol{E}_A + \varepsilon^3 Y_B\boldsymbol{E}_B + \mathrm{cc} + \varepsilon^3 M, \\ \varepsilon^4\Lambda_j(t) &= \varepsilon^4 A_{4,3}\boldsymbol{E}_A^3 + \varepsilon^5 A_{5,3}\boldsymbol{E}_A^3 + \varepsilon^5 A_{5,5}\boldsymbol{E}_A^5 \\ &\quad + \varepsilon^4 B_{4,3}\boldsymbol{E}_B^3 + \varepsilon^5 B_{5,3}\boldsymbol{E}_B^3 + \varepsilon^5 B_{5,5}\boldsymbol{E}_B^5 + \mathrm{cc}.\end{aligned}$$

Obviously we have $\varepsilon^3\Upsilon_j(t) = \mathcal{O}(\varepsilon^3)$ and $\varepsilon^4\Lambda_j(t) = \mathcal{O}(\varepsilon^4)$. From (3.3), (3.4), and the triangle inequality we have

$$\begin{aligned}\|U(t) - \Xi(t)\|_\infty &\leq \|U(t) - Z(t)\|_\infty + \left\|\varepsilon^3\Upsilon(t) + \varepsilon^4\Lambda(t)\right\|_\infty \\ &\leq C_{\mathrm{Norm}}\left\|\tilde{U}(t) - \tilde{Z}(t)\right\|_Y + \tilde{C}_1\varepsilon^3 \leq (C_{\mathrm{Norm}}C_A + \tilde{C}_1)\varepsilon^3.\end{aligned}$$

Thus if the conditions of Lemma 3.3 hold, then we obtain (3.7), which is the statement of the Theorem 3.5.

It remains to show that the assumptions of Lemma 3.3 follow from those of Theorem 3.5. At time $t = 0$ we have

$$\begin{aligned}\|\tilde{U}(0) - \tilde{Z}(0)\|_Y &\leq \left\|\tilde{U}(0) - \tilde{\Xi}(0)\right\|_Y + \left\|\varepsilon^3\Upsilon_j(0)\right\|_Y + \left\|\varepsilon^4\Lambda_j(t)\right\|_Y \\ &\leq C_1\varepsilon^{7/2} + \left\|\varepsilon^3(Y_A(0,j) + Y_B(0,j) + \mathrm{cc} + M(0,j))\right\|_Y + \tilde{C}_2\varepsilon^4 \\ &\leq C_A\varepsilon^{7/2}\end{aligned} \tag{3.17}$$

with a constant C_A, because by Lemma 2.3 we can choose $Y_A(0) \equiv Y_B(0) \equiv 0$, and the mixed terms may be estimated as follows: Each term in $\varepsilon^3 M$ with the order $\mathcal{O}(\varepsilon^3)$ has the form

$$\varepsilon^3 T = \varepsilon^3\alpha A_1^\kappa \bar{A}_1^\lambda B_1^\mu \bar{B}_1^\nu \boldsymbol{E}_A^{\kappa-\lambda}\boldsymbol{E}_B^{\mu-\nu}, \tag{3.18}$$

where $\kappa - \lambda \neq 0$ and $\mu - \nu \neq 0$. Since the two pulses are separated in the beginning, the product in equation (3.18) is exponentially small. Therefore we have $\|\varepsilon^3 M(0,j)\|_Y \leq \tilde{C}_3 \varepsilon^{7/2}$ for a suitable constant C_3. Thus, from (3.17) and (3.6) we have $\left\|\tilde{U}(0) - \tilde{Z}(0)\right\|_Y \leq C_1 \varepsilon^{7/2}$, and Lemma 3.3 gives the desired result. □

References

[AS81] M. J. Ablowitz and H. Segur. *Solitons and the inverse scattering transform.* SIAM, Philadelphia, 1981.

[BF06] A. Babin and A. Figotin. Linear Superposition in Nonlinear Wave Dynamics. *Rev. Math. Phys.*, 18(9):971–1053, 2006.

[CBCSU08] M. Chirilus-Bruckner, Ch. Chong, G. Schneider, and H. Uecker. Separation of Internal and Interaction Dynamics for NLS-described Wave Packets with Different Carrier Waves. *J. Math. Anal. Appl.*, 347(1):304-314, 2008.

[CBSU07] M. Chirilus-Bruckner, G. Schneider, and H. Uecker. On the Interaction of NLS-Described Modulating Pulses with Different Carrier Waves. *Math. Meth. Appl. Sci.*, 30:1965 – 1978, 2007.

[FPU55] E. Fermi, J. Pasta, and S. Ulam. Studies of nonlinear problems. Technical Report I., Los Alamos Rep, LA1940, Los Alamos, 1955. Reproduced in *Nonlinear Wave Motion*, A. C. Newell, ed. AMS, Providence, RI, 1974.

[GM04] J. Giannoulis and A. Mielke. The Nonlinear Schrödinger Equation as a macroscopic limit for an oscillator chain with cubic nonlinearities. *Nonlinearity*, 17(2):551–565, 2004.

[GM06] J. Giannoulis and A. Mielke. Dispersive evolution of pulses in oscillator chains with general interaction potentials. *Discrete and Continuous Dynamical Systems - Series B*, 6(3):493–523, 2006.

[McM05] E. McMillan. On the impact of FPU lattice discreteness upon solitary wave interactions. Physica D, 200:25-46, 2005.

[Sch98] G. Schneider. Justification of modulation equations for hyperbolic systems via normal forms. *NoDEA*, 5:69–82, 1998.

[SW00] G. Schneider and C.E. Wayne. Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. Fiedler, B. (ed.) et al., Proceedings of Equadiff '99, Vol. 1. Singapore: World Scientific. 390-404, 2000.

[TPB04] L. Tkeshelashvili, S. Pereira and K. Busch. General theory of nonresonant wave interaction: Giant soliton shift in photonic band gap materials *Europhysics Letter*, 68:205-211, 2004.

[Tsu89] T. Tsuboi, Phase shift in the collision of two solitons propagating in a nonlinear transmission line. *Phys. Rev. A*, 40(5):2753-2755, 1989.

[ZK65] N. J. Zabusky and M. D. Kruskal. Interactions of solitons in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 15:240–243, 1965.

Addresses of the authors:

Guido Schneider, Institut für Analysis, Dynamik und Modellierung, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart.

Hannes Uecker, Institut für Mathematik, Universität Oldenburg, D-26111 Oldenburg.

Michael Wand, Cognitive Systems Lab, Fakultät für Informatik, Universität Karlsruhe, Am Fasanengarten 5, D-76131 Karlsruhe.