Nonlinear Interaction of Pulses

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der Fakultät für Mathematik der Universität Karlsruhe (TH) genehmigte

DISSERTATION

von

Dipl.-Math. techn. Martina Chirilus-Bruckner

aus Vatra Dornei, Rumänien

Tag der mündlichen Prüfung: 23. Juli 2009

Referent:	Prof. Dr. Guido Schneider
Korreferent:	Prof. Dr. Michael Plum

Introduction

Consider the initial value problem for a nonlinear wave equation, say, the cubic Klein-Gordon equation

$$\partial_t^2 u = \partial_x^2 u - u + u^3,$$

where $x, t, u = u(x,t) \in \mathbb{R}$, with the initial profile (as depicted in Fig. 0.1) given by

$$\begin{aligned} u(x,t)|_{t=0} &= u_{pulse}(x,t)|_{t=0}, \\ \partial_t u(x,t)|_{t=0} &= \partial_t u_{pulse}(x,t)|_{t=0} \end{aligned}$$

with

$$u_{pulse}(x,t) = \varepsilon A_1(\varepsilon(x - c_1 t), \varepsilon^2 t) e^{ik_1 x + i\omega_1 t} + \varepsilon A_2(\varepsilon(x - c_2 t + d), \varepsilon^2 t) e^{ik_2 x + i\omega_2 t} + c.c.,$$

where $c_j, k_j, \omega_j, d \in \mathbb{R}, 0 < \varepsilon \ll 1$, and $A_j(\cdot, t)$ is a localized function for all $t \in \mathbb{R}$.

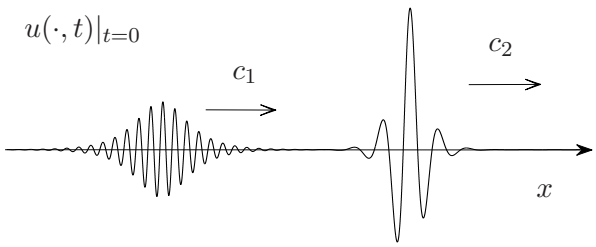

Figure 0.1: Initial profile given by two localized structures traveling with different velocities

Now, how does this initial profile evolve in time? How do the localized structures interact? Will a collision destroy them? What if they travel with the same speed? What if we take more than two?

The present work analyzes all these natural questions in the context of various nonlinear wave equations which arise in applications such as nonlinear fiber optics and photonics, where the localized structures represent light pulses which can be used to transport and process digital data (see Ch. 7).

In what follows, we will call these localized structures as "pulses". They constitute the central theme of this work.

Brief summary of results

Our main achievement can be stated as follows: By means of perturbation analysis we show that the two pulses do not interact at all in first order, i.e. they pass through each other undistorted. Hence, at

leading order the interaction is elastic, as one would expect from a linear equation. The nonlinearity shows its influence only at higher orders through interaction effects like a shift of the pulse carrier wave, a shift of the pulse envelope and a deformation of the pulse shape (see Fig. 2.3- Fig. 2.5).

We give a very precise analytical description of this interaction behavior by means of a reduction of the original problem to an initial value problem for a

- recursively solvable, decoupled "modulation system" for the envelopes and their higher order corrections and
- explicit formulas for the interaction effects.

Consequently, we completely separate the description of the internal dynamics of each pulse and the description of pure interaction dynamics (well, at least at leading order). At the top of this modulation system is a Nonlinear Schrödinger equation

$$\partial_2 A_j = \nu_{j,1}\partial_1^2 A_j + \nu_{j,2}|A_j|^2 A_j, \quad \nu_{j,1}, \nu_{j,1} \in \mathbb{R}, j = 1, 2,$$

for each of the two envelopes. The Nonlinear Schrödinger equation has extensively been studied and is known to exhibit stable, spatially localized solutions — so-called solitons — which can be specified in terms of elementary functions. This already indicates that the reduced system is much easier to handle both from an analytical and from a numerical point of view.

Outline of the work

This work is organized in two major parts:

Part I. Concepts of pulse interaction — Separation of internal and interaction dynamics

In Chapter 1 we give an introduction to the main analytical tools that we make use of throughout the work. We explain how to derive and rigorously justify the Nonlinear Schrödinger equation for envelopes of one-pulse solutions for the cubic Klein-Gordon equation. Most of the material has already been discussed in [KSM92].
Chapter 2 presents a refined perturbation approach that leads to the separation of internal and interaction dynamics for a multipulse in the setting of a cubic Klein-Gordon equation.

Part II. Pulse interaction in "natural coordinates" and applications

Chapter 3 illustrates a frame work which allows one to transfer the results from the first part of the work to a more general setting. In fact, this chapter forms the central piece of the present work.
In the next chapters, i.e. Chapter 4-6, we perform exactly this transfer for the Maxwell-Lorentz system and a nonlinear wave equation with periodic coefficients. We also revisit the cubic Klein-Gordon case and reproduce the results from Chapter 2 with the new tools developed in Chapter 3. In this sense, Chapter 4 links Part I and Part II of this work.
Finally, we discuss in Chapter 7 the consequences of the presented results for applications.

Note, that every chapter concludes with a "Chapter Summary" that briefly reviews the main points of the corresponding chapter.

Motivation of the problem

Our research was initiated by the work of Tkeshelashvili *et al* in [TPB04] where the above problem was posed for a nonlinear wave equation with periodic coefficients, which is used as a model for light propagation in nonlinear "photonic crystals" (see Ch. 7).
Photonic crystals consist of a periodic arrangement of materials with different refraction properties. They are believed to have the ability to support "standing light pulses" which could be used as optical storage (however, so far, this phenomenon has not been observed experimentally). But how can such standing pulses be detected? One of the ideas expressed in [TPB04] is that the detection can be realized through a collision with another pulse, which, however, requires a very deep understanding of the interaction behavior. For this reason [TPB04] contains a formal derivation of explicit formulas for the leading order interaction effects and a numerical validation of these formulas for special pulse shapes. Our work is an extension of these results. In particular, we can deal with arbitrary shaped pulses and derive an explicit formula for an additional interaction effect — the shape deformation.

Another motivation comes from the field of optical communications which is concerned with the transportation of digital data via light pulses (see Ch. 7). To increase the bit rate of the optical fiber that guides the light pulses, one can simultaneously send pulses on different carrier waves (and, hence, with different wavelengths), a technique called "Wavelength-Division Multiplexing". Since optical fibers exhibit a nonlinear behavior (such as an intensity dependent refraction index), one needs a precise understanding of the nonlinear interaction both between pulses within one channel and between pulses from different channels. Such a "Wavelength-Division Multiplexing" scenario can be described by a multipulse (see Fig. 2.6). In this work we managed to clearly separate the description of the internal and interaction dynamics associated with such a multipulse and characterize the relation between the most important parameters that are needed to optimize the multiplexing (see Ch. 2).

Acknowledgements

First and foremost I would like to thank Prof. Dr. Guido Schneider for his extensive advice, support and encouragement over the last several years and all the fruitful discussions concerning the contents of my thesis!
I would also like to thank Prof. Michael Plum for agreeing to be a member of my thesis committee and providing helpful comments and suggestions on my work.
Moreover, I would like to thank my colleague and friend Christopher Chong. I always enjoyed our discussions not only about mathematics, but also about what it means to be a PhD student!
I thank Prof. Dr. Hannes Uecker for his mathematical advice concerning this thesis.
I am very grateful to Prof. Dr. Kurt Busch and his research group for the stimulating discussions on the physical problem that initated my research.
Furthermore, I would like to thank Prof. Dr. Willy Dörfler for his constant support.
Finally, I want to thank Simon Friedberger who provided invaluable support during my turbulent time as a PhD student. His assistance with LaTeX issues is appreciated.

This work was supported by the German Research Foundation (DFG) through the GRK 1294: Analysis, Simulation and Design of Nanotechnological Processes.

Contents

List of Figures

Notation

- All constants that appear which are independent of the perturbation parameter ε are denoted by C.
- The slow spatial variable is defined by $\underline{X} := \varepsilon x$ and the slow temporal variable by $\underline{T} := \varepsilon^2 t$.
- For vectors V the notation $[V]_n$ means the n-th coordinate.
- Fourier transform is defined by $\mathcal{F}\{u\}(k) := \hat{u}(k) := \frac{1}{2\pi}\int_{\mathbb{R}} u(x)e^{-ikx}\,dx$.
- Sobolev spaces H^s are equipped with the norm

$$\|u\|_{H^s} = \left(\int_{\mathbb{R}} (1+|k|^2)^s |\hat{u}(k)|^2\,dk\right)^{1/2}.$$

- The spaces C_b^s are equipped with the norm

$$\|u\|_{C_b^s} = \sum_{j=0}^{s} \|\partial_x^j u\|_{C_b^0}, \quad \|u\|_{C_b^0} = \sup_{x\in\mathbb{R}} |u(x)|.$$

- The expressions $\partial_x f$ and $\partial_t f$ denote the partial derivative of a function $f = f(x,t)$ w.r.t. the first variable x and the second variable t, respectively.
- For functions $f = f(x + g(x))$ we will have to distinguish between $\partial_x f$ and $\partial_1 f$.

Part I

Concepts of pulse interaction — Separation of internal and interaction dynamics

1 Description of basic tools — One pulse solutions

This introductory chapter explains a method which can be employed for constructing approximate pulse solutions for partial differential equations. The presented ideas will extensively be used throughout this work. The approximation is realized by a reduction of the original partial differential equation to a "simpler" one — the so-called "modulation system".

The presented method essentially follows the work of Kirrmann, Mielke and Schneider from 1992 [KSM92] (well, with slight modifications).

To keep technicalities on a reasonable level, we make use of the cubic Klein–Gordon equation

$$\partial_t^2 u = \partial_x^2 u - u + u^3, \quad x, t, u = u(x,t) \in \mathbb{R}, \tag{1.0.1}$$

for illustration and also restrict our analysis to one-pulse solutions (see Fig. 1.1). In this sense, this chapter can be viewed as a warm-up for the more complex setting we encounter when dealing with multipulse solutions for "more complicated" equations (meaning with a more complicated dispersion relation, cf. Fig. 3.1). However, the basic concepts introduced here will play a major role throughout the present work.

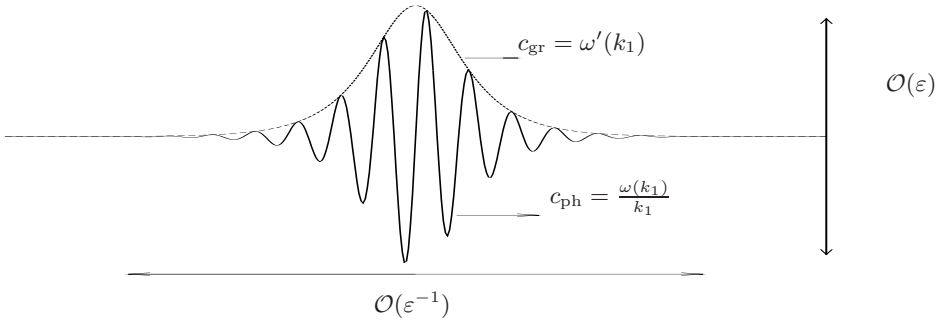

Figure 1.1: A pulse consisting of a sinusoidal carrier wave traveling with velocity c_{ph} and a spatially localized envelope traveling with velocity c_{gr}

A rigorous construction and justification of approximate solutions consists of two steps:

1. The *formal analysis*, which includes the construction of the approximation — hereafter called "ansatz" — and the derivation of solvability conditions for its components — the so-called "modulation system".

2. The *rigorous justification of the approximation*, i.e. giving an error bound (in some suitable norm) and an estimate for the time scale on which the approximation is valid.

This chapter devotes a section to each step.

Note, that there is no general theory available which would guide the way for the first step. It is usually based on a lot of experience and intuition, which often comes from the underlying physical problem.

1.1 Formal derivation of modulations equations

1.1.1 The ansatz — Approximating pulse solutions

By "pulse" we mean a solution as depicted in Fig. 1.1, i.e. consisting of a sinusoidal carrier wave and a spatially localized envelope. Since in practice the nonlinearity of the equation is comparably weak, we make use of the solutions $e^{ikx}e^{\pm i\omega(k)t}$ of the linearized problem ,whose spatial and temporal wave numbers are linked through the so-called "dispersion relation"

$$\omega(k)^2 = k^2 + 1, \tag{1.1.1}$$

to construct an *ansatz* approximating such a pulse, i.e. we take a complex function $\tilde{A} = \tilde{A}(k)$ which is localized around a wave number $k_1 \in \mathbb{R}$ with width ε where $0 < \varepsilon \ll 1$ (see Fig. 1.2), and set

$$\begin{aligned} u^{an}(x,t) &= \mathrm{Re}\{\int_{\mathbb{R}} \tilde{A}(k) e^{ikx+i\omega(k)t}\, dk\} \\ &= \varepsilon\, e^{i(k_1 x+\omega(k_1)t)} \int_{\mathbb{R}} \tilde{A}(k_1+\varepsilon\underline{K}) e^{i(\varepsilon(x+\omega'(k_1)t)\underline{K}+\frac{1}{2}\varepsilon^2 t\omega''(k_1)\underline{K}^2)}\, d\underline{K} \\ &\qquad\qquad +\mathcal{O}(\varepsilon^4 t) + c.c. \\ &= \varepsilon A(\varepsilon(x+\omega'(k_1)t), \varepsilon^2 t) e^{ik_1 x+i\omega(k_1)t} + \mathcal{O}(\varepsilon^4 t) + c.c. \end{aligned} \tag{1.1.2}$$

with $\omega(k) := \sqrt{k^2+1}$ and a spatially localized function

$$A(\varepsilon(x+\omega'(k_1)t), \varepsilon^2 t) := \int_{\mathbb{R}} \tilde{A}(k_1+\varepsilon\underline{K}) e^{i(\varepsilon(x+\omega'(k_1)t)\underline{K}+\frac{1}{2}\varepsilon^2 t\omega''(k_1)K^2)}\, d\underline{K}.$$

This ansatz approximates what is often referred to as *small–amplitude solutions*. It describes how the zero solution of the linear problem evolves under the influence of the nonlinearity.

Now we need to derive conditions for A to make the approximation as good as possible. To this end we introduce the so-called residual.

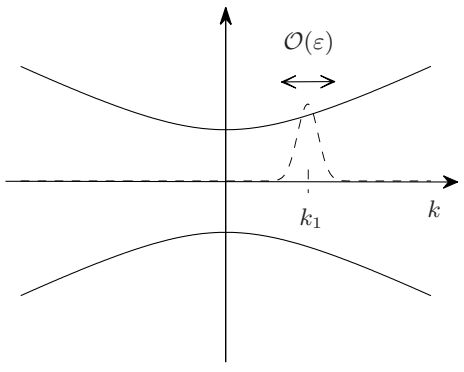

Figure 1.2: *Left:* The dispersion relation $\omega^2 = k^2 + 1$ of the cubic Klein-Gordon equation vs. the scaling in Fourier space of $\tilde{A}$ (dashed); *Right:* Scaling in x-space of the pulse u

1.1.2 The residual — Measuring the quality of the ansatz

Our desire is to construct the ansatz u^{an} for the cubic Klein-Gordon equation (1.0.1) such that there is an exact solution u "near by", i.e. fulfilling

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u-u^{an}\| \leq C\varepsilon^{\delta}, \quad \delta>1, C>0.$$

on the natural time scale $\mathcal{O}(1/\varepsilon^2)$. To put this in words, we want the error $R := u - u^{an}$ to be as small as possible and, of course, at least smaller than the approximated solution itself (hence $\delta > 1$). Specifying what we mean by "near by", i.e. specifying norms (resp. functions spaces) at this point would only distract from the overview we want to convey, so we postpone this discussion to a later paragraph.
Plugging the representation of a solution in terms of its approximation

$$u = u^{an} + R$$

into the cubic Klein-Gordon equation results in the so–called "error equation"

$$\partial_t^2 R = \partial_x^2 R - R + G(R, u^{an}) + \mathrm{Res}\{u^{an}\}$$

with

$$G(R,u^{an}) := (u^{an}+R)^3 - (u^{an})^3 = 3(u^{an})^2 R + 3u^{an}R^2 + R^3,$$

and with the so-called "residual"

$$\mathrm{Res}\{u^{an}\} := -\partial_t^2 u^{an} + \partial_x^2 u^{an} - u^{an} + (u^{an})^3. \tag{1.1.3}$$

This shows that the magnitude of the approximation error R is determined by the residual $\mathrm{Res}\{u^{an}\}$. Hence, we will use $\mathrm{Res}\{u^{an}\}$ to come up with a good construction of u^{an}. Clearly, if u is an exact solution, then $\mathrm{Res}\{u\} = 0$. Otherwise, we have to find a way of

1.1.3 Getting the residual small — Derivation of modulation equations

The ansatz (1.1.2) we constructed above leads to

$$\begin{aligned}\mathrm{Res}\{u^{an}\} = \quad & \varepsilon\,[\omega(k_1)^2 - k_1^2 - 1]AE + \varepsilon^2\,[-\omega(k_1)\omega'(k_1) + k_1](\partial_1 A)E \\ & + \varepsilon^3[-2i\omega(k_1)\partial_2 A + (1-\omega'(k_1)^2)\partial_1^2 A + 3|A|^2 A]E + \varepsilon^3[3A^3E^3] \\ & + \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E] - \varepsilon^5[(\partial_2^2 A)E] + \mathcal{O}(\varepsilon^6) + c.c..\end{aligned}$$

where we introduced the notation $E := e^{ik_1x + i\omega(k_1)t}$. Now, to diminish the residual, we will successively set the terms in front of each ε order to zero.
At $\mathcal{O}(\varepsilon)$ we get the dispersion relation (cf. (1.1.1))

$$\omega(k_1)^2 = k_1^2 + 1,$$

and at $\mathcal{O}(\varepsilon^2)$ we get the formula for the group velocity

$$\omega'(k_1) = \frac{k_1}{\omega(k_1)}$$

which are fulfilled by definition due to the construction in (1.1.2). This results into the cancelation

$$\begin{aligned}\mathrm{Res}\{u^{an}\} = \quad & \varepsilon\,\cancel{[\omega(k_1)^2 - k_1^2 - 1]}AE + \varepsilon^2\,\cancel{[-\omega(k_1)\omega'(k_1) + k_1]}(\partial_1 A)E \\ & + \varepsilon^3[-2i\omega(k_1)\partial_2 A + (1-\omega'(k_1)^2)\partial_1^2 A + 3|A|^2 A]E + \varepsilon^3[3A^3E^3] \\ & + \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E] - \varepsilon^5[(\partial_2^2 A)E] + \mathcal{O}(\varepsilon^6) + c.c..\end{aligned}$$

so we are left with

$$\begin{aligned}\mathrm{Res}\{u^{an}\} = \quad & \varepsilon^3[-2i\omega(k_1)\partial_2 A + (1-\omega'(k_1)^2)\partial_1^2 A + 3|A|^2 A]E + \varepsilon^3[3A^3E^3] \\ & + \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E] - \varepsilon^5[(\partial_2^2 A)E] + \mathcal{O}(\varepsilon^6) + c.c..\end{aligned}$$

Choosing A to be a spatially localized solution of the Nonlinear Schrödinger (NLS) equation

$$-2i\omega(k_1)\partial_2 A + (1-\omega'(k_1)^2)\partial_1^2 A + 3|A|^2 A = 0 \tag{1.1.4}$$

we get

$$\begin{aligned}\mathrm{Res}\{u^{an}\} = \quad & \varepsilon^3\cancel{[-2i\omega(k_1)\partial_2 A + (1-\omega'(k_1)^2)\partial_1^2 A + 3|A|^2 A]E} \\ & + \varepsilon^3[A^3E^3] \\ & + \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E] - \varepsilon^5[(\partial_2^2 A)E] + \mathcal{O}(\varepsilon^6) + c.c..\end{aligned}$$

Hence, in order to diminish the residual, the ansatz has to be constructed such that the carrier wave is modulated by a spatially localized solution A of the above NLS — which, in this context, is referred to as "modulation" or "amplitude equation" for obvious reasons.

1.1.4 Canceling higher order harmonics — Algebraic solvability conditions

To further cancel the higher order harmonic term, i.e. the one proportional to $E^3 = e^{3ik_1x+3i\omega(k_1)t}$ at $\mathcal{O}(\varepsilon^3)$ we modify the ansatz by

$$\tilde{u}^{an} := u^{an} + \varepsilon^3\alpha A_3 E^3 + c.c.,$$

to arrive at

$$\begin{aligned}\mathrm{Res}\{\tilde{u}^{an}\} = & \varepsilon^3[(\alpha(9\omega(k_1)^2 - 9k_1^2 - 1) + 1)A_3]E^3 \\ & + \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E + 3i\alpha(-\omega(k_1)\omega'(k_1) + k_1)(\partial_1 A_3)E^3] \\ & + \varepsilon^5[-(\partial_2^2 A + 3\alpha A_3(\overline{A})^2)E + (\alpha(1-\omega'(k_1)^2)(\partial_1^2 A_3) + 6\alpha A_3|A|^2)E^3] + \mathcal{O}(\varepsilon^6) + c.c..\end{aligned}$$

Recalling that $\omega'(k_1) = k_1/\omega(k_1)$ and choosing

$$\alpha = (9\omega(k_1)^2 - 9k_1^2 - 1)^{-1} \tag{1.1.5}$$

we finally get

$$\begin{aligned}\mathrm{Res}\{\tilde{u}^{an}\} = & \varepsilon^3\cancel{[(\alpha(9\omega(k_1)^2 - 9k_1^2 - 1) + 1)A_3]}E^3 \\ & + \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E + 3i\alpha\cancel{(-\omega(k_1)\omega'(k_1) + k_1)}(\partial_1 A_3)E^3] \\ & + \varepsilon^5[-(\partial_2^2 A + 3\alpha A_3(\overline{A})^2)E + (\alpha(1-\omega'(k_1)^2)(\partial_1^2 A_3) + 6\alpha A_3|A|^2)E^3] + \mathcal{O}(\varepsilon^6) + c.c..\end{aligned}$$

In summary, we are left with

$$\mathrm{Res}\{\tilde{u}^{an}\} = \varepsilon^4[-2\omega'(k_1)(\partial_{1,2}A)E] + \mathcal{O}(\varepsilon^5) + c.c.,$$

which makes the residual — at least formally — of order $\mathcal{O}(\varepsilon^4)$.
Note, that the definition of α in (1.1.5) is only well-defined if the so-called "non-resonance condition"

$$|9\omega(k_1)^2 - 9k_1^2 - 1| > 0, \tag{1.1.6}$$

is fulfilled, which is the case for the dispersion relation ω of the cubic Klein-Gordon equation. We will encounter such conditions every time we need to cancel higher order (and also mixed for the two-pulse case) harmonics. Indeed, in the following chapters we will make extensive use of this mechanism to cancel terms in the residual which are not proportional to the pure harmonic E.

The general procedure goes as follows: Sort the residual first for different powers of ε and then for pure and higher order (or mixed in the two pulse case, cf. Sec. 2.2.4) harmonics. Derive modulation equations by using the terms proportional to the pure harmonics. Then cancel the mixed harmonics by adding them with real factors — in the above case α — to the ansatz and deriving algebraic equations — in the above case (1.1.5) — for these factors. Throughout this work, we will call the higher order (and mixed harmonics in the two-pulse case) in the residual M_{mixed} and the terms that we add in order to cancel them $M_{mixed,\alpha}$.
Note, that this procedure will always work for cubic nonlinearities. In case of quadratic nonlinearities, this is different (see [Sch05]).

1.1.5 The choice of norms

The natural spaces for existence and uniqueness for our modulation equation — the Nonlinear Schrödinger equation — are Sobolev spaces $H^{s_A}, s_A \geq 1$, (cf. [Sul99]). However, since the modulation equation is posed in the slow variable $\underline{X} = \varepsilon x$, when bounding the residual w.r.t. the original time scale, i.e. the fast variable x, we will loose $1/2$ in the exponent since

$$\|A(\varepsilon\cdot)\|_{L^2} = \left(\int_{\mathbb{R}} |A(\varepsilon x)|^2\, dx\right)^{1/2} = \varepsilon^{-1/2}\left(\int_{\mathbb{R}} |A(\underline{X})|^2\, d\underline{X}\right)^{1/2} = \varepsilon^{-1/2}\|A(\cdot)\|_{L^2}, \tag{1.1.7}$$

and, similarly,

$$\|A(\varepsilon\cdot)\|_{H^s} \leq C\varepsilon^{-1/2}\|A(\cdot)\|_{H^s}. \tag{1.1.8}$$

Furthermore, it will turn out in the justification proof that we also need to bound the ansatz w.r.t. a norm which does not "feel" the scaling. Hence, we will also work in C_b^s and make use of the Sobolev embedding (cf. [Ada07]) to transfer the estimates, i.e.

$$\|A(\cdot)\|_{C_b^{s-1}} \leq C\|A(\cdot)\|_{H^s}. \tag{1.1.9}$$

1.1.6 Modulations system for one pulse solutions — Internal dynamics

Now the question arises if our ansatz can be improved such that the residual becomes even smaller. The answer is yes!

Remark. In the subsequent Lemma 1.1.1 the superscript (r) is actually an abuse of notation, since (r) does not refer to derivatives, it only refers to the order of correction.

Lemma 1.1.1. *Fix $s, s_A \in \mathbb{R}$ for arbitrary $r_{acc} \in \mathbb{N}$ according to $1 \leq s \leq s_A - 3(r_{acc}-1) - 4$. If we set*

$$u^{an}(x,t) = \sum_{r=1}^{r_{acc}} \varepsilon^r A^{(r)}(\varepsilon(x+\omega'(k_1)t), \varepsilon^2 t) e^{ik_1 x + i\omega(k_1)t} + M_{mixed,\alpha} + c.c$$

where $A^{(r)} \in C([0,T_0], H^{s_A-3(r-1)})$ solves the equation

$$2i\omega(k_1)\partial_2 A^{(r)} = (1-\omega'(k_1)^2)\partial_1^2 A^{(r)} - 2\omega'(k_1)\partial_{1,2} A^{(r-1)} - \partial_2^2 A^{(r-2)} + G^{(r)}_{internal} \tag{1.1.10}$$

with the so-called internal dynamics terms given by

$$G^{(r)}_{internal} := 3 \sum_{2r_1+r_2=r+2} (A^{(r_1)})^2 \overline{A^{(r_2)}} + 6 \sum_{\substack{r_1+r_2+r_3=r+2, \\ r_1 \neq r_3}} A^{(r_1)} \overline{A^{(r_2)}} A^{(r_3)},$$

(where we used the convention $A^{(r)} = 0, r \leq 0$) and $M_{mixed,\alpha}$ is according to Sec. 1.1.4, we have that

$$\|\mathrm{Res}\{u^{an}(\cdot,t)\}\|_{H^s} \leq C\varepsilon^{res}, \quad res := r_{acc} + 3 - 1/2,$$

uniformly in time, with $C > 0$ independent of ε.

Remarks. The system of modulation equations (1.1.10) has the following structure:

- The equation for $A^{(r)}$ only involves $A^{(l)}$ with $l \leq r$. Hence, the system can be solved recursively.
- Since the terms $A^{(r)}$ are defined recursively, this results in a loss of regularity.
- At $\mathcal{O}(\varepsilon^3)$, i.e. $r = 1$, we get a Nonlinear Schrödinger equation

$$2i\omega(k_1)\partial_2 A^{(1)} = (1-\omega'(k_1)^2)\partial_1^2 A^{(1)} + 3|A^{(1)}|^2 A^{(1)}. \tag{1.1.11}$$

  Hence, $G^{(1)}_{internal} = 3|A^{(1)}|^2 A^{(1)}$. From the summation in $G^{(r)}_{internal}$ it is clear that this is the only nonlinear equation that can occur in the modulation system, since $a_3 = a_4 = r \Rightarrow a_5 = 2 - r \Rightarrow r = 1$. For $r \geq 2$ only linear Schrödinger equations driven by the solutions of the lower order equations can appear.

Proof. As already alluded to, the natural spaces for existence and uniqueness for Schrödinger type equations are the Sobolev spaces $H^s, s \geq 1$ (cf. [Sul99]).

For $r = 1$ we get (1.1.11), i.e. a Nonlinear Schrödinger equation, for which spatially localized solutions with $A^{(1)} \in C([0,T_0], H^{s_A})$, $s_A \geq 1$ exist (see Rem. 1.1.2).
For $r = 2$ (1.1.10) has the form

$$2i\omega(k_1)\partial_2 A^{(2)} = (1-\omega'(k_1)^2)\partial_1^2 A^{(2)} - 2\omega'(k_1)\partial_{1,2} A^{(1)} \underbrace{+3(A^{(1)})^2\overline{A}^{(2)} + 6A^{(2)}A^{(1)}\overline{A}^{(1)}}_{=G^{(2)}_{internal}},$$

which is a linear Schrödinger equation with inhomogeneous terms involving derivatives of $A^{(1)}$. Dissolving temporal derivatives of $A^{(1)}$ by using that it solves the NLS (1.1.11) and combining this with the fact that the regularity of solutions depends on the regularity of the right-hand side gives existence and uniqueness with $A^{(2)} \in C([0,T_0], H^{s_A-3})$.
Iterating this procedure gives that for all $r = 1, \ldots, r_{acc}$ the system (1.1.10) has unique solutions

$A^{(r)} \in C([0,T_0], H^{s_A-3(r-1)}), s_A \geq 3(r_{acc}-1)$. Moreover, since the equations do not explicitly depend on ε, all solutions are $\mathcal{O}(1)$.

If $A^{(r)}$ solve the modulation system (1.1.10) for $r = 1,\dots r_{acc}$, we canceled all terms up to order $\mathcal{O}(\varepsilon^{r_{acc}+3})$. To make the residual bound rigorous, we have to control all the terms left at order $\mathcal{O}(\varepsilon^{r_{acc}+3})$ and higher. This means we have to get, for example,

$$\|\partial_{1,2}A^{(r_{acc}-1)}\|_{H^s} \leq C, \quad \|\partial_2^2 A^{(r_{acc}-1)}\|_{H^s} \leq C, \dots,$$

for some $C > 0$. Now tracking terms involving derivatives in the residual at higher orders finally gives the condition

$$s \leq s_A - 3(r_{acc}-1) - 4,$$

since the highest regularity loss comes from

$$\partial_2^2 A^{(r_{acc})},$$

which can be seen by recalling that $A^{(r_{acc})} \in C([0,T_0], H^{s_A-3(r_{acc}-1)})$ and using the equation for $A^{(r_{acc})}$ to dissolve the temporal derivatives (which gives a loss of 4).
Hence, we get

$$\|\mathrm{Res}\{u^{an}\}\|_{H^s} \leq C\varepsilon^{r_{acc}+3}\varepsilon^{-1/2},$$

for some $C > 0$, if $s_A - 3(r_{acc}-1) - 4$.
Note, that the loss of $1/2$ in the exponent is due to the scaling property (1.1.8) of L^2-based norms. □

Remark 1.1.2. *It is well-known that for the Nonlinear Schrödinger equation*

$$2i\omega(k_1)\partial_2 A^{(1)} = (1-\omega'(k_1)^2)\partial_1^2 A^{(1)} + 3|A^{(1)}|^2 A^{(1)}$$

the condition $3(1-\omega'(k_1)^2) > 0$ *ensures that spatially localized solutions exist. If this condition is fulfilled the equation is called "focussing". In our case, we have* $\omega'(k) = \frac{k}{\sqrt{k^2+1}} < 1$ *for all* $k \in \mathbb{R}$, *so the condition is fulfilled for arbitrary* $k_1 \in \mathbb{R}$.
The most famous localized solution is the so-called "soliton" which can be specified in terms of elementary functions. It is given by

$$A^{(1)}(\underline{X}_1, \underline{T}) = C_1 \mathrm{sech}(C_2\underline{X})e^{\gamma_0 \underline{T}}e^{i\phi_0}, \quad \underline{X}_1 := \varepsilon(x+\omega'(k_1)t), \quad \underline{T} := \varepsilon^2 t,$$

with the constants $C_1 := \sqrt{\frac{4\gamma_0\omega(k_1)}{3}}, C_2 := \frac{2\gamma_0\omega(k_1)}{1-\omega'(k_1)}$, *for arbitrary* $\gamma_0 > 0$ *(see [Pol04] for a collection of solutions).*

1.2 Justification of modulation equations for long time scales

In order to construct a good approximation it is crucial to have a clear understanding of the interplay between the order of the residual, the approximation error and the time scale of validity. We sum up exactly this trade–off in the following approximation theorem.

Theorem 1.2.1. *Given the assumptions of Lemma 1.1.1, for all* $T_0 > 0$ *there is* $\varepsilon_0 > 0$ *and a constant* $C > 0$ *independent of* ε *such that*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(\cdot,t) - u^{an}(\cdot,t)\|_{C_b^0(\mathbb{R})} \leq C\varepsilon^\delta, \quad \delta := res - 2, \tag{1.2.1}$$

for all $\varepsilon \in (0,\varepsilon_0)$, *where* u *is an exact solution of the cubic Klein-Gordon equation* (1.0.1).

Proof. Define $u - u_{ansatz} =: \varepsilon^\delta R$, and plug $u = u_{ansatz} + \varepsilon^\delta R$ into the cubic Klein-Gordon equation (1.0.1) to derive the error equation

$$\partial_t^2 R = \partial_x^2 R - R + f$$

with

$$f = \varepsilon^{-\delta}(\mathrm{Res}\{u^{an}\} + \varepsilon^{3\delta} R^3 + 3\,\varepsilon^{2\delta} R^2 u^{an} + 3\,\varepsilon^{\delta} R u^{an2}).$$

Now we turn this equation into a first order system for $(R, \partial_t R)$ and work in the space $H^1 \times L^2$ equipped with the norm

$$\|(R, \partial_t R)\|^2_{H^1 \times L^2} = \|R\|^2_{H^1} + \|\partial_t R\|^2_{L^2} = \int R^2 + (\partial_x R)^2 + (\partial_t R)^2 \; dx.$$

Note, that this special choice enables us to use the structure of the linear part. To be precise, with the notation

$$E(R) := \|(R, \partial_t R)\|^2_{H^1 \times L^2},$$

which indicates that this quantity represents an "energy" - hence the term *energy estimates* - we have

$$\begin{aligned}
\frac{d}{dt} E(R) &= 2 \int R(\partial_t R) + (\partial_x R)(\partial_t \partial_x R) + (\partial_x R)(\partial_t^2 R) \; dx \\
&= 2 \int R(\partial_t R) - (\partial_x^2 R)(\partial_t R) + (\partial_x R)(\partial_t^2 R \|\partial_t R\|_{L^2} f) \; dx \\
&= 2 \int (\partial_t R) f \; dx \leq 2 \|\partial_t R\|_{L^2} \|f\|_{L^2} \leq 2 E(R)^{1/2} \|f\|_{L^2}.
\end{aligned}$$

Now the strategy is to bound f in terms of E such that we get

$$\frac{d}{dt} E(R) \leq \alpha + \beta E(R) + \gamma E(R)^2, \quad t \in [0, T_0/\varepsilon^2],$$

and thereby conclude with the differential inequality from the subsequent Thm. 1.2.2 and conditions on α, β, γ that

$$E(R) \leq C, \quad t \in [0, T_0/\varepsilon^2],$$

with $C > 0$ independent of ε. This would conclude the proof since $\|R\|_{H^1} \leq E(R)^{1/2}$. Now the rest of the proof is devoted to the computation of α, β and γ:

We can estimate f using

$$\|f\|_{H^1} \leq C_{res} \varepsilon^{res-\delta} + \varepsilon^{2\delta} \|R\|^3_{H^1} + 3\; \varepsilon^{\delta+1} C_{an} \|R\|^2_{H^1} + 3\; \varepsilon^2 C^2_{an} \|R\|_{H^1},$$

where $C_{an} := \|u^{an}\|_{C^0_b}$ and $C_{res} > 0$ is defined by $\|\mathrm{Res}\{u^{an}\}\|_{H^1} \leq C_{res} \varepsilon^{res}$. Thus we get (uniformly in time)

$$\begin{aligned}
&\frac{d}{dt} E(R) \\
&\leq 2 E(R)^{1/2} [C_{res} \varepsilon^{res-\delta} + \varepsilon^{2\delta} \|R\|^3_{H^1} + 3\; \varepsilon^{\delta+1} C_{an} \|R\|^2_{H^1} + 3\; \varepsilon^2 C^2_{an} \|R\|_{H^1}] \\
&\leq 2 E(R)^{1/2} [C_{res} \varepsilon^{res-\delta} + \varepsilon^{2\delta} E(R)^{3/2} + 3\; \varepsilon^{\delta+1} C_{an} E(R) + 3\; \varepsilon^2 C^2_{an} E(R)^{1/2}] \\
&= 2 [C_{res} \varepsilon^{res-\delta} E(R)^{1/2} + \varepsilon^{2\delta} E(R)^2 + 3\; \varepsilon^{\delta+1} C_{an} E(R)^{3/2} + 3\; \varepsilon^2 C^2_{an} E(R)] \\
&\leq \underbrace{C \varepsilon^{res-\delta}}_{=:\alpha} + \underbrace{(C \varepsilon^{res-\delta} + 3C\varepsilon^2)}_{=:\beta} E(R) + \underbrace{\varepsilon^{\delta+1}(1 + 3C)}_{=:\gamma} E(R)^2.
\end{aligned}$$

Now the differential inequality from the subsequent Thm. 1.2.2 gives

$$E(R) \leq \beta \frac{\alpha T_0/\varepsilon^2}{\beta + \gamma \alpha T_0/\varepsilon} e^{\beta T_0/\varepsilon^2} \frac{1}{1 - \gamma \frac{\alpha T_0/\varepsilon^2}{\beta + \gamma \alpha T_0/\varepsilon} e^{\beta T_0/\varepsilon^2}}.$$

Since, $\beta = \mathcal{O}(\varepsilon^{\min\{res-\delta,2\}})$ and $e^{\beta T_0/\varepsilon^2} \stackrel{!}{=} \mathcal{O}(1)$, we get as condition $\beta = \mathcal{O}(\varepsilon^2)$ and, hence, $res - \delta \geq 2$. This simplifies the inequality for $E(R)$ to

$$E(R) \leq C \frac{\alpha}{\varepsilon^2},$$

such that we have to demand $\alpha \stackrel{!}{=} \Theta(\varepsilon^2)$ and therefore $res - \delta = 2$ to finally get

$$E(R) \leq C.$$

Note, that $\gamma = \mathcal{O}(\varepsilon^{\delta+1})$ can be chosen arbitrarily and so, indeed, decreasing the magnitude of the residual will directly translate into an improved error bound. However, there is no way of improving the length of the time scale of validity.
Finally, using the Sobolev embedding (1.1.9) concludes the proof.

□

Some remarks about the limitations of the proof

1. *Tracking where the exponent* 2 *in the time scale comes from makes it evident that the time scale is limited by the order of* ε *that is generated by the nonlinearity. This is a special feature of our equation. The prove does not work for e.g. quadratic nonlinearities. See [Kal87],[Sch05] for an improved version.*

2. *Diminishing the residual improves the accuracy of the ansatz, but cannot increase the time scale of validity!*

A differential inequality. In the proof we used the following differential inequality (see [Pac98], Thm. 2.5.1, p.126):

Theorem 1.2.2. *Let* a, b *be continuous and positive functions defined on* $\mathbb{R}_+$ *and* $w : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ *be a continuous, nondecreasing function such that* $w(0) = 0$ *and* $w(x) > 0$ *for* $x > 0$. *If* u *is a positive differentiable function on* $\mathbb{R}_+$ *that satisfies*

$$u'(t) \leq a(t)w(u(t)) + b(t), \quad t \in \mathbb{R}_+$$

then we have

$$u(t) \leq G^{-1}\left[G\left(u(0) + \int_0^t b(s)ds\right) + \int_0^t a(s)ds\right],$$

with

$$G(r) = \int_{r_0}^{r} \frac{1}{w(z)} dz,$$

for all $t \in \mathbb{R}_+$ *where the right-hand side is defined.*

We have used this inequality in the proof of Thm. 1.2.1 for

$$\frac{d}{dt}E(R) \leq \alpha + \beta E(R) + \gamma E(R)^2.$$

Hence, in our case $w(u) = \beta u + \gamma u^2$ and $a(t) = 1, \quad b(t) = \alpha, \quad t \in [0, T_0/\varepsilon^2]$.

Proof for H^s Since for our later results we will need more regularity than H^1 we also report how to carry out the proof for H^s for arbitrary $s \geq 1$. To this end we only have to introduce a slightly modified energy

$$E^s(R) := \sum_{j=0}^{s} \int_{\mathbb{R}} (\partial_x^j \partial_t R)^2 + (\partial_x^{j+1} R)^2 + (\partial_x^j R)^2 \, dx,$$

and so

$$\begin{aligned}
\frac{1}{2}\partial_t E^s(R) &= \sum_{j=0}^{s} \int_{\mathbb{R}} (\partial_x^j \partial_t R)(\partial_x^j [\partial_t^2 R]) + \{(\partial_x^{j+1} R)(\partial_t \partial_x^{j+1} R)\} + (\partial_x^j R)(\partial_t \partial_x^j R) \, dx \\
&= \sum_{j=0}^{s} \int_{\mathbb{R}} (\partial_x^j \partial_t R)(\partial_x^j [\partial_x^2 R - R + f]) - (\partial_x^{j+2} R)(\partial_t \partial_x^j R) + (\partial_x^j R)(\partial_t \partial_x^j R) \, dx \\
&= \sum_{j=0}^{s} \int_{\mathbb{R}} (\partial_x^j \partial_t R)(\partial_x^j f) \, dx \leq E^s(R)^{1/2} \|f\|_{L^2} + E^s(R)^{1/2} \|\partial_x f\|_{L^2} + \dots \\
&= E^s(R)^{1/2} \|f\|_{H^s} \leq (1 + E^s(R)) \|f\|_{H^s}.
\end{aligned}$$

1.3 Chapter Summary

This chapter gives an introduction to the main analytical tools that we make use of throughout the work. We explain how to derive and rigorously justify the Nonlinear Schrödinger equation for envelopes of one-pulse solutions for the cubic Klein-Gordon equation. Let us briefly recapitulate our findings:

The cubic Klein–Gordon equation

$$\partial_t^2 u = \partial_x^2 u - u + u^3, \quad x, t, u = u(x,t) \in \mathbb{R},$$

possesses one-pulse solutions as depicted in Fig. 1.1 which can be approximated to any polynomial order $r_{acc} \geq 1$ in the perturbation parameter $0 < \varepsilon \ll 1$ by the ansatz

$$u^{an}(x,t) = \sum_{r=1}^{r_{acc}} \varepsilon^r A^{(r)}(\varepsilon(x + \omega'(k_1)t), \varepsilon^2 t) e^{ik_1 x + i\omega(k_1)t} + M_{mixed,\alpha} + c.c.,$$

with $A^{(r)}$ fulfilling the recursively solvable modulation system (1.1.10), i.e.

$$\begin{aligned}
-2i\omega(k_1)\partial_2 A^{(1)} &= (1 - \omega'(k_1)^2)\partial_1^2 A^{(1)} + 3|A^{(1)}|^2 A^{(1)}, \\
-2i\omega(k_1)\partial_2 A^{(2)} &= (1 - \omega'(k_1)^2)\partial_1^2 A^{(2)} - 2\omega'(k_1)\partial_{1,2} A^{(1)} + 3(A^{(1)})^2 \overline{A}^{(2)} + 6A^{(2)} A^{(1)} \overline{A}^{(1)}, \\
&\vdots
\end{aligned}$$

which is headed by a Nonlinear Schrödinger equation for $A^{(1)}$ and followed by a system of linear Schrödinger equations for $A^{(r)}, r \geq 2$. The approximation is valid on time scale of order $\mathcal{O}(1/\varepsilon^2)$ which can be proved by a method described in [KSM92] involving energy estimates (or equivalently the right choice of function space) and a Gronwall type inequality.

The great achievement of this result can be stated as follows: One can reduce the problem of finding *one-pulse solutions for the cubic Klein–Gordon equation,* which is impossible to attack directly, in leading order to the problem of finding *localized solutions of the Nonlinear Schrödinger equation*, a problem

that has extensively been studied (cf. [Sul99], [Bou99]) and is easier to attack both analytically and numerically. In particular, the Nonlinear Schrödinger equation (under some condition on the coefficients) is known to possess localized solutions — so-called "solitons" — that can be explicitly specified in terms of elementary functions (see Rem. 1.1.2), which is a very rare situation when dealing with partial differential equations (see [Pol04] for a collection of explicitly known solutions of the NLS). From a numerical point of view, the reduced problem is much easier to handle since it only describes the envelope, i.e. the highly oscillatory part of the pulse, which demands a very high spatial resolution for the numerical scheme, is removed.

2 Separation of internal and interaction dynamics of multipulse solutions

This chapter explains a method which we developed to analyze the interaction of pulses in the setting of nonlinear wave equations. As in the previous chapter, we will make use of the cubic Klein-Gordon equation (1.0.1) for illustrations. This will prevent the main notions from drowning in the crowd of technicalities. Nevertheless, we want to point out that the method is by no means limited to the cubic Klein-Gordon case, but in fact can be applied in very general settings (see Ch. 3).

We distinguish between two types of pulse interactions (where the names are inspired by applications in nonlinear fiber optics, see Sec. 2.3 and Ch. 7):

1. *Intrachannel interaction*, where the interacting pulses travel on the same carrier and are distributed with equidistant spacing (see Fig. 2.1).

2. *Interchannel interaction*, where the interacting pulses travel on different carriers and hence with different velocities (see Fig. 2.2).

Now the first section of this chapter is dedicated to the analysis of intrachannel interaction, while the second is devoted to interchannel interaction. The third chapter merges these results to a general $N \times M$ *multipulse*, i.e. N pulse trains each consisting of M equidistant pulses (see Fig. 2.6), which is a model for so-called "Wavelength-Division multiplexing" techniques in optical communication lines (see Ch. 7).

Remark 2.0.1. *For the analysis of the multipulse in Sec. 2.3 it will turn out that we need an accuracy of $\mathcal{O}(\varepsilon^\delta), \delta > 3$, for our ansatz to capture the leading order interaction effects (see Rem. 2.2.3). Expressed in the language of Lemma 1.1.1 and Thm. 1.2.1, we will choose in the following $r_{acc} = 3$ and so the bound on the residual will be $res = r_{acc}+3-1/2 = 11/2$ such that the error $\delta = res-2 = 7/2 > 3$.*

2.1 Interaction of pulses with the same carrier wave - Intrachannel interaction

Two pulses traveling on the same carrier wave (see Fig. 2.1) can be approximated by the ansatz

$$u^{an}(x,t) = \sum_{n=1,2} \varepsilon A_n^{(1)}(\underline{X}_1 + \varepsilon d_n, \underline{T}) E_1 + c.c.,$$
$$E_1 = e^{ik_1 x + i\omega(k_1)t}, \quad \underline{X}_1 = \underline{X} + \varepsilon\omega'(k_1)t, \quad \underline{X} := \varepsilon x, \quad \underline{T} = \varepsilon^2 t,$$

with the spacing

$$|d_1 - d_2| \gg 0$$

which is nothing but a superposition of two single pulses as given in (1.1.2). Note, that due to the localization of the pulses and the assumption $|d_1 - d_2| \gg 0$, the superposition is feasible (after all, we are in the weakly nonlinear regime). Hence, this ansatz is really a straightforward generalization of

the one-pulse ansatz from Ch. 1 to the two-pulse case.

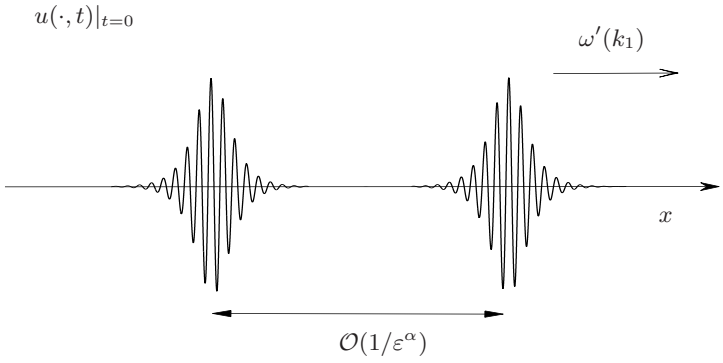

Figure 2.1: Initial condition representing intrachannel interaction

Plugging this ansatz into the residual (1.1.3) gives

$$\begin{aligned}\mathrm{Res}\{u^{an}\} = \; & \varepsilon^3 E_1 \sum_{n=1,2} \left(-2i\omega(k_1)\partial_2 A_n^{(1)} + (\omega'(k_1)^2 - 1)\partial_1^2 A_n^{(1)} \underbrace{+3|A_n^{(1)}|^2 A_n^{(1)}}_{=:G_{n,internal}^{(1)}} \right) \\ & + \; \varepsilon^3 E_1 \left(\underbrace{6|A_2^{(1)}|^2 A_1^{(1)} + 6|A_1^{(1)}|^2 A_2^{(1)}}_{=:G_{intra}^{(1)}} \right) + \mathcal{O}(\varepsilon^4) + M_{mixed},\end{aligned}$$

where $G_{n,internal}^{(r)}$ describe the internal dynamics terms already introduced in Ch. 1 for the one-pulse ansatz and where $G_{intra}^{(r)}$ contains all couplings caused by the interaction of the two pulses. The prefix "intra" should emphasize that both pulses travel on the same carrier and, hence, belong to the same "channel".

The interaction of two such pulses is expected to be very weak since they travel with the same speed and therefore really only interact through their "tails". Consequently, if they are sufficiently far apart and their spatial decay is sufficiently fast, the interaction should actually be negligible. We can put this intuition into formulas by fixing the localization strength $l \in \mathbb{R}$ of the pulses and make use of weighted spaces Sobolev spaces

$$H^s(l) := \{u \in H^s \mid \|(1+(\cdot)^2)^{l/2} u(\cdot)\|_{H^s} < \infty\}, \tag{2.1.1}$$

to get

$$\begin{aligned}
&\||A_2^{(1)}(\underline{X}_1-\varepsilon d_2,\tau)|^2A_1^{(1)}(\underline{X}_1-\varepsilon d_1,\tau)\|_{H^s}\\
&\leq \sup_t\|A_2^{(1)}\|^2_{H^s(l)}\sup_t\|A_1^{(1)}\|_{H^s(l)}\sup_{\underline{X}_1\in\mathbb{R}}\left|\frac{1}{(1+(\underline{X}_1-\varepsilon d_1)^2)^l(1+(\underline{X}_1-\varepsilon d_2)^2)^{r/2}}\right|\\
&\leq \sup_t\|A_2^{(1)}\|^2_{H^s(l)}\sup_t\|A_1^{(1)}\|_{H^s(l)}\sup_{\underline{X}_1\in\mathbb{R}}\left|\frac{1}{(1+(\underline{X}_1-\varepsilon d_1)^2)(1+(\underline{X}_1-\varepsilon d_2)^2)}\right|^{l/2}\\
&= \sup_t\|A_2^{(1)}\|^2_{H^s(l)}\sup_t\|A_1^{(1)}\|_{H^s(l)}\left(\frac{1}{(1+(\frac{\varepsilon d_0}{2})^2)}\right)^{l/2},
\end{aligned}$$

where we introduced the notation $d_0:=|d_1-d_2|$. If we set $d_0=\mathcal{O}(1/\varepsilon^\alpha),\alpha>1$, (the lower bound for α is due to the fact that the pulse width is $\mathcal{O}(1/\varepsilon)$) we can conclude

$$\left(\frac{1}{(1+(\frac{\varepsilon d_0}{2})^2)}\right)^{l/2}=\mathcal{O}(\varepsilon^{l(\alpha-1)}),$$

and, consequently, if

$$l(\alpha-1)-2\geq m_0\geq 1,$$

then

$$\sup_{\underline{T}\in[0,T_0]}\int_0^{\underline{T}}\||A_2(\underline{X}_1-\varepsilon d_0,\tau)|^2A_1(\underline{X}_2,\tau)\|_{H^s}\,d\tau=\mathcal{O}(\varepsilon^{m_0}). \tag{2.1.2}$$

Hence, we can shift the coupling terms into higher orders in the residual and choose the amplitudes $A_n^{(1)}$ to solve each an Nonlinear Schrödinger equation (without couplings). For example, if we want to shift these terms beyond $\mathcal{O}(\varepsilon^5)$, we simply choose $l(\alpha-1)-2\geq 3$, i.e. $m_0=3$ such that the terms appear at $\mathcal{O}(\varepsilon^{3+3})$.
To sum up, in first order, the two pulses do not interact at all!

Higher order corrections As already alluded to in Rem.2.0.1 (and will be explained in more detail in Rem.2.2.3) for the analysis of the multipulse in Sec. 2.3 it will turn out that we need an accuracy of $\mathcal{O}(\varepsilon^\delta),\delta>3$. Now to incorporate higher order corrections, we can employ the following two-pulse ansatz

$$\begin{aligned}
u^{an}(x,t)&=\sum_{r=1}^{3}\sum_{n=1,2}\varepsilon^r A_n^{(r)}(\underline{X}_1+\varepsilon d_n,T)E_1+M_{mixed,\alpha}+c.c.,\\
E_1&=e^{ik_1x+i\omega(k_1)t},\quad \underline{X}_1=\underline{X}+\varepsilon\omega'(k_1)t,\quad \underline{X}:=\varepsilon x,\quad \underline{T}=\varepsilon^2t,
\end{aligned} \tag{2.1.3}$$

with the spacing

$$|d_n-d_m|=d_0\gg 0$$

where the term $M_{mixed,\alpha}$ in our ansatz (2.1.3) is constructed analogously to the one-pulse case in Sec. 1.1.4.
With the same reasoning as above, we can immediately prove the following residual bound.

Lemma 2.1.1. *Let u^{an} be the two-pulse ansatz* (2.1.3) *where*

$$A_n^{(r)} \in C([0,T_0]; H^{s_A-3(r-1)}(l) \cap H^{s_A-3(r-1)+l}(0)), s_A \geq 6, l \geq 2,$$

solve the modulation system (which is actually (1.1.10) *for each $n = 1, 2$)*

$$-2i\omega(k_1)\partial_2 A_n^{(r)} = (1-\omega'(k_n)^2-1)\partial_1^2 A_n^{(r)} - 2\omega'(k_n)\partial_{1,2} A_n^{(r-1)} \partial_2^2 A_n^{(r-2)} + G_{n,internal}^{(r)},$$

with

$$G_{n,internal}^{(r)} := 3 \sum_{2r_1+r_2=r+2} (A_n^{(r_1)})^2 \overline{A_n^{(r_2)}} + 6 \sum_{\substack{r_1+r_2+r_3=r+2,\\ r_1 \neq r_3}} A_n^{(r_1)} \overline{A_n^{(r_2)}} A_n^{(r_3)}$$

and the convention $A_n^{(r)} = 0, r \leq 0$. Then for $1 \leq s \leq s_A - 10$ we have

$$\|\mathrm{Res}\{u^{an}\}\|_{H^s} \leq C\varepsilon^{11/2} + C\varepsilon^{11/2} \| \left(\varepsilon^{-3} G_{intra}^{(1)} + \varepsilon^{-2} G_{intra}^{(2)} + \varepsilon^{-1} G_{intra}^{(3)} \right) \|_{H^s}, \tag{2.1.4}$$

where

$$G_{intra}^{(r)} := 6 \sum_{\substack{r_1+r_2+r_3=r+2,\\ (r_1,n_1)\neq(r_3,n_3)}} A_{n_1}^{(r_1)} \overline{A_{n_2}^{(r_2)}} A_{n_3}^{(r_3)},$$

and if $l(\alpha-1)-2 \geq 3$, we even get

$$\sup_{t\in[0,t_0/\varepsilon^2]} \|\mathrm{Res}\{u^{an}\}\|_{H^s} \leq C\varepsilon^{11/2}.$$

Proof. The proof follows exactly the proof of Thm. 1.2.1. The only new thing is that we shifted all coupling term beyond the ε-order we are interested in. Now by using the estimate (2.1.2) with $m_0 = 3$ (since we want to shift the terms at most three orders higher), and, hence, $l(\alpha-1)-2 \geq 3$, we can show that the shifted terms $\varepsilon^{-3} G_{intra}^{(1)}, \ldots$, are $\mathcal{O}(1)$ on the long time scale, which concludes the proof. □

Remark 2.1.2. *Note, that local existence and uniqueness for weighted spaces can be obtained by a procedure like in [CKS95], which basically makes use of the fact that the propagator*

$$\mathcal{S}(\underline{T}) := \mathcal{F}^{-1} e^{-ik^2 \underline{T}} \mathcal{F}$$

of the linear Schrödinger equation is an isometry from $H^s(l) \cap H^{s+l}(0)$ into itself.

With this estimate for the residual we can immediately formulate the following approximation result for two pulses traveling on the same carrier wave.

Theorem 2.1.3. *Let u^{an} be according to Lemma 2.1.1. Then for all $T_0 > 0$ there is $C > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(\cdot,t) - u^{an}(\cdot,t)\|_{C_b^{s-1}(\mathbb{R})} \leq C\varepsilon^{7/2}, \tag{2.1.5}$$

where u is a two-pulse solution of the cubic Klein-Gordon equation (1.0.1).

In summary, by means of weighted Sobolev spaces the trade–off between spacing and decay rates becomes evident: In order to get a weak interaction, a small spacing, i.e. $\alpha > 1$ small, demands a high decay rate, i.e. l large, and vice versa. Consequently, if we only choose l and α right, we can neglect the coupling terms $G_{n,intra}^{(r)}$ completely!

Note, that the constants in front of the shifted coupling terms in (2.1.4) become ε-dependent, if deal with M pulses (instead of just two) where $M = \mathcal{O}(1/\varepsilon^\gamma)$.

2.2 Interaction of pulses with different carrier waves - Interchannel interaction

This section is concerned with the propagation of a two-pulse as depicted in Fig. 2.2, i.e. two pulses with different carrier waves (thus the name "interchannel") and, hence, traveling at different speeds. Before we report about the refined perturbation approach that allows a very precise description of pulse interaction, we first want to point out the failure of the classical approach. We do this not only to motivate our method, but also to get the reader acquainted to our notation.

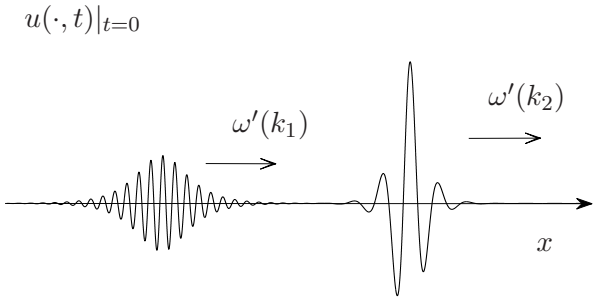

Figure 2.2: Initial condition representing interchannel interaction

2.2.1 Classical approach - The naive ansatz for a two-pulse

To analyze the interaction of two pulses with different carrier waves we can employ the "naive ansatz"

$$\begin{aligned} u^{an}(x,t) &= A_1^{(1)}(\underline{X}_1,\underline{T})E_1 + A_2^{(1)}(\underline{X}_2,\underline{T})E_2 + M_{mixed,\alpha} + c.c., \\ E_j &= e^{ik_jx+i\omega(k_j)t}, \quad \underline{X}_j = \underline{X} + \varepsilon\omega'(k_j)t, \quad \underline{X} := \varepsilon x, \quad \underline{T} = \varepsilon^2 t, \end{aligned} \tag{2.2.1}$$

where the structure of $M_{mixed,\alpha}$ will be discussed at the end of this section (see Sec. 2.2.4).
As we will see, in contrast to the intrachannel interaction case described in the last section, this straightforward generalization of the ansatz from Ch. 1 will not provide much information on interchannel interaction and, thus, will have to be refined!

Plugging the two-pulse ansatz (2.2.1) into the residual (1.1.3) gives rise to a coupled system of Nonlinear Schrödinger equations (instead of just one as in the one-pulse case)

$$\mathcal{O}(\varepsilon^3E_1): \quad -2i\omega(k_1)\partial_2A_1^{(1)} = (1-\omega'(k_1)^2)\partial_1^2A_1^{(1)} \underbrace{+3|A_1^{(1)}|^2A_1^{(1)}}_{=:G^{(1)}_{1,internal}} \underbrace{+6|A_2^{(1)}|^2A_1^{(1)}}_{=:G^{(1)}_{1,inter}}, \tag{2.2.2}$$

$$\mathcal{O}(\varepsilon^3E_2): \quad -2i\omega(k_2)\partial_2A_2^{(1)} = (1-\omega'(k_2)^2)\partial_1^2A_2^{(1)} \underbrace{+3|A_2^{(1)}|^2A_2^{(1)}}_{=:G^{(1)}_{2,internal}} \underbrace{+6|A_1^{(1)}|^2A_2^{(1)}}_{=:G^{(1)}_{2,inter}}. \tag{2.2.3}$$

Clearly, understanding interaction means finding a clever way to deal with the coupling terms $G^{(r)}_{j,inter}$.

Qualitative analysis of coupling terms. Intuitively one would expect that the interaction of two pulses with a $\mathcal{O}(1/\varepsilon)$ width, traveling with $\mathcal{O}(1)$ speeds is very weak on the long time scale of $\mathcal{O}(1/\varepsilon^2)$, since the interaction time is very short. Inspired by the estimate of the last section, we can employ the following reasoning to confirm this intuition.

Fix the spatial localization strength of two functions $f = f(\underline{X}_1, T), g = g(\underline{X}_2, T)$ by demanding that they lie in a spatially weighted Sobolev space (see (2.1.1)) for some $l \geq 2$. Then for fixed $\tau \in \mathbb{R}$ we have

$$\begin{aligned}
&\|f(\underline{X}_1,\tau)g(\underline{X}_2,\tau)\|_{H^{s_A}} \\
&\quad\leq \sup_t \|f\|_{H^{s_A}(l)} \sup_t \|g\|_{H^{s_A}(l)} \sup_{\underline{X}_1\in\mathbb{R}} \left| \frac{1}{(1+(\underline{X}_1 + D\frac{\tau}{\varepsilon})^2)^{l/2}(1+\underline{X}_1^2)^{l/2}} \right| \\
&\quad= \sup_t \|f\|_{H^{s_A}(l)} \sup_t \|g\|_{H^{s_A}(l)} \left(\frac{1}{(1+(D\frac{\tau}{2\varepsilon})^2)} \right)^{l/2} \\
&\quad\leq \sup_t \|f\|_{H^{s_A}(l)} \sup_t \|g\|_{H^{s_A}(l)} \left(\frac{1}{(1+(D\frac{\tau}{2\varepsilon})^2)} \right). \qquad (2.2.4)
\end{aligned}$$

where we used the abbreviation $D := \omega'(k_1) - \omega'(k_2)$ for the group velocity difference. As a consequence, we get

$$\begin{aligned}
&\int_0^{\underline{T}} \|f(\underline{X}_1,\tau)g(\underline{X}_2,\tau)\|_{H^{s_A}}\, d\tau \\
&\quad\leq \sup_t \|f\|_{H^{s_A}(l)} \sup_t \|g\|_{H^{s_A}(l)} \int_0^{\underline{T}} \left(\frac{1}{(1+(D\frac{\tau}{2\varepsilon})^2)} \right) d\tau \\
&\quad\leq \sup_t \|f\|_{H^{s_A}(l)} \sup_t \|g\|_{H^{s_A}(l)} \frac{2\varepsilon}{D} \arctan\left(\frac{D}{2\varepsilon}\underline{T} \right) \\
&\quad\leq \sup_t \|f\|_{H^{s_A}(l)} \sup_t \|g\|_{H^{s_A}(l)} \frac{2\varepsilon}{D} C.
\end{aligned}$$

So, again, we can shift the coupling terms — exactly as in the last section — with a factor ε^{-1} into the next order to get

Theorem 2.2.1. *Let u^{an} be according to* (2.2.1) *and let $A_j^{(1)} \in H^{s_A}(l) \cap H^{s_A+l}(0), j = 1, 2, l \geq 2, s_A \geq 5$, solve the decoupled Nonlinear Schrödinger equations*

$$\begin{aligned}
-2i\omega(k_1)\partial_2 A_1^{(1)} &= (1-\omega'(k_1)^2)\partial_1^2 A_1^{(1)} + 3|A_1^{(1)}|^2 A_1^{(1)}, \\
-2i\omega(k_2)\partial_2 A_2^{(1)} &= (1-\omega'(k_2)^2)\partial_1^2 A_2^{(1)} + 3|A_2^{(1)}|^2 A_2^{(1)}.
\end{aligned}$$

Then there is some $C > 0$ with

$$\|\mathrm{Res}\{u^{an}\}\|_{H^{s_A}} \leq C\varepsilon^{7/2} + C\varepsilon^{7/2}\varepsilon^{-1}\left(\frac{1}{(1+(D\frac{\tau}{2\varepsilon})^2)} \right),$$

and, thus,

$$\sup_{t\in[0,\underline{T}_0/\varepsilon^2]} \|\mathrm{Res}\{u^{an}\}\|_{H^{s_A}} \leq C\varepsilon^{7/2}.$$

Hence, for every $T_0 > 0$ there is $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ there is a solution u of the cubic Klein-Gordon equation (1.0.1) *with*

$$\|u - u^{an}\|_{H^s} \leq C\varepsilon^{3/2},$$

if $1 \leq s \leq s_A - 2$.

Proof. The proof goes line-by-line as in Ch. 1 for Thm. 1.2.1 with $r_{acc} = 1, res = 7/2$ and $\delta = 3/2$, the only additional difficulty being the estimation of the coupling terms. However, we have already discussed this issue in (2.2.4). □

To put this in words, at first order the two pulses do not feel each others presence, i.e. there is no interaction at all!

Higher order corrections. Since the two pulses are decoupled in first order, we will have to increase the accuracy to describe interaction effects (which have to exist due to the nonlinearity). However, we will see that, when sticking to the classical approach, higher order approximations $A^{(2)}, A^{(3)}, \ldots$ will have to satisfy equations that are far too complicated to render a deeper insight into interaction behavior:

To capture interaction effects of second and third order, we can employ the "naive ansatz"

$$\begin{aligned} u^{an}(x,t) &= \sum_{r=1}^{3} \sum_{j=1,2} \varepsilon^r A_j^{(n)}(X_j, \underline{T}) E_j + c.c., \\ E_j &= e^{ik_j x + i\omega(k_j)t}, \quad \underline{X}_j = \underline{X} + \varepsilon\omega'(k_j)t, \quad \underline{X} := \varepsilon x, \quad \underline{T} = \varepsilon^2 t. \end{aligned} \tag{2.2.5}$$

Analyzing the corresponding residual, we get the modulation system for $A_j^{(r)}, r = 1, 2, 3, j = 1, 2$ given by

$$-2i\omega(k_j)\partial_2 A_j^{(r)} = (1 - \omega'(k_j)^2)\partial_1^2 A_j^{(r)} - 2\omega'(k_j)\partial_{1,2} A_j^{(r-1)} - \partial_2^2 A_j^{(r-2)} + G_{j,internal}^{(r)} + G_{j,inter}^{(r)},$$

again with the convention $A_j^{(r)} = 0, r \leq 0$. The *internal dynamics terms* are given by

$$G_{j,internal}^{(r)} := 3 \sum_{2r_1 + r_2 = r+2} (A_j^{(r_1)})^2 \overline{A_j^{(r_2)}} + 6 \sum_{\substack{r_1 + r_2 + r_3 = r+2, \\ r_1 \neq r_3}} A_j^{(r_1)} \overline{A_j^{(r_2)}} A_j^{(r_3)}. \tag{2.2.6}$$

The *interaction dynamics terms* are given by

$$G_{j,inter}^{(r)} := 6 \sum_{r_1 + r_2 + r_3 = r+2} A_{3-j}^{(r_1)} \overline{A_{3-j}^{(r_2)}} A_j^{(r_3)}. \tag{2.2.7}$$

Note, that we cannot shift all coupling terms beyond $\mathcal{O}(\varepsilon^5)$ since the estimate (2.2.4) only allows to shift the terms *at most one order* higher (in contrast to the case for intrachannel interaction where we were able to shift them arbitrarily by choosing m_0, see Sec. 2.1). So we are really left with a highly coupled system of equations which do not disclose anything about the higher order interaction behavior. Not even the $\mathcal{O}(1)$ boundedness of $A_j^{(r)}$ is clear, since by estimate (2.2.4) the equations actually depend on ε in a singular way.

2.2.2 Separation of internal and interaction dynamics

Since we deal with a nonlinear equation we anticipate the interaction to cause some effects, as for example a shape deformation or a position shift. We will incorporate exactly these expectations into the "naive ansatz" (2.2.5). This procedure has been introduced to us by the work of L. Tkeshelashvili *et al* from 2004 [TPB04], while it originates in the work of M. Oikawa and N. Yajima from 1974 [OY74]. We have already discussed this in [CBCSU08].

The extended ansatz reads (see the Rem. 2.2.2 below for a motivation for the choice of variables)

$$\begin{aligned} u_{ansatz} &= \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r A_j^{(r)}(\underline{X}_j,\underline{T})E_j + M_{mixed,\alpha} + c.c., \\ E_j &= e^{ik_jx - i\omega(k_j)t + i\sum_{r=1,2}\varepsilon^r\Omega_j^{(r)}(\underline{X}_{3-j},\underline{T})}, \quad \underline{T} = \varepsilon^2 t, \\ \underline{X}_j &= \varepsilon(x + \omega'(k_j)t + \varepsilon\Psi_j^{(1)}(\underline{X}_{3-j},\underline{T})), \end{aligned} \tag{2.2.8}$$

where the structure of $M_{mixed,\alpha}$ will be discussed at the end of this section (see Sec. 2.2.4).

Remark 2.2.2. *Note, that the new quantities $\Omega_j^{(1)}, \Omega_j^{(2)}$ and $\Psi_j^{(1)}$ account for pure interaction effects (we explain this in great detail in the course of this section) and therefore depend on the position of the respectively other pulse, i.e.*

$$\Omega_1^{(1)} = \Omega_1^{(1)}(\underline{X}_2), \quad \Omega_2^{(1)} = \Omega_2^{(1)}(\underline{X}_1),$$

and so on. To be able to use the more compact notation with sums, we write instead

$$\Omega_j^{(1)} = \Omega_j^{(1)}(\underline{X}_{3-j}), \quad j = 1, 2.$$

To convey the notion why this complicated ansatz is indeed very natural and really simplifies the problem, we go through each new term separately and explain its purpose. We do this by successively adding $\Omega_j^{(1)}, \Omega_j^{(2)}$ and $\Psi_j^{(1)}$ to the "naive ansatz" (2.2.1) and tracking the corresponding terms in the residual.

Remark 2.2.3. *The connection between the extended ansatz and classical approach becomes clear, if we perform Taylor expansion around the "classical variables" $\underline{X}_j = \varepsilon(x + \omega'(k_j))$ and $Y_j = ik_jx + i\omega(k_j)t$*

$$\begin{aligned} u^{an}(x,t) = \quad & \varepsilon[A_1^{(1)}]e^{Y_1} + \varepsilon^2[A_1^{(2)} + iA_1^{(1)}\Omega_1^{(1)}]e^{Y_1} \\ & + \varepsilon^3[A_1^{(3)} + (\partial_1 A_1^{(1)})\Psi_1^{(1)} + iA^{(2)}\Omega_1^{(1)} + iA_1^{(1)}\Omega_1^{(2)} - \frac{1}{2}A_1^{(1)}(\Omega_1^{(1)})^2]e^{Y_1} + \dots \end{aligned}$$

Evidently, the term $\Omega_1^{(1)}$ is actually a correction of second order, while $\Omega_1^{(2)}, \Psi_1^{(1)}$ really embody corrections of third order.
Consequently, keeping in mind the requirements of Thm. 1.2.1 we will need $r_{acc} \geq 3$ to assure that $\Omega_j^{(1)}, \Omega_j^{(2)}$ and $\Psi_j^{(1)}$ are correct.
Furthermore, to guarantee that the quantities $A_j^{(r)}$ do not describe interaction effects, but merely account for the internal dynamics of the pulse, we will have to show that $A_j^{(r)}$ can be chosen to satisfy decoupled equations.

What effect does $\Omega_j^{(1)}$ describe?

As already discussed, when considering the naive ansatz (2.2.5) we are faced at order $\mathcal{O}(\varepsilon^3 E_j)$ in the residual with a system of coupled Nonlinear Schrödinger equations

$$-2i\omega(k_j)\partial_2 A_j^{(1)} = (1-\omega'(k_j)^2)\partial_1^2 A_j^{(1)} + 3|A_j^{(1)}|^2 A_j^{(1)} + 6|A_{3-j}^{(1)}|^2 A_j^{(1)}, \quad j=1,2.$$

The computations of the last sections have demontrated that this system is effectively decoupled (cf. Thm. 2.2.1). But instead of merely shifting the coupling terms one order higher — as we have done it in Thm. 2.2.1 — we use them to extract more specific information about the interaction:

Adding the terms $\Omega_j^{(1)}$ to the classical ansatz (2.2.5) modifies the coupled Nonlinear Schrödinger equations to

$$\begin{aligned}-2i\omega(k_j)\partial_2 A_j^{(1)} =&(1-\omega'(k_j)^2)\partial_1^2 A_j^{(1)} + 3|A_j^{(1)}|^2 A_j^{(1)} \\ &+ 6|A_{3-j}^{(1)}|^2 A_j^{(1)} + 2(\omega(k_j)\omega'(k_{3-j}) - k_j)(\partial_1\Omega_j^{(1)})A_j^{(1)},\end{aligned}$$

which leads us to the following natural choice of solvability conditions:
The amplitude $A_j^{(1)}$ can be chosen to solve the *decoupled* Nonlinear Schrödinger equation

$$-2i\omega(k_j)\partial_2 A_j^{(1)} =(1-\omega'(k_j)^2)\partial_1^2 A_j^{(1)} + 3|A_j^{(1)}|^2 A_j^{(1)} \tag{2.2.9}$$

and $\Omega_j^{(1)}$ has to satisfy a formula reading

$$\partial_1\Omega_j^{(1)} = \frac{3}{\omega(k_j)D_j}|A_{3-j}^{(1)}|^2. \tag{2.2.10}$$

with the difference of the group velocities denoted by

$$D_j := \omega'(k_{3-j}) - \omega'(k_j).$$

Clearly, $\Omega_j^{(1)}$ is a real quantity and, hence, the modification $e^{i\Omega_j^{(1)}}$ has the effect of a *carrier phase shift* (see Fig. 2.3). To give an example, if $A_j^{(1)}$ are sech-profiles, then $\Omega_j^{(1)}$ have tanh-profiles.

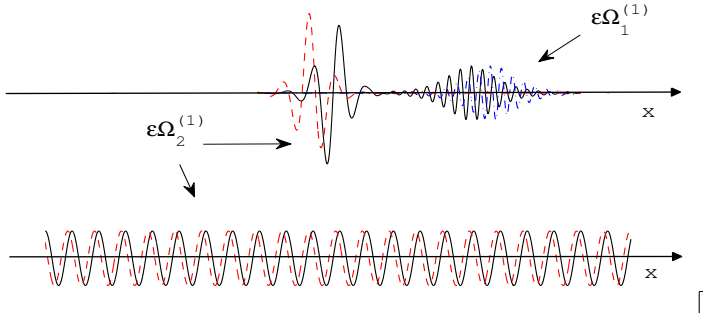

Figure 2.3: *Illustration of the carrier shift after collision. Top:* The two-pulse after interaction (solid) vs. the two pulse traveling by themselves (dashed); *Bottom:* The carrier belonging to the slower pulse in the presence of the second pulse (solid) and without it (dashed)

Some remarks.

1. *Note, that making a modification of the carrier of the form*

$$E_j = e^{ik_j x - i\omega(k_j)t + i\Omega_j^{(0)}(\underline{X}_{3-j},T)}$$

would have resulted in a term $2(\omega(k_j)\omega'(k_{3-j}) - k_j)(\partial_1\Omega_j^{(0)})A_j^{(1)}$ *at order* $\mathcal{O}(\varepsilon^2)$*. But since there are no terms at all at this order, it is impossible to derive a solvability condition other than*

$$\Omega_j^{(0)} = 0.$$

This is in perfect correspondence with what we have already found out by the qualitative analysis summerized in Thm. 2.2.1, i.e. at first order the pulses are completely decoupled and the interaction effects only become visible at second order (and, indeed, $\Omega_j^{(1)}$ *is a second order effect due to Rem. 2.2.3).*

2. *At this point it is not clear if* $\Omega_j^{(1)}$ *captures all the interaction effects at second order (e.g. the entire shift), since - as pointed out before -* $\Omega_j^{(1)}$ *really lives at second order, so* $A_j^{(2)}$ *could also contain a part of the interaction effects. It will turn out that this is not the case, since* $A_j^{(2)}$ *can be chosen to solve an equation without any coupling terms (see next paragraph).*

3. *So far, we accomplished to formally cancel all terms in the residual up to order* $\mathcal{O}(\varepsilon^4)$ *(in front of the pure hamonics).*

What effects do $\Omega_j^{(2)}$ and $\Psi_j^{(1)}$ describe?

Consider the coupled modulation equations for $A_j^{(2)}, j = 1, 2$, at order $\mathcal{O}(\varepsilon^4 E_j)$ in the residual for the naive ansatz (2.2.5), i.e. (2.2.6) for $r = 2$. It is a system of coupled linear Schrödinger equations with rather nasty inhomogeneous terms

$$\begin{aligned} -2\omega(k_j)i\partial_2 A_j^{(2)} = \; & (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} \\ & + (6A_j^{(2)}\overline{A_j^{(1)}} + 3A_j^{(1)}\overline{A_j^{(2)}})A_j^{(1)} - 2\omega'(k_j)\partial_{1,2}A_j^{(1)} \\ & + 6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}})A_j^{(1)} \end{aligned}$$

Remember, we added $\Omega_j^{(1)}, j = 1, 2$ to the ansatz, which helped us to deal with the coupling at third order. However, this modification also adds terms at fourth order, i.e. the naive ansatz with the terms $\Omega_j^{(1)}, n = 1, 2$, added results in

$$\begin{aligned} -2\omega(k_j)i\partial_2\, A_j^{(2)} = \; & (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} \\ & + (6A_j^{(2)}\overline{A_j^{(1)}} + 3A_j^{(1)}\overline{A_j^{(2)}})A_j^{(1)} - 2\omega'(k_j)\partial_{1,2}A_j^{(1)} \\ & + [6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}}) - 2\omega\partial_2\Omega_j^{(1)} + i(1-\omega'(k_j)^2)\partial_1^2\Omega_j^{(1)}]A_j^{(1)} \\ & + [2i(1-\omega'(k_j)\omega'(k_{3-j}))\partial_1\Omega_j^{(1)}]\partial_1 A_j^{(1)} \end{aligned}$$

So it seems that we have gotten us into even more trouble at this order. But this can be elegantly resolved by introducing $\Psi_j^{(1)}, j = 1, 2$, to the ansatz such that the equations read

$$\begin{aligned}
-2\omega(k_j)i\partial_2 A_j^{(2)} = \; & (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} \\
& + (6A_j^{(2)}\overline{A_j^{(1)}} + 3A_j^{(1)}\overline{A_j^{(2)}})A_j^{(1)} - 2\omega'(k_j)\partial_{1,2}A_j^{(1)} \\
& + [6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}}) - 2\omega(k_j)\partial_2\Omega_j^{(1)} \\
& \qquad + i(1-\omega'(k_j)^2)\partial_1^2\Omega_j^{(1)}]A_j^{(1)} \\
& + 2i[(1-\omega'(k_j)\omega'(k_{3-j}))\partial_1\Omega_j^{(1)} \\
& \qquad + (k_j - \omega(k_j)\omega'(k_{3-j}))\partial_1\Psi_j^{(1)}]\partial_1 A_j^{(1)}
\end{aligned}$$

Note, that in order to emphasize the structure of the equation, we sorted the equation for internal dynamics and interaction dynamics terms, and then also sorted for terms proportional to $A_j^{(1)}$ and $\partial_1 A_j^{(1)}$. Now this structure makes

$$\begin{aligned}
\partial_1\Psi_j^{(1)} &= \frac{1}{\omega(k_j)D_j}\left(1-\omega'(k_j)\omega'(k_{3-j})\right)\partial_1\Omega_j^{(1)} \\
&= \frac{3}{\omega(k_j)D_j^2}\left(1-\omega'(k_j)\omega'(k_{3-j})\right)|A_{3-j}^{(1)}|^2. \qquad (2.2.11)
\end{aligned}$$

a natural solvability condition. Note, that $\Psi_j^{(1)}$ is a real quantity and therefore plays the role of an *envelope position shift* (see Fig. 2.4). More specific, if we choose for $A_j^{(1)}$ sech-profiles, then $\Omega_j^{(1)}$ would have tanh-profiles.

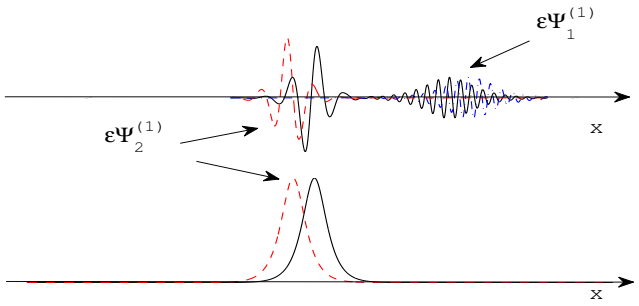

Figure 2.4: *Illustration of the envelope shift after collision. Top:* The two-pulse after interaction (solid) vs. the two pulse traveling by themselves (dashed); *Bottom:* The envelope belonging to the slower pulse in the presence of the second pulse (solid) and without it (dashed)

Through the choice of the above solvability conditions, the above equation reduces to

$$\begin{aligned}-2\omega(k_j)i\partial_2 A_j^{(2)} = \; & (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} \\ & + (6A_j^{(2)}\overline{A_j^{(1)}} + 3A_j^{(1)}\overline{A_j^{(2)}})A_j^{(1)} - 2\omega'(k_j)\partial_{1,2}A_j^{(1)} \\ & + [6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-n}^{(1)}\overline{A_{3-j}^{(2)}}) - 2\omega(k_j)\partial_2\Omega_j^{(1)} \\ & \quad + i(1-\omega'(k_j)^2)\partial_1^2\Omega_j^{(1)}]A_j^{(1)}.\end{aligned}$$

Now further adding $\Omega_j^{(2)}$ will result in

$$\begin{aligned}-2\omega(k_j)i\partial_2\; A_j^{(2)} = \; & (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} \\ & + (6A_j^{(2)}\overline{A_j^{(1)}} + 3A_j^{(1)}\overline{A_j^{(2)}})A_j^{(1)} - 2\omega'(k_j)\partial_{1,2}A_j^{(1)} \\ & + [6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-n}^{(2)}}) - 2\omega(k_j)\partial_2\Omega_j^{(1)} \\ & \quad + i(1-\omega'(k_j)^2)\partial_1^2\Omega_j^{(1)} \\ & \quad\quad + 2(\omega'(k_j)\omega'(k_{3-j}) - k_j)\partial_1\Omega_j^{(2)}]A_j^{(1)},\end{aligned}$$

and evidently this structure indicates the formula

$$\partial_1\Omega_j^{(2)} = \frac{1}{2\omega(k_j)D_j}[6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}}) - 2\omega(k_j)\partial_2\Omega_j^{(1)} + i(1-\omega'(k_j)^2)\partial_1^2\Omega_j^{(1)}]. \tag{2.2.12}$$

Note, that $\Omega_j^{(2)}$ has both a real and an imaginary part. To be precise,

$$\begin{aligned}\mathrm{Re}\{\Omega_j^{(2)}(\underline{X}_{3-j},\underline{T})\} &= \frac{1}{2\omega(k_j)D_j}\int_{-\infty}^{\underline{X}_{3-j}}[6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}}) - 2\omega(k_j)\partial_2\Omega_j^{(1)}](\xi,\underline{T})\; d\xi, \\ \mathrm{Im}\{\Omega_j^{(2)}(\underline{X}_{3-j},\underline{T})\} &= \frac{1}{2\omega(k_j)D_j}(1-\omega'(k_j)^2)\partial_1\Omega_j^{(1)} \\ &= \frac{3}{2\omega(k_j)^2D_j^2}(1-\omega'(k_j)^2)|A_{3-j}^{(1)}|^2,\end{aligned} \tag{2.2.13}$$

since $\Omega_j^{(1)}$ is real and $\overline{A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}}} = A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}}$. Consequently, $\mathrm{Re}\{\Omega_j^{(2)}\}$ will play the role of a second order carrier shift, while $\mathrm{Im}\{\Omega_j^{(2)}\}$ will account for a deformation of the pulse shape (see Fig. 2.5). Note, that $\mathrm{Im}\{\Omega_j^{(2)}\}$ inherits the spatial localization of $A_{3-j}^{(1)}$.

In summary, by introducing $\Omega_j^{(2)}$ and $\Psi_j^{(1)}$ we can choose $A_j^{(2)}$ to satisfy

$$\begin{aligned}-2\omega(k_j)i\partial_2\; A_j^{(2)} = \; & (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} \\ & + (6A_j^{(2)}\overline{A_j^{(1)}} + 3A_j^{(1)}\overline{A_j^{(2)}})A_j^{(1)} - 2\omega'(k_j)\partial_{1,2}A_j^{(1)},\end{aligned} \tag{2.2.14}$$

which exclusively features internal dynamics terms. Hence, we managed to decouple the system of Linear Schrödinger equations (much like we have done it at $\mathcal{O}(\varepsilon^3)$ for the coupled Nonlinear Schrödinger equations).

Some remarks.

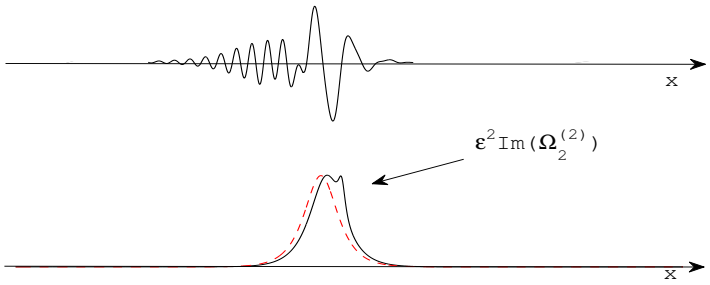

Figure 2.5: *Illustration of the shape deformation after collision. Top:* The two-pulse after interaction (solid) vs. the two pulse traveling by themselves (dashed); *Bottom:* The envelope belonging to the slower pulse in the presence of the second pulse (solid) and without it (dashed)

1. *Again, as already explained for* $\Omega_j^{(0)}$ *in the remark of the last paragraph,* $\Psi_j^{(0)}$ *would have shown up at third order and could have not been associated with any term since it is proportional to* $\partial_1 A_j^{(1)}$.

2. *We were worried in the last paragraph that* $\Omega_j^{(1)}$ *might not capture all interaction effects, since it is essentially a second order correction and* $A_j^{(2)}$ *might also contribute to capturing interaction effects. Clearly, this is not the case since* $A_j^{(2)}$ *has to satisfy* (2.2.14) *which does not contain interaction terms.*

3. *We have to worry now about the validity of* $\Omega_j^{(2)}$ *and* $\Psi_j^{(1)}$ *instead, which live at third order (again by Taylor expansion). Hence, we have to make sure that* $A_j^{(3)}$ *does not describe interaction effects, either. This is discussed in the next paragraph*

4. *At this point, we are left with a residual of order* $\mathcal{O}(\varepsilon^5)$ *(well, at least for the residual part involving pure harmonics, see Sec. 2.2.4 a discussion of the mixed and higher order harmonics).*

Qualitative decoupling

Considering the coupled modulation equations for $A_j^{(2)}, j = 1, 2$, at order $\mathcal{O}(\varepsilon^5 E_j)$ in the residual for the naive ansatz (2.2.5), i.e. (2.2.6) for $r = 3$, we see that we have to deal with highly coupled equations. Hence, we have difficulties to "sort and associate" terms in the spirit of the last sections. Therefore, we only explain the structure of the equations for $A_j^{(3)}$ and explain why they actually decouple in first order.

Our extended ansatz (2.2.8) leads to an equation for $A_j^{(3)}$ of the form

$$-2\omega(k_j)i\partial_2 A_j^{(3)} = (1-\omega'(k_j)^2)\partial_1^2 A_j^{(3)} - 2\omega'(k_j)\partial_{1,2} A_j^{(2)} - \partial_2^2 A_j^{(1)} + G_{j,internal}^{(3)} + G_{j,inter}^{(3)} + F_{j,extended}$$

where $F_{j,extended}$ embodies all the coupling terms that are caused by the introduction of the new quantities $\Omega_j^{(1)}, \Omega_j^{(2)}$ and $\Psi_j^{(1)}$. Note, that these coupling terms are generated by the linear part of the residual as opposed to $G_{j,inter}^{(3)}$ and $G_{j,internal}^{(3)}$ which are caused by the nonlinear part.

Due to the mechanism of the coupling term estimate (2.2.4) we can simply demand that $A_j^{(3)}, j = 1, 2$, solves the decoupled modulation equation

$$-2\omega(k_j)i\partial_2 A_j^{(3)} = (1-\omega'(k_j)^2)\partial_1^2 A_j^{(3)} + 2\omega'(k_j)\partial_{1,2}A_j^{(2)} - \partial_2^2 A_j^{(1)} + G_{j,internal}^{(3)} \tag{2.2.15}$$

and shift the coupling terms

$$\varepsilon^{-1}(G_{j,inter}^{(3)} + F_{j,extended})$$

into the next order (exactly as we have done it in Thm. 2.2.1, were we showed that the Nonlinear Schrödinger equations for $A_j^{(1)}$ are really decoupled).

As a consequence, we are left with a residual of order $\mathcal{O}(\varepsilon^6)$ for the residual part involving pure harmonics! Now, let us turn our attention to the cancelation of the terms involving mixed and higher order harmonics.

Remark 2.2.4. (Canceling mixed and higher order harmonics) *Exactly as in Sec. 1.1 the term $M_{\text{mixed},\alpha}$ in (2.2.1) accounts for terms involving higher order or mixed harmonics, i.e. for the frequencies which are generated by the nonlinearity. Again, we refrain from giving the exact formula for $M_{mixed,\alpha}$ and rather explain its structure by an example:*
At $\varepsilon^3 E_1^2 E_2$ for example the term $(A_1^{(1)})^2 A_2^{(1)}$ appears. To cancel this term we extend the ansatz by $\alpha_{21}\varepsilon^3 (A_1^{(1)})^2 A_2^{(1)} E_1^2 E_2$ and get an algebraic equation for α_{21} of the form

$$\left(1 + (2\mathrm{i}\omega(k_1) + \mathrm{i}\omega(k_2))^2 + (2\mathrm{i}k_1 + \mathrm{i}k_2)^2\right)\alpha_{21} = 3.$$

The procedure is essentially the same for each such term leading to equations of the form

$$\left(1 + (l\omega(k_1) + j\omega(k_2))^2 + (lk_1 + jk_2)^2\right)\alpha_{lj} = \beta_{lj}.$$

for the unknowns α_{lj}. Now $M_{mixed,\alpha}$ contains all these extensions. For α_{lj} to be well-defined, we need to assure the so-called "non-resonance" condition

$$|\left(1 + (l\omega(k_1) + j\omega(k_2))^2 + (lk_1 + jk_2)^2\right)| > 0.$$

Approximation result

Putting everything from this section together we get the extended modulation system given by a system describing *internal dynamics*

$$-2i\omega(k_j)\partial_2 A_j^{(1)} = (1-\omega'(k_j)^2)\partial_1^2 A_j^{(1)} + 3|A_j^{(1)}|^2 A_j^{(1)}, \tag{2.2.16}$$

$$-2\omega(k_j)i\partial_2 A_j^{(2)} = (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} + 2\omega'(k_j)\partial_{1,2}A_j^{(1)} + G_{j,internal}^{(2)}, \tag{2.2.17}$$

$$-2\omega(k_j)i\partial_2 A_j^{(3)} = (1-\omega'(k_j)^2)\partial_1^2 A_j^{(3)} + 2\omega'(k_j)\partial_{1,2}A_j^{(2)} - \partial_2^2 A_j^{(1)} + G_{j,internal}^{(3)}, \tag{2.2.18}$$

and a system describing *interaction dynamics*

$$\partial_1 \Omega_j^{(1)} = -\frac{3}{\omega(k_j)D_j}|A_{3-j}^{(1)}|^2, \tag{2.2.19}$$

$$\partial_1 \Psi_j^{(1)} = -\left(\frac{1-\omega'(k_j)\omega'(k_{3-j})}{\omega(k_j)D_j}\right)\partial_1 \Omega_j^{(1)}, \tag{2.2.20}$$

$$\partial_1 \Omega_j^{(2)} = \frac{1}{2\omega(k_j)D_j}[6(A_{3-j}^{(2)}\overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)}\overline{A_{3-j}^{(2)}}) - 2\omega(k_j)\partial_2\Omega_j^{(1)} + i(1-\omega'(k_j)^2)\partial_1^2\Omega_j^{(1)}]. \tag{2.2.21}$$

As alluded to before, this choice make the residual formally of order $\mathcal{O}(\varepsilon^6)$. Let us make this rigorous:

Lemma 2.2.5. *Fix $s, s_A \geq 1$ with $1 \leq s \leq s_A - 10$. Let u^{an} be according to (2.2.8) and let*

$$A_{n,m}^{(r)} \in C([0,T_0]; H^{s_A-3(r-1)}(l) \cap H^{s_A-3(r-1)+l}(0)), s_A \geq 10, l \geq 2,$$

and $\Omega_j^{(r)}, \Psi_j^{(1)}$ fulfill the extended modulation system (2.2.16)-(2.2.21). Then

$$\|\mathrm{Res}(u^{an})\|_{H^s} \leq C\varepsilon^{11/2} + C\varepsilon^{11/2}\varepsilon^{-1}\left(\frac{1}{(1+(D\frac{\tau}{2\varepsilon})^2)}\right)$$

and hence

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\mathrm{Res}(u^{an})\|_{H^s} \leq C\varepsilon^{11/2}.$$

Proof. The proof only amounts to verifying formal calculations and use the estimate (2.2.4). □

Theorem 2.2.6. *Given the conditions of Lemma (2.2.5), for all $T_0 > 0$ there is $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(\cdot,t) - u^{an}(\cdot,t)\|_{C_b^{s-1}(\mathbb{R})} \leq C\varepsilon^{7/2}, \tag{2.2.22}$$

with $C > 0$ a constant independent of ε and where u is an exact two-pulse solution of the cubic Klein-Gordon equation (1.0.1).

Proof. The proof goes line-by-line as in Ch. 1 for Thm. 1.2.1 with $r_{acc} = 3, res = 11/2$ and $\delta = 7/2$. □

Validity of interaction effect formulas

Since $A_n^{(1)}, A_n^{(2)}$ and $A_n^{(3)}$ can be chosen to solve decoupled equations, the terms $\Omega_n^{(1)}, \Omega_n^{(2)}$ and $\Psi_n^{(1)}$ capture all interaction effects of first, second and third order. In other words, we managed to *separate internal and interaction dynamics* up to third order!

Remark 2.2.7. *A major difficulty to overcome in our computations was finding the right way of "sorting terms" at order $\mathcal{O}(\varepsilon^4)$ such that the formerly coupled equations for $A_1^{(2)}, A_2^{(2)}$, i.e. (2.2.6) for $r = 2$ given by*

$$-2i\omega(k_1)\partial_2 A_1^{(2)} = (1-\omega'(k_1)^2)\partial_1^2 A_1^{(2)} - 2\omega'(k_1)\partial_{1,2}A_1^{(1)} + G_{1,internal}^{(2)} + G_{1,inter}^{(2)},$$

$$-2i\omega(k_2)\partial_2 A_2^{(2)} = (1-\omega'(k_2)^2)\partial_1^2 A_2^{(2)} - 2\omega'(k_2)\partial_{1,2}A_2^{(1)} + G_{2,internal}^{(2)} + G_{2,inter}^{(2)}$$

decouple. We managed to do so by introducing the quantities $\Omega_j^{(2)}$ and $\Psi_j^{(1)}$ and associating the corresponding terms in the residual to the interaction dynamics terms $G_{j,inter}^{(2)}$. We have explained this in great detail in the last paragraphs. Note, that this sorting procedure is very difficult to see at first sight. However, motivated by numerical simulations of pulse interaction conducted by Christopher Chong, we were convinced that a decoupling should be possible (see [Ch09]).
Nevertheless, so far there is no indication that the sorting will work in a similar way also for the coupled equations for $A_1^{(3)}, A_2^{(3)}$.

2.3 Multipulse solutions

This section will merge the results on inter- and intrachannel interaction to describe solutions which are important for application in nonlinear optics (see Ch. 7), namely N pulse trains each consisting of M equidistant pulses, hereafter simply referred to as *multipulse*.

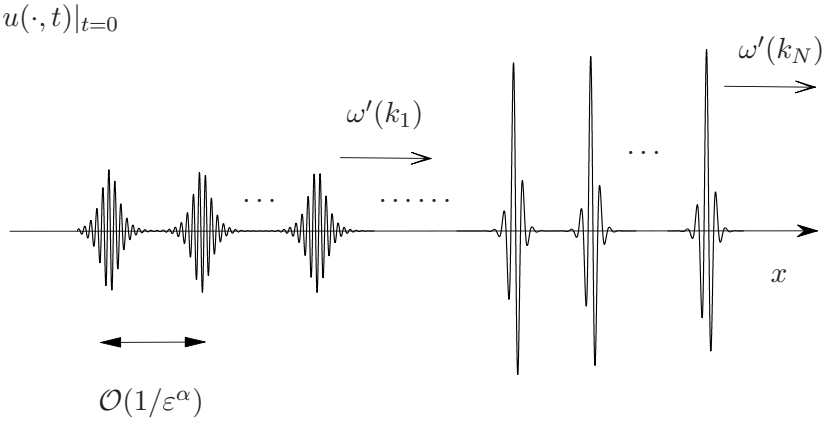

Figure 2.6: A multipulse consisting of N pulse trains each consisting of M equidistant pulses

In order to approximate such a multipulse we make the ansatz

$$
\begin{aligned}
u^{an} &= \sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{r=1,2,3}\varepsilon^r A^{(r)}_{n,m}(\underline{X}_{n,m},\underline{T})E_{n,m} + M_{mixed,\alpha} + c.c., \qquad (2.3.1)\\
E_{n,m} &= e^{ik_nx+i\omega(k_n)t+i\sum_{r=1,2}\varepsilon^r\Omega^{(r)}_{n,m}(\underline{Y}_{n,m},\underline{T})}, \quad \underline{T}=\varepsilon^2 t,\\
\underline{X}_{n,m} &= \varepsilon(x+\omega'(k_n)t+\varepsilon d_m+\varepsilon\Psi^{(1)}_{n,m}(\underline{Y}_{n,m},\underline{T})),\\
\underline{Y}_{n,m} &= (\ldots;\underline{X}_{n-1,1},\ldots,\underline{X}_{n-1,M};\underline{X}_{n+1,1},\ldots,\underline{X}_{n+1,M};\ldots),
\end{aligned}
$$

with $|d_n-d_m|=\mathcal{O}(1/\varepsilon^\alpha), n\neq m$, and where $M_{mixed,\alpha}$ is constructed as described in Sec. (2.2.4). The variable $\underline{Y}_{n,m}$ should express, that the functions $\Omega^{(r)}_{n,m},\Psi^{(1)}_{n,m}$ depend on the positions of all pulses except the ones belonging to the same channel, i.e. they only describe pure interchannel interaction effects, since we deal with intrachannel interaction differently (see Sec. 2.1).

Plugging this ansatz into the cubic Klein-Gordon equation (1.0.1) gives the following *modulation system*:
The *modulation equations* for the envelopes $A^{(r)}_{n,m}$, i.e. the *internal dynamics equations*

$$
-2i\omega(k_n)\partial_2 A^{(r)}_{n,m} = (1-\omega'(k_n)^2)\partial_1^2 A^{(r)}_{n,m} - 2\omega'(k_n)\partial_{1,2}A^{(r-1)}_{n,m} - \partial_2^2 A^{(r-2)}_{n,m} + G^{(r)}_{n,m,internal}, \qquad (2.3.2)
$$

where we again used the convention $A^{(r)}_{n,m}=0, r\leq 0$, and with

$$
G^{(r)}_{n,m,internal} := 3\sum_{2r_1+r_2=r+2}(A^{(r_1)}_{n,m})^2\overline{A^{(r_2)}_{n,m}} + 6\sum_{\substack{r_1+r_2+r_3=r+2,\\ r_1\neq r_3}} A^{(r_1)}_{n,m}\overline{A^{(r_2)}_{n,m}}A^{(r_3)}_{n,m}.
$$

The formulas for the effects of interaction, namely the *first order carrier shift*

$$\Omega^{(1)}_{n,m}(\underline{Y}_{n,m},\underline{T}) = 3\sum_{(\tilde{n},\tilde{m}),\tilde{n}\neq n}\frac{1}{\omega(k_n)D_{n,m}}\int_{-\infty}^{\underline{X}_{\tilde{n},\tilde{m}}}|A^{(1)}_{\tilde{n},\tilde{m}}(\xi,\underline{T})|^2\,d\xi, \tag{2.3.3}$$

where $D_{n,m} := \omega'(k_n) - \omega'(k_{3-m})$, the *first order envelope shift*

$$\Psi^{(1)}_{n,m}(\underline{Y}_{n,m},\underline{T}) = 3\sum_{(\tilde{n},\tilde{m}),\tilde{n}\neq n}\frac{1-\omega'(k_n)\omega'(k_{3-n})}{\omega(k_n)^2D^2_{n,m}}\int_{-\infty}^{\underline{X}_{\tilde{n},\tilde{m}}}|A^{(1)}_{\tilde{n},\tilde{m}}(\xi,\underline{T})|^2\,d\xi, \tag{2.3.4}$$

the *second order correction of the pulse shape*

$$\mathrm{Im}\{\Omega^{(2)}_{n,m}(\underline{Y}_{n,m},\underline{T})\} = \frac{(1-\omega'(k_{3-n})^2}{2\omega(k_n)^2D^2_{n,m}}\partial_1\Omega^{(1)}_{n,m}(\underline{Y}_{n,m},\underline{T}), \tag{2.3.5}$$

and the *second order carrier shift*

$$\begin{aligned}&\mathrm{Re}\{\Omega^{(2)}_{n,m}(\underline{Y}_{n,m},\underline{T})\}\\ &= -3\sum_{(\tilde{n},\tilde{m}),\tilde{n}\neq n}[\frac{1}{\omega(k_n)D_{n,m}}\int_{-\infty}^{\underline{X}_{\tilde{n},\tilde{m}}}(A^{(1)}_{\tilde{n},\tilde{m}}\overline{A^{(2)}_{\tilde{n},\tilde{m}}} + A^{(2)}_{\tilde{n},\tilde{m}}\overline{A^{(1)}_{\tilde{n},\tilde{m}}}) - \frac{\omega(k_n)}{3}\partial_2\Omega^{(1)}_{n,m}\,d\xi].\end{aligned} \tag{2.3.6}$$

Lemma 2.3.1. *Let u^{an} be the multipulse ansatz* (2.3.1) *such that*

$$A^{(r)}_{n,m} \in C\left([0,T_0];H^{s_A-3(r-1)}(l)\cap H^{s_A-3(r-1)+l}(0)\right),\quad s_A\geq 10, l\geq 2,$$

solve the internal dynamics equations (2.3.2) *and $\Omega^{(r)}_{n,m}, \Psi^{(1)}_{n,m}$ solve the interaction dynamics equations* (2.3.3)-(2.3.6)*. Then for $1\leq s\leq s_A-10$ and $l(\alpha-1)-2\geq 3$, we get*

$$\|\mathrm{Res}\{u^{an}\}\|_{H^s}\leq C_{N,M}\varepsilon^{11/2},$$

uniformly in time, with $C_{N,M}>0$ independent of ε, but depending on the number of pulses.

Proof. Let us first analyze a special case of the extended ansatz, namely with the assumption

$$\Omega^{(1)}_{1,1}(Y_{1,1},T) = \Omega^{(1)}_{1,1}(\underline{X}_{1,2},\dots,\underline{X}_{N,M},T) = \sum_{(n,m),n\neq 1}\widetilde{\Omega^{(1)}_{1,1}}(\underline{X}_{n,m})$$

and respectively for all (n,m), i.e. we simply assume that the dependence is additively separable. This makes it possible to directly generalize our computations from the previous section from two to arbitrarily many pulses by simply adding the formulas. In detail, we get for each $\Omega^{(1)}_{1,1}$ a total of $N(M-1)$ formulas

$$\widetilde{\Omega^{(1)}_{1,1}}(\underline{X}_{n,m}) = \frac{3}{(k_1-\omega(k_1)\omega'(k_n))}\int_{-\infty}^{\underline{X}_{n,m}}|A^{(1)}_{n,m}(\xi,T)|^2\,d\xi, \tag{2.3.7}$$

and hence

$$\Omega^{(1)}_{1,1}(Y_{1,1},T) = \sum_{(\tilde{n},\tilde{m}),\tilde{n}\neq 1}\frac{3}{(\omega(k_1)D_{1,n}}\int_{-\infty}^{\underline{X}_{\tilde{n},\tilde{m}}}|A^{(1)}_{\tilde{n},\tilde{m}}(\xi,T)|^2\,d\xi,$$

with $D_{1,n} := \omega'(k_1) - \omega'(k_n)$, and respectively for each $\Omega^{(1)}_{n,m}$. Consequently, the total carrier shift is the sum of the carrier shifts gained through interaction with all pulses traveling on other carrier waves (remember that we deal with intrachannel interaction in a different way, see Sec. 2.1).

Similarly, assuming the additive separability also for $\Omega^{(2)}_{n,m}$ and $\Psi^{(1)}_{n,m}$ we get by the same reasoning the corresponding formulas (2.3.4)-(2.3.6).

If we proceed to the general case, i.e. leave the case of additive separability, the only new terms which appear in the residual are mixed derivatives, for example

$$\partial_{\underline{X}_{2,1}} \partial_{\underline{X}_{3,4}} \Omega^{(1)}_{1,1}(\underline{Y}_{n,m}) A_{n,m}.$$

However, since this is again a product of spatially localized functions with different velocities, we can again shifte it one order higher by using the estimate (2.2.4). □

In the end, by using exactly the proof technique as for Thm. 1.2.1 the following approximation theorem for multipulses:

Theorem 2.3.2. *Given the conditions of Lemma 2.3.1, for all $T_0 > 0$ there are $\varepsilon_0 > 0$ and $C > 0$ such that*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(\cdot,t) - u^{an}(\cdot,t)\|_{C_b^{s-1}(\mathbb{R})} \leq C\varepsilon^{7/2}, \tag{2.3.8}$$

where u is a multipulse solution of the cubic Klein-Gordon equation (1.0.1).

2.3.1 ε-dependent number of carriers

The question arises (see Ch. 7 for a motivation) what happens, if $N = \mathcal{O}(1/\varepsilon^{\alpha_N})$ for $\alpha_N > 0$. One way of achieving this is to distribute the wave numbers k_n between some arbitrary, but fixed K_1 and K_2 with $K_n = \mathcal{O}(1)$. Then, of course, $|k_n - k_l| = \mathcal{O}(\varepsilon^{\alpha_N})$ and, so,

$$|\omega'(k_n) - \omega'(k_l)| = \mathcal{O}(\varepsilon^{\alpha_N}).$$

However, for the carrier shift we have that

$$\Omega^{(1)}_n = \mathcal{O}(\frac{1}{|\omega'(k_n) - \omega'(k_l)|}) = \mathcal{O}(1/\varepsilon^{\alpha_N}),$$

and for the envelope shift we have that

$$\Psi^{(1)}_n = \mathcal{O}(\frac{1}{|\omega'(k_n) - \omega'(k_l)|^2}) = \mathcal{O}(1/\varepsilon^{2\alpha_N}).$$

Consequently, this choice leads to an increased shift.
Another way of realizing $N = \mathcal{O}(1/\varepsilon^{\alpha_N})$ is by setting $|k_n - k_l| = \mathcal{O}(1)$ and choosing $K_1 = \mathcal{O}(1), K_2 = \mathcal{O}(1/\varepsilon^{\alpha_N})$. However, this will destroy the multiscale character of the modulation equation, since

$$\omega''(1/\varepsilon^{\alpha_N}) \longrightarrow 0, \quad \varepsilon \longrightarrow 0.$$

Note, that this is not a problem when dealing with a nonlinear wave equation

$$\partial_t^2 u = \partial_x^2 u - \alpha\partial_x^4 u - u + u^3, \quad \alpha > 0,$$

since its dispersion relation is given by $\omega^2 = \alpha k^4 + k^2 + 1$ and so

$$\omega''(1/\varepsilon^{\alpha_N}) \longrightarrow \sqrt{\alpha}, \quad \varepsilon \longrightarrow 0.$$

In summary, in the case of the cubic Klein-Gordon equation an ε-dependent number of wave numbers is not possible if one wants to keep the bounds on the shifts caused by interaction.

2.4 Chapter Summary

This chapter presents a refined perturbation approach that leads to the separation of internal and interaction dynamics for a multipulse in the setting of a cubic Klein-Gordon equation

$$\partial_t^2 u = \partial_x^2 u - u + u^3, \quad x, t, u = u(x,t) \in \mathbb{R}.$$

We break down the analysis of the multipulse by separately addressing two different types of pulse interactions (were we borrowed the terminology from optical communications, cf. Ch. 7). Let us give a brief account of what we have found:

Intrachannel interaction. Pulses traveling on the same carrier wave (as depicted in Fig. 2.1) can be approximated on the long time scale of order $\mathcal{O}(1/\varepsilon^2)$ by the ansatz

$$u^{an}(x,t) = \sum_{n=1,2} \sum_{r=1}^{3} \varepsilon^r A_n^{(r)}(\varepsilon(x + \omega'(k_1)t + d_n), \varepsilon^2 t) e^{ik_1 x + i\omega(k_1)t} + M_{mixed,\alpha} + c.c.,$$

if each $A_n^{(r)}$ fulfills the recursively solvable modulation system for internal dynamics, i.e.

$$\begin{aligned}
-2i\omega(k_1)\partial_2 A_n^{(1)} &= (1-\omega'(k_1)^2)\partial_1^2 A_n^{(1)} + 3|A_n^{(1)}|^2 A_n^{(1)},\\
-2i\omega(k_1)\partial_2 A_n^{(2)} &= (1-\omega'(k_1)^2)\partial_1^2 A_n^{(2)} - 2\omega'(k_1)\partial_{1,2} A_n^{(1)} + 3(A_n^{(1)})^2\overline{A}_n^{(2)} + 6A_n^{(2)}A_n^{(1)}\overline{A}_n^{(1)},\\
-2\omega(k_1)i\partial_2 A_n^{(3)} &= (1-\omega'(k_1)^2)\partial_1^2 A_n^{(3)} - 2\omega'(k_1)\partial_{1,2} A_n^{(2)} - \partial_2^2 A_n^{(1)} + G_{n,internal}^{(3)},
\end{aligned}$$

and if the relation $l(1-\alpha) \geq 5$ between the localization strength l and the spacing parameter α with $|d_2 - d_1| = \mathcal{O}(1/\varepsilon^\alpha)$ is respected. Again, we can prove the validity of the approximation for the long time scale of order $\mathcal{O}(1/\varepsilon^2)$ by a slightly modified version of the method in [KSM92].
In other words, we managed to derive a condition that ensures a complete suppression of interaction effects at leading approximation orders for a long time.

Interchannel interaction. Pulses traveling on different carriers (as depicted in Fig. 2.2) and, hence, with different velocities can be approximated by the extended ansatz

$$\begin{aligned}
u_{ansatz} &= \sum_{j=1,2} \sum_{r=1}^{3} \varepsilon^r A_j^{(r)}(\underline{X}_j, \underline{T}) E_j + M_{mixed,\alpha} + c.c.,\\
E_j &= e^{ik_j x - i\omega(k_j)t + i\sum_{r=1,2}\varepsilon^r \Omega_j^{(r)}(\underline{X}_{3-j}, \underline{T})}, \quad \underline{T} = \varepsilon^2 t,\\
\underline{X}_j &= \varepsilon(x + \omega'(k_j)t + \varepsilon\Psi_j^{(1)}(\underline{X}_{3-j}, \underline{T})),
\end{aligned}$$

if each $A_n^{(r)}$ fulfills the recursively solvable modulation system for internal dynamics, i.e.

$$\begin{aligned}
-2i\omega(k_j)\partial_2 A_j^{(1)} &= (1-\omega'(k_j)^2)\partial_1^2 A_j^{(1)} + 3|A_j^{(1)}|^2 A_j^{(1)},\\
-2i\omega(k_j)\partial_2 A_j^{(2)} &= (1-\omega'(k_j)^2)\partial_1^2 A_j^{(2)} - 2\omega'(k_j)\partial_{1,2} A_j^{(1)} + 3(A_j^{(1)})^2\overline{A}_j^{(2)} + 6A_j^{(2)}A_j^{(1)}\overline{A}_j^{(1)},\\
-2\omega(k_j)i\partial_2 A_j^{(3)} &= (1-\omega'(k_j)^2)\partial_1^2 A_j^{(3)} - 2\omega'(k_j)\partial_{1,2} A_j^{(2)} - \partial_2^2 A_j^{(1)} + G_{j,internal}^{(3)}
\end{aligned}$$

and the interaction quantities solve the following equations:
The **carrier shift** $\Omega_j^{(1)}$ (see Fig. 2.3) has to satisfy the ordinary differential equation

$$\partial_1 \Omega_j^{(1)} = \frac{3}{\omega(k_j) D_j} |A_{3-j}^{(1)}|^2, \quad D_j := \omega'(k_{3-j}) - \omega'(k_j).$$

The **envelope position shift** $\Psi_j^{(1)}$ (see Fig. 2.4) has to satisfy the ordinary differential equation

$$\partial_1 \Psi_j^{(1)} = \frac{1}{\omega(k_j) D_j} \left(1 - \omega'(k_j)\omega'(k_{3-j})\right) \partial_1 \Omega_j^{(1)}.$$

Furthermore, the **second order carrier shift and shape correction** are described by the real and imaginary part of $\Omega_j^{(2)}$ which has to fulfill

$$\partial_1 \Omega_j^{(2)} = \frac{1}{2\omega(k_j) D_j} [6(A_{3-j}^{(2)} \overline{A_{3-j}^{(1)}} + A_{3-j}^{(1)} \overline{A_{3-j}^{(2)}}) - 2\omega(k_j)\partial_2 \Omega_j^{(1)} + i(1 - \omega'(k_j)^2)\partial_1^2 \Omega_j^{(1)}].$$

Note, that the for intrachannel interaction both pulse envelopes have to obey the same modulation system (meaning with the same coefficients), whereas for the interchannel interaction case, each pulse has its "own" modulation system.

To sum up, the separate description of *pure internal dynamics of each pulse given by the modulation system for* $A_j^{(r)}, r = 1, 2, 3$, *on the one side* and of *pure interaction dynamics given by the three ordinary differential equations for* $\Omega_j^{(1)}, \Omega_j^{(2)}$ *and* $\Psi_j^{(1)}$ *on the other side* embodies what we hereafter refer to as "separation of internal and interaction dynamics". This separation made it possible to merge our results for intra- and interchannel interaction to the description of a $N \times M$ *multipulse* (see Fig. 2.6), i.e. N pulse trains each consisting of M equidistant pulses, which is a scenario that models "Wavelength-Division multiplexing" techniques in optical communication lines (cf. Ch. 7).

Note, that the rest of the present work will be concerned with transferring the "separation of internal and interaction dynamics of multipulses" to more complex settings than the cubic Klein-Gordon equation. We will, however, only explain how the analysis for a two-pulse representing interchannel interaction can be carried out, since this is indeed the major difficulty to overcome for the description of a multipulse.

Part II

Pulse interaction in "natural coordinates" and applications

3 Pulse interaction for a system in "natural coordinates"

In this chapter we report about the framework that allows us to transfer the "separation of internal and interaction dynamics" presented in Ch. 2 from the setting of a cubic Klein-Gordon equation to equations exhibiting a more complex dispersion relation (see Fig. 3.1).

The chapter is organized as follows:

In the first section we introduce the notion of a "system in natural coordinates" and explain how it helps us to transfer the results from the last chapter to a very general setting.

The second section contains exactly this transfer, i.e. we derive an extended modulation system like (2.2.16)-(2.2.21) for a two-pulse ansatz like (2.2.8)[1]. In fact, this is one of the center pieces of our work and we will make use of the outcome of the calculations in each of the next chapters!

Finally, we briefly display the general strategy for pulse interaction analysis in the context of nonlinear wave equations with constant coefficients. This will facilitate the exposition in Ch. 4 and Ch. 5.

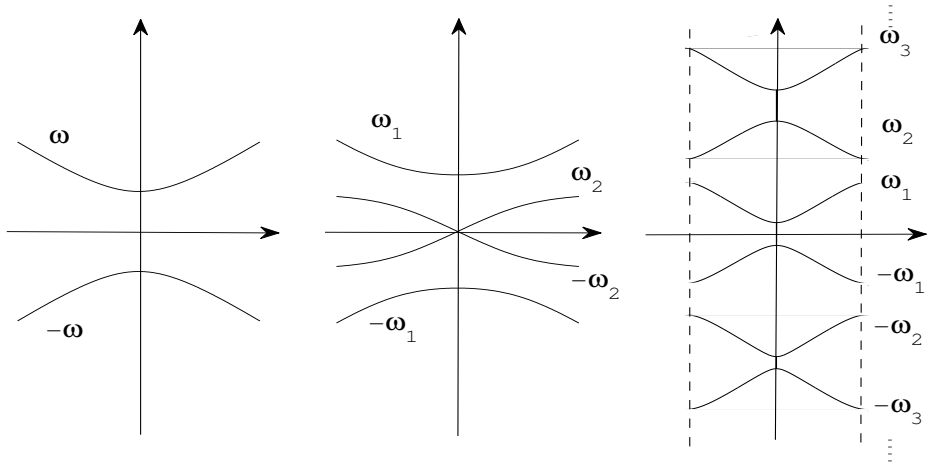

Figure 3.1: *Left panel:* Dispersion relation of a cubic Klein-Gordon equation (1.0.1); *Middle panel:* Dispersion relation of a Maxwell-Lorentz system (5.0.1)-(5.0.2); *Right panel:* Dispersion relation of a cubic Klein-Gordon equation with periodic coefficients (6.0.1)

[1]From now on we will only focus on two-pulse solutions representing interchannel interaction, i.e. two pulses with different carrier waves (see Sec. 2.2), even though we could really transfer all the results from Ch. 2. However, we refrain from doing this, since it would not involve any new mathematical challenges.

3.1 Transformation to "natural coordinates"

We motivate the idea of transforming an equation to its representation in "natural coordinates" by reviewing one of the main aspects of the "method of deriving modulation equations", in general.

Spectral concentration. Consider the spectral content of a two-pulse ansatz

$$u^{an}(x,t) := \varepsilon A_1(\underline{X}_1, \underline{T})e^{ik_1 x + i\omega_1(k_1)t} + \varepsilon A_2(\underline{X}_2, \underline{T})e^{ik_2 x + i\omega_1(k_2)t} + c.c.$$

with

$$\underline{X}_j := \underline{X} + \varepsilon\omega'(k_j)t, \quad \underline{X} := \varepsilon x, \quad \underline{T} := \varepsilon^2 t,$$

as one would make for the cubic Klein-Gordon (1.0.1) equation (cf. Sec. 2.2). Its spectral content is concentrated in an ε-neighborhood around the corresponding wave numbers (well, that is how we constructed it in first place, cf. (1.1.2)), and, hence, we can deduce that they really only feel these local properties of the dispersion relation. Indeed, this is the key observation for the "derivation of modulation equations", in general, and will lead the way to the desired transfer from the cubic Klein-Gordon equation to a "more difficult" setting. By "more difficult" we mean a setting where the dispersion relation is more complex, e.g. it has more branches (see Fig. 3.1) as it happens, for example, for the Maxwell-Lorentz system (see Ch. 5) or a nonlinear wave equation with periodic coefficients (see Ch. 6).

The transfer. Now, what we have reported about in Ch. 2 is how to carry out pulse interaction analysis for the cubic Klein-Gordon equation (1.0.1) whose dispersion relation has just one pair of branches. But, in fact, a two-pulse ansatz for a more complicated equation will actually also only "feel" the pair(s) of branches to which the temporal frequencies $\omega_{n_j}(k_j)$ belong (see Fig. 3.2). Intuitively, one would expect that, regardless of the complexity of the dispersion relation, we can always restrict the analysis either, if $n_1 = n_2$, to a two-dimensional subsystem, or, if $n_1 \neq n_2$, to a four-dimensional subsystem belonging to the respective pair(s) of branches. Such a reduction would directly enable us to carry over the methods from the last chapter. However, in order to implement this reasoning to a given equation, we will have to perform a transformation that brings it to a form where the branches of the dispersion relation are explicitly visible. It is this form that we refer to as "system in natural coordinates".
For notational simplicity we will restrict ourselves throughout this work to the case where the temporal frequencies belong to the same branch, i.e. $n_1 = n_2$. The other case, where $n_1 \neq n_2$, can also be treated by the subsequent analysis, but this case is not very well suited for exposition, since it would lead to a four- instead of a two-dimensional system.

Natural coordinates. The "system in natural coordinates" is assumed to have the structure

$$\begin{aligned}
\partial_t \tilde{V}_1 &= \begin{pmatrix} i\omega_1 & 0 \\ 0 & -i\omega_1 \end{pmatrix} \tilde{V}_1 + \tilde{N}_1[(\tilde{V}_n)_{n\in\mathcal{N}}], \\
\partial_t \tilde{V}_2 &= \begin{pmatrix} i\omega_2 & 0 \\ 0 & -i\omega_2 \end{pmatrix} \tilde{V}_2 + \tilde{N}_2[(\tilde{V}_n)_{n\in\mathcal{N}}], \\
&\vdots
\end{aligned}$$

or, written in a more compact form,

$$\partial_t \tilde{V}_n = \tilde{D}_n \tilde{V}_n + \tilde{N}_n[(\tilde{V}_n)_{n\in\mathcal{N}}], \tag{3.1.1}$$

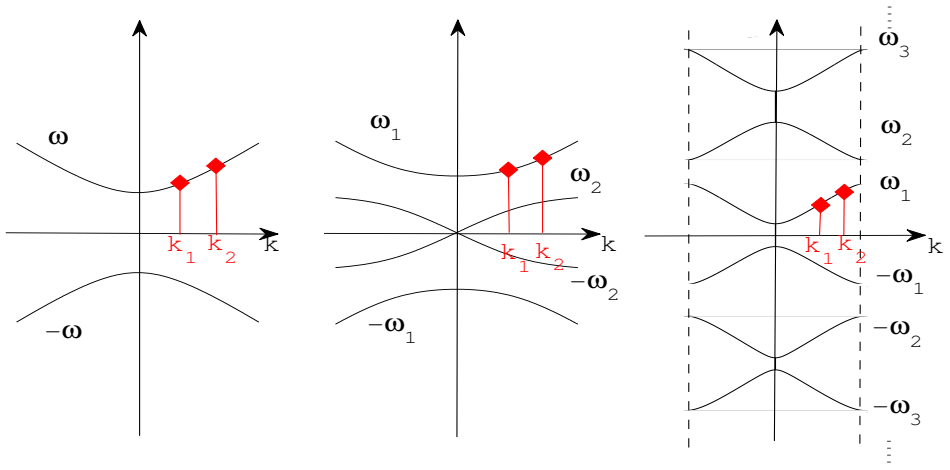

Figure 3.2: Concentration of spectral content for two-pulse solutions

for $\tilde{V}_n := (\tilde{v}_{n,1}, \tilde{v}_{n,2})^T, \tilde{v}_{n,l} = \tilde{v}_{n,l}(k,t) \in \mathbb{C}, n \in \mathcal{N}, l = 1,2$, where $\mathcal{N}$ is an at most countable index set, with linear part

$$\tilde{D}_n = \tilde{D}_n(k) := \begin{pmatrix} i\omega_n(k) & 0 \\ 0 & -i\omega_n(k) \end{pmatrix},$$

where $\omega_n : \mathbb{R} \longrightarrow \mathbb{R}$ are nonlinear, piece-wise analytic functions, and with the nonlinear operator $\tilde{N}_n$ which is at least cubic and whose k-dependence is also piece-wise analytic. To breathe life into this abstract setting, we close this section by forestalling for the three equations which will be considered in the next three chapters the structure of the transformation and the dimension of the corresponding "system in natural coordinates".

For the case of the constant coefficient cubic Klein-Gordon equation (1.0.1), we have the index set $\mathcal{N} = \{1\}$. Thus, the corresponding "system in natural coordinates" (3.1.1) is two-dimensional, which is consistent with the fact that the dispersion relation has one pair of branches (see Fig. 3.1). The transformation to natural coordinates is given by $u = [\mathcal{F}\{Q\tilde{V}\}]_1$ where $\mathcal{F}$ is the Fourier transform and $\tilde{Q}$ is the 2×2-matrix that diagonalizes the linear part of the equation converted into a first order system (cf. Ch. 4).

For the case of the Maxwell-Lorentz system (5.0.1)-(5.0.2), we have the index set $\mathcal{N} = \{1,2\}$ and, so, the corresponding "system in natural coordinates" (3.1.1) has two two-dimensional subsystems, which is consistent with the fact that the dispersion relation has two pairs of branches (see Fig. 3.1). The transformation to natural coordinates is given by $u = [\mathcal{F}\{Q(k)\tilde{V}\}]_1$ where $\mathcal{F}$ is the Fourier transform and $\tilde{Q}(k)$ is the 4×4-matrix that diagonalizes the linear part of the equation converted into a first order system (cf. Ch. 5).

For the case of a periodic coefficient nonlinear wave equation we have the index set $\mathcal{N} = \mathbb{N}$ and so the corresponding "system in natural coordinates" (3.1.1) is a countably infinite collection of two-dimensional subsystems, which is in correspondence with the fact that the dispersion relation has infinitely many pairs of branches (see Fig. 3.1). The transformation to natural coordinates is a little more involved consisting of Bloch transformation followed by an eigenfunction representation and a diagonalizing transformation (cf. Ch. 6).

Note that this brief listing is only meant to appeal to the readers intuition, even though he might not understand the details, which will be elaborated in the next chapters.

3.2 Derivation of modulation equations

Guided by the ideas from the last chapter, we will construct a two-pulse ansatz $\tilde{V}^{an} := (\tilde{V}_n^{an})_{n\in\mathcal{N}}$ that diminishes the residual corresponding to the "system in natural coordinates" (3.1.1), i.e.

$$\widetilde{\mathrm{Res}}\{\tilde{V}\} := \left(-\partial_t \tilde{V}_n + \tilde{D}_n \tilde{V}_n + \tilde{N}_n[\tilde{V}]\right)_{n\in\mathcal{N}} := \left(\widetilde{\mathrm{Res}_n}\{\tilde{V}\}\right)_{n\in\mathcal{N}}, \tag{3.2.1}$$

with $\tilde{V} := (\tilde{V}_n)_{n\in\mathcal{N}}$ and thereby derive a modulation system as (2.2.16)-(2.2.21).

Note that we cannot directly employ the extended two-pulse ansatz (2.2.8) from the last chapter, since it is not possible to Fourier transform it due to the appearance of nested functions like $e^{i\varepsilon\Omega_j^{(r)}(\underline{X}_{3-j})}$. Hence, we will have to go back to x-space in order to be able to use the extended ansatz. But before returning to x-space, we will first make explicit use of the spectral concentration property of pulses that we have already alluded to in the first section.

Remark on the Fourier transform of pulse solutions Note that when we switch from x-space to Fourier space, we lose ε^{-1} due to the integration involved in the Fourier transform. In detail, we have

$$\begin{aligned}
\mathcal{F}\{\varepsilon A_1(\varepsilon(\cdot+\omega'(k_j)t),\underline{T})e^{ik_1(\cdot)+i\omega_1(k_1)t}\}(k) &= \varepsilon\int_{\mathbb{R}} A_1(\underline{X}_1,\underline{T})e^{ik_1x+i\omega_1(k_1)t}e^{-ikx}\,dx\\
&= \varepsilon\int_{\mathbb{R}} A_1(\underline{X}_1,\underline{T})e^{i\frac{k_1-k}{\varepsilon}\varepsilon x}\,dx\,e^{i\omega_1(k_1)t}\\
&= \varepsilon\,\varepsilon^{-1}\int_{\mathbb{R}} A_1(\underline{X}_1,\underline{T})e^{i\underline{K}_1\underline{X}_1}\,d\underline{X}_1\,e^{i\omega_1(k_1)t-i\varepsilon\underline{K}_1\omega'(k_1)t}\\
&= \hat{A}_1(\underline{K}_1)\,\tilde{E}_1
\end{aligned} \tag{3.2.2}$$

where we used the notation $\underline{X} := \varepsilon x, \underline{K}_1 := \frac{k-k_1}{\varepsilon}, \tilde{E}_1 := e^{i\omega_1(k_1)t-i\varepsilon\underline{K}_1\omega'(k_1)t}$.

3.2.1 Spectral concentration of pulse solutions

As already illustrated in Sec. 3.1, the spectral content of pulse solutions is concentrated around the wave number of the carrier wave. Now the heart of the "method of deriving modulation equations" is taking advantage of exactly this concentration by expanding and truncating the dispersion relation around the corresponding wave numbers. It is this "expansion and truncation" that causes the reduction of the original equation to the modulation system. This mechanism also takes place when deriving the modulation system in x-space directly (as we have done it in the last chapters), however, it is more transparent when carrying out the derivation in "natural coordinates".

We want to illustrate how one can exploit the spectral concentration by means of an instructive example that already has most of the features that will appear for the equations that we deal with in the next chapters.

An instructive example.

Let us consider the residual

$$\widetilde{\mathrm{Res}}\{\tilde{V}\} := \widetilde{D}\tilde{V} + (\beta_1\tilde{v}_1 + \beta_2\tilde{v}_2)^{*3}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}, \quad \tilde{V} = (\tilde{v}_1, \tilde{v}_2)^T, \tag{3.2.3}$$

with linear part

$$\widetilde{D}(k)\tilde{V}(k) := \begin{pmatrix} i\omega(k) & 0\\ 0 & -i\omega(k)\end{pmatrix}\tilde{V}(k),$$

where $\omega : \mathbb{R} \to \mathbb{R}$ is piece-wise analytic, and nonlinear part

$$\widetilde{N}[\tilde{V}](k) := (\beta_1(\cdot)\tilde{v}_1(\cdot) + \beta_2(\cdot)\tilde{v}_2(\cdot))^{*3}(k)\begin{pmatrix}\alpha_1(k)\\ \alpha_2(k)\end{pmatrix}, \tag{3.2.4}$$

where $\alpha_j, \beta_j : \mathbb{R} \to \mathbb{C}$ are piece-wise analytic.

Let us first focus on the linear part. If we choose for $k \in \mathbb{R}$

$$\begin{aligned}\tilde{v}_1(k) &= \tilde{B}_1(\underline{K}_1), \quad \underline{K}_1 := \frac{k-k_1}{\varepsilon},\\ \tilde{v}_2(k) &= \tilde{B}_2(\underline{K}_{1-}), \quad \underline{K}_1 := \frac{k-(-k_1)}{\varepsilon},\end{aligned}$$

that is, $\tilde{v}_1$ and $\tilde{v}_2$ localized in an ε-neighborhood around k_1 and $-k_1$ respectively, then

$$\begin{aligned}[\widetilde{D}(k)\tilde{V}(k)]_1 &= i\omega(k)\tilde{v}_1(k)\\ &= i\omega(k)\tilde{B}_1(\underline{K}_1)\\ &= i\sum_{m=0}^{4}\frac{\omega^{(m)}(k_1)}{m!}\varepsilon^m\underline{K}_1^m\tilde{B}_1(\underline{K}_1) + \mathcal{O}(\varepsilon^5),\end{aligned}$$

and similar for $[\widetilde{D}(k)\tilde{V}(k)]_2$, but with expansion around $-k_1$.
Inverse Fourier transformation then yields

$$[\mathcal{F}^{-1}\{\widetilde{D}(\cdot)\tilde{V}(\cdot)\}(x)]_1 = i\sum_{m=0}^{4}\frac{\omega^{(m)}(k_1)}{m!}\varepsilon^{m+1}(-i\partial_{\underline{X}})^m[B_1(\underline{X})]e^{ik_1x} + \mathcal{O}(\varepsilon^6),$$

with B_1 defined by

$$\mathcal{F}^{-1}\{\tilde{B}_1(\frac{\cdot-k_1}{\varepsilon})\}(x) = \int \tilde{B}_1(\frac{k-k_1}{\varepsilon})e^{ikx}\,dk = \varepsilon\underbrace{\int \tilde{B}_1(\underline{K}_1)e^{i\underline{K}_1\underline{X}}\,d\underline{K}_1}_{=:B_1(\underline{X})}e^{ik_1x}, \quad \underline{X} := \varepsilon x.$$

Hence, for $V = (B_1e^{ik_1x}, B_2e^{-ik_1x})^T$ we get with the definition

$$[\underline{D}[V]]_1 := i\sum_{m=0}^{4}\frac{\omega^{(m)}(k_1)}{m!}\varepsilon^{m+1}(-i\partial_{\underline{X}})^m[B_1]e^{ik_1x}, \tag{3.2.5}$$

and respectively for $[\underline{D}[V]]_2$, that, formally,

$$\widetilde{D}[\mathcal{F}\{V\}] = \mathcal{F}\{\underline{D}[V]\} + \mathcal{O}(\varepsilon^5).$$

As a consequence, we see that, in order to get $\widetilde{D}\tilde{V} = \mathcal{O}(\varepsilon^5)$, we simply have to find V such that $\mathcal{F}\{\underline{D}[V]\} = \mathcal{O}(\varepsilon^5)$ or, equivalently,

$$\underline{D}\{V\} = \mathcal{O}(\varepsilon^6),$$

where the loss in the exponent comes from the scaling caused by the integration in the Fourier transform (see (3.2.2)). Note that $\underline{D}$ only involves x-space expressions and is, hence, perfectly suited for our extended ansatz.

Now let us turn our attention to the nonlinear part. Motivated by the observations for the linear part, we set

$$\tilde{V} = \mathcal{F}\{V^{an}\}, \quad V^{an} := (v^{an}, \overline{v^{an}}),$$

with the "classical" two-pulse ansatz (we comment on this choice in a moment)

$$\begin{aligned} v^{an}(x,t) &= \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r B_j^{(r)}(\underline{X}_j, \underline{T}) e^{i\underline{Y}_j}, \\ \underline{X}_j &:= \underline{X} + \varepsilon\omega_1'(k_j)t, \quad \underline{Y}_j := k_j x + \omega_1(k_j)t, \quad \underline{X} := \varepsilon x, \quad \underline{T} := \varepsilon^2 t. \end{aligned}$$

In the following we will make use of the notation from (3.2.2), i.e.

$$\mathcal{F}\{V^{an}\} = (\hat{v}^{an}, \hat{\overline{v}}^{an})^T, \quad \mathcal{F}\{B_j^{(r)}(\varepsilon(\cdot + \omega'(k_j)t), \underline{T}) e^{ik_j\cdot + i\omega_1(k_j)t}\} =: \hat{B}_j^{(r)} \tilde{E}_j$$

and respectively for the complex conjugate terms.
Note that we will solely focus on the terms occurring in the first component and only at pure harmonics, since the rest of the terms can be treated by the same ideas.
Plugging this ansatz into the residual gives

$$\begin{aligned} [\widetilde{N}[\tilde{V}]]_1 \; &= \alpha_1[\beta_1(\hat{B}_1^{(1)}\tilde{E}_1 + \hat{B}_2^{(1)}\tilde{E}_2 + \varepsilon\hat{B}_1^{(2)}\tilde{E}_1 + \varepsilon\hat{B}_2^{(2)}\tilde{E}_2 + \ldots) \\ &\qquad + \beta_2(\hat{\overline{B}}_1^{(1)}\tilde{E}_1 + \hat{\overline{B}}_2^{(1)}\tilde{E}_2 + \varepsilon\hat{\overline{B}}_1^{(2)}\tilde{E}_1 + \varepsilon\hat{\overline{B}}_2^{(2)}\tilde{E}_2 + \ldots)]^{*3}. \end{aligned}$$

Dissolving the convolution and making use of the spectral concentration — by first expanding β_1, β_2 and then α_1 — this further "simplifies" to

$$\begin{aligned} [\widetilde{N}\{\tilde{V}\}]_1 \; &= \varepsilon^2\tilde{E}_1\,(\alpha_1(k_1)\,[3(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_2(-k_1)\hat{\overline{B}}_1^{(1)}) \\ &\qquad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_2)\hat{B}_2^{(1)}) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(1)})] \\ &\quad + \varepsilon^3\tilde{E}_1\,(\alpha_1(k_1)[3(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_2(-k_1)\hat{\overline{B}}_1^{(2)}) \\ &\qquad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_1)\hat{B}_1^{(2)}) * (\beta_2(-k_1)\hat{\overline{B}}_1^{(1)}) \\ &\qquad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_2)\hat{B}_2^{(1)}) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(2)}) \\ &\qquad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_2)\hat{B}_2^{(2)}) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(1)}) \\ &\qquad + 6(\beta_1(k_1)\hat{B}_1^{(2)}) * (\beta_1(k_2)\hat{B}_2^{(1)}) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(1)})] \\ &\quad + \alpha_1(k_1)\,[3(\beta_1'(k_1)(\underline{K}_1\hat{B}_1^{(1)})) * (\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_2(-k_1)\hat{\overline{B}}_1^{(1)}) \end{aligned}$$

$$
\begin{aligned}
. \qquad & + 3(\beta_1(k_1)(\underline{K}_1\hat{B}_1^{(1)})) * (\beta_1'(k_1)(\underline{K}_1\hat{B}_1^{(1)})) * (\beta_2(-k_1)\hat{\overline{B}}_1^{(1)}) \\
& \quad + 3(\beta_1(k_1)(\underline{K}_1\hat{B}_1^{(1)})) * (\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_2'(-k_1)(\underline{K}_{-1}\hat{\overline{B}}_1^{(1)})) \\
& + \alpha_1(k_1)\,[6(\beta_1'(k_1)(\underline{K}_1\hat{B}_1^{(1)})) * (\beta_1(k_2)\hat{B}_2^{(1)}) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(1)}) \\
& \quad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1'(k_2)(\underline{K}_2\hat{B}_2^{(1)})) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(1)}) \\
& \quad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_2)\hat{B}_2^{(1)}) * (\beta_2(-k_2)(\underline{K}_{2-}\hat{\overline{B}}_2^{(1)})) \\
& + \alpha_1'(k_1)\underline{K}_1\,[3(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_2(-k_1)\hat{\overline{B}}_1^{(1)}) \\
& \quad + 6(\beta_1(k_1)\hat{B}_1^{(1)}) * (\beta_1(k_2)\hat{B}_2^{(1)}) * (\beta_2(-k_2)\hat{\overline{B}}_2^{(1)})] \\
& + \dots,
\end{aligned}
$$

where we again introduced the notation $\underline{K}_j := \frac{k-k_j}{\varepsilon}, \underline{K}_{j-} := \frac{k-(-k_j)}{\varepsilon}$ and the prime denotes a derivative w.r.t. k. Note that by $\underline{K}_1\hat{B}_1^{(1)}$ we really mean $\cdot\hat{B}_1^{(1)}(\cdot)$. Moreover, we used

$$
\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{B}(\frac{k-l-m-k_1}{\varepsilon})\tilde{B}(\frac{m-k_1}{\varepsilon})\tilde{B}(\frac{l+k_1}{\varepsilon})e^{ic_1\frac{k-l-m-k_1+l-k_1+m+k_1}{\varepsilon}}\,dm\,dl \\
&\quad = \varepsilon^2\left(\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{B}(\underline{K}_1-\underline{M}-\underline{L})\tilde{B}(\underline{M})\tilde{B}(\underline{L})\,d\underline{M}\,d\underline{L}\right)\tilde{E}_1.
\end{aligned}
$$

Exactly as in our first example, we will define the operator $\underline{N}$ by truncating the expansion at $\mathcal{O}(\varepsilon^4)$ (in Fourier space) and performing inverse Fourier transformation, to get again

$$\widetilde{N}[\mathcal{F}\{V\}] = \mathcal{F}\underline{N}[V] + \mathcal{O}(\varepsilon^5),$$

where

$$
\begin{aligned}
[\underline{N}\{V\}]_1 \; &= \varepsilon^3 E_1\,\alpha_1(k_1)\,[3(\beta_1(k_1)^2\beta_2(-k_1)|B_1^{(1)}|^2B_1^{(1)} + 6\beta_1(k_1)\beta_1(k_2)\beta_2(-k_2)|B_2^{(1)}|^2B_1^{(1)}] \\
&+ \varepsilon^4 E_1\,\alpha_1(k_1)[3\beta_1(k_1)^2\beta_2(-k_1)B_1^{(1)}B_1^{(1)}\overline{B}_1^{(2)} \\
&\quad + 6\beta_1(k_1)^2\beta_2(-k_1)B_1^{(1)}B_1^{(2)}\overline{B}_1^{(1)} \\
&\quad + 6\beta_1(k_1)\beta_1(k_2)\beta_2(-k_2)B_1^{(1)}B_2^{(1)}\overline{B}_2^{(2)} \\
&\quad + 6\beta_1(k_1)\beta_1(k_2)\beta_2(-k_2)B_1^{(1)}B_2^{(2)}\overline{B}_2^{(1)} \\
&\quad + 6\beta_1(k_1)\beta_1(k_2)\beta_2(-k_2)B_1^{(2)}B_2^{(1)}\overline{B}_2^{(1)}] \\
&+ \alpha_1(k_1)\,[3\beta_1'(k_1)\beta_1(k_1)\beta_2(-k_1)(-i\partial_{\underline{X}}B_1^{(1)})B_1^{(1)}\overline{B}_1^{(1)} \\
&\quad + 3\beta_1(k_1)\beta_1'(k_1)\beta_2(-k_1)B_1^{(1)}(-i\partial_{\underline{X}}B_1^{(1)})\overline{B}_1^{(1)} \\
&\quad + 3\beta_1(k_1)\beta_1(k_1)\beta_2'(-k_1)B_1^{(1)}B_1^{(1)}(-i\partial_{\underline{X}}\overline{B}_1^{(1)}) \\
&+ \alpha_1(k_1)\,[6\beta_1'(k_1)\beta_1(k_2)\beta_2(-k_2)(-i\partial_{\underline{X}}B_1^{(1)})B_2^{(1)}\overline{B}_2^{(1)} \\
&\quad + 6\beta_1(k_1)\beta_1'(k_2)\beta_2(-k_2)B_1^{(1)}(-i\partial_{\underline{X}}B_2^{(1)})\overline{B}_2^{(1)}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + 6\beta_1(k_1)\beta_1(k_2)\beta_2'(-k_2)B_1^{(1)}B_2^{(1)}(-i\partial_{\underline{X}}\overline{B}_2^{(1)}) \\
&+ \alpha_1'(k_1)(-i\partial_{\underline{X}})\,[3\beta_1(k_1)\beta_1(k_1)\beta_2(-k_1)B_1^{(1)}B_1^{(1)}\overline{B}_1^{(1)} \\
&\qquad + 6\beta_1(k_1)\beta_1(k_2)\beta_2(-k_2)B_1^{(1)}B_2^{(1)}\overline{B}_2^{(1)})] \\
&+ \dots . \qquad (3.2.6)
\end{aligned}
$$

Let us point out again that we have not explicitly mentioned all the mixed and higher order harmonics that also occur. We will treat them exactly the same, the expansion being around $3k_1, 2k_1 + k_2, \dots$. However, opposed to the nonlinear terms that we have mentioned explicitly in (3.2.6), they do not really play a major role for our analysis and can be canceled from the residual in the same spirit as in the last chapters (see Sec. 3.2.2).

Some remarks.

1. It suffices to consider the "classical" two-pulse ansatz for the nonlinearity, since the derivatives in the nonlinearity only appear at order $\mathcal{O}(\varepsilon^4)$ and higher and

$$
\begin{aligned}
&\partial_{\underline{X}}\left(B_j^{(r)}(\underline{X}_j + \varepsilon^2\Psi_j^{(1)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T}), \underline{T})e^{ik_jx + i\omega_1(k_j)t + i\sum_{l=1,2}\varepsilon^l\Omega_j^{(l)}(\underline{X}+\varepsilon\omega_1'(k_{3-j})t, \underline{T})}\right) \\
&\quad = \partial_{\underline{X}}\left(B_j^{(r)}(\underline{X}_j, \underline{T})e^{ik_jx + i\omega_1(k_j)t}\right) + \mathcal{O}(\varepsilon^2).
\end{aligned}
$$

   Consequently, the first differences will be visible at $\mathcal{O}(\varepsilon^{4+2})$, which is beyond the order we are interested in (recall Rem. 2.2.3).

2. As already explained when we return to x-space we get an additional factor ε (cf. 3.2.2). Therefore, what formally is $\mathcal{O}(\varepsilon^2)$ and $\mathcal{O}(\varepsilon^3)$ in Fourier space will be formally $\mathcal{O}(\varepsilon^3)$ and $\mathcal{O}(\varepsilon^4)$ in x-space. This why we truncate the expansion at $\mathcal{O}(\varepsilon^4)$. Moreover, we only specified terms up to $\mathcal{O}(\varepsilon^4)$ in (3.2.6), since they will contribute to the modulation equations and interaction effect formulas. The terms at $\mathcal{O}(\varepsilon^5)$ will only be treated qualitatively (compare Sec. 2.2.2 or the subsequent Sec. 3.2.3, Remark 3.2.1) and, hence, play a minor role.

In summary, we can introduce the x-space residual

$$
\underline{\mathrm{Res}}\{V\} := \underline{D}[V] + \underline{N}[V], \qquad (3.2.7)
$$

construct an ansatz V^{an} such that $\underline{\mathrm{Res}}\{V^{an}\} = \mathcal{O}(\varepsilon^6)$, and then conclude by the relation

$$
\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\} = \mathcal{F}\underline{\mathrm{Res}}\{V\} + \mathcal{O}(\varepsilon^5), \qquad (3.2.8)
$$

that

$$
\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\} = \mathcal{O}(\varepsilon^5).
$$

Recall that the loss in the exponent is due to the scaling property of the Fourier transform (3.2.2).

Remark 3.2.1. *When we carry out* **formal** *calculations in Fourier space, we only have to cancel everything up to $\mathcal{O}(\varepsilon^5)$ to have that the corresponding residual in x-space is formally $\mathcal{O}(\varepsilon^6)$. Nevertheless, when rigorously bounding the residual the bounds are the same both in Fourier space and in x-space since*

$$
\|\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\|_{L^2(s)} = \|\mathcal{F}^{-1}\{\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\}\|_{H^s} = \|\underline{\mathrm{Res}}\{V^{an}\}\|_{H^s} + C\varepsilon^{11/2} \leq C\varepsilon^{11/2}.
$$

if

$$\|\underline{\mathrm{Res}}\{V^{an}\}\|_{H^s} \leq C\varepsilon^{11/2}.$$

Note that the residual (3.2.4) from this example will reappear in all of the next chapters, since it has the structure of the "main part" of the residual "in natural coordinates" for all the equations treated in the next chapters. By "main part" we mean the two-dimensional subsystem which belongs to the pair of branches in the dispersion relation to which the temporal frequencies of the pulses belong (see Fig. 3.2 and Sec. 3.1). Hence, after bringing each equation to its "natural coordinates" representation (3.1.1) and trying to diminish the corresponding $\widetilde{\mathrm{Res}}$, we will always reference the above computations when introducing $\underline{\mathrm{Res}}$, but then, of course, with specific α_n, β_n.

Justification of expansion and truncation

The main estimate employed to justify the expansion and truncation of the dispersion relation is given in the following basic lemma which is often used in the theory of modulation equations.

Lemma 3.2.2. *Let $\mu : \mathbb{R} \longrightarrow \mathbb{R}$ be analytic and*

$$v(x) := \varepsilon B(\varepsilon x)e^{ik_1 x},$$

with $0 < \varepsilon \ll 1, k_1 \in \mathbb{R}$ and where $B \in H^{s_B}$. Then

$$\|(\mathcal{F}^{-1}\mu\mathcal{F})v - (\mathcal{F}^{-1}(\sum_{m=0}^{n} \frac{\mu^{(m)}(k_1)}{m!}(\cdot - k_1)^m)\mathcal{F})v\|_{H^s} \leq C\varepsilon^{n+1+1/2}\|B\|_{H^{s_B}}, \tag{3.2.9}$$

for $s < s_B - (n+1)$ and $C > 0$.

Proof. The proof merely uses Taylor expansion and that $\mathcal{F}$ is an isometry between H^s and $L^2(s)$. In detail,

$$\begin{aligned}
&\|(\mathcal{F}^{-1}\mu\mathcal{F})v - (\mathcal{F}^{-1}(\sum_{m=0}^{n} \frac{\mu^{(m)}(k_1)}{m!}(\cdot - k_1)^m)\mathcal{F})v\|_{H^s} \\
&\leq C\varepsilon^{1/2}\|(\sum_{m=n+1}^{\infty} \frac{\mu^{(m)}(k_1)}{m!}\varepsilon^m(\cdot)^m)\hat{B}(\cdot)\|_{L^2(s)} \\
&\leq C\varepsilon^{1/2+(n+1)} \sup_{\underline{K}_1 \in \mathbb{R}} \left| \frac{|\underline{K}_1|^s \underline{K}_1^{n+1}}{(1+\underline{K}_1^2)^{s_B/2}} \right| \|\hat{B}\|_{L^2(s_B)} \\
&\leq C\varepsilon^{1/2+(n+1)}\|B\|_{H^{s_B}},
\end{aligned}$$

if $s < s_B - (n+1)$. □

Remark. It is actually sufficient to demand that μ is $s_B \geq 1$ times differentiable.

As already explained for the example, we can conclude — and rigorously prove using Lemma 3.2.2 — that for a two-pulse ansatz $(V_n^{an})_{n\in\mathbb{N}}$

$$\widetilde{\mathrm{Res}}\{(\mathcal{F}\{V_n^{an}\})_{n\in\mathcal{N}}\} = \mathcal{F}\underline{\mathrm{Res}}\{(V_n)_{n\in\mathcal{N}}\} + \mathcal{O}(\varepsilon^5),$$

where $\underline{\mathrm{Res}}\{(V_n)_{n\in\mathcal{N}}\}$ according to (3.2.7) is obtained by expansion of the k-dependent components of $\widetilde{\mathrm{Res}}$, truncation and subsequent inverse Fourier transformation (see the example in Sec. 3.2.1).

We want to point out again, that the necessity of introducing $\underline{\mathrm{Res}}\{(V_n)_{n\in\mathcal{N}}\}$ comes from the fact that it only involves x-space expressions, which enables us to use the extended two-wave ansatz (2.2.8), which turned out to be the key to understanding pulse interaction, i.e. to the "separation of internal and interaction dynamics".

3.2.2 Extended modulation system

Consider the extended two-pulse ansatz (motivated by the results in Sec. 2.2)

$$V_1^{an} = (v^{an}, \overline{v^{an}})^T, \quad V_n^{an} := (M^n_{mixed,\alpha}, \overline{M^n_{mixed,\alpha}})^T, n \geq 2, \tag{3.2.10}$$

with

$$\begin{aligned}
v^{an}(x,t) &= \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r B_j^{(r)}(\underline{X}_j, \underline{T}) e^{i\underline{Y}_j} + M^1_{mixed,\alpha},\\
\underline{X}_j &:= \underline{X} + \varepsilon\omega_1'(k_j)t + \varepsilon^2\Psi_j^{(1)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T}),\\
\underline{Y}_j &:= k_j x + \omega_1(k_j)t + \sum_{l=1,2}\varepsilon^l\Omega_j^{(l)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T}),\\
\underline{X} &:= \varepsilon x, \quad \underline{T} := \varepsilon^2 t,
\end{aligned}$$

and the mixed and higher order harmonic terms

$$M^n_{mixed,\alpha} = \mathcal{O}(\varepsilon^3), \quad n \geq 1,$$

whose exact structure will be explained at the end of this section.

Note that this ansatz directly implements the reasoning illustrated in the first section, where we explained that, no matter how complex the dispersion relation might be, the analysis can by restricted to the two-dimensional subsystem belonging to the branches where the pulses "live". Since for a "system in natural coordinates" the branches are explicitly visible, we can directly employ this reduction through the two-pulse ansatz (3.2.10).

Plugging this ansatz into the residual $\underline{\mathrm{Res}}$ as introduced in (3.2.8) — i.e. using the spectral concentration property and inverse Fourier transformation— we can derive solvability conditions exactly as in the last chapters to get the extended modulation system given by the equations for *internal dynamics*

$$\partial_2 B_j^{(1)} = -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(1)} + \tilde{N}_{1,j}^{internal}(B_j^{(1)}), \tag{3.2.11}$$

$$\partial_2 B_j^{(2)} = -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2}, \tag{3.2.12}$$

$$\partial_2 B_j^{(3)} = -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega_1^{(4)}(k_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\leq 3}, \tag{3.2.13}$$

and those for *interaction dynamics*

$$\partial_1 \Omega_j^{(1)} = \frac{i}{D_j} \tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)}), \tag{3.2.14}$$

$$\partial_1 \Psi_j^{(1)} = \frac{1}{D_j}(\tilde{N}_{2,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(q)})_{q\leq 2} + \omega_1''(k_j)(\partial_1 \Omega_j^{(1)})), \tag{3.2.15}$$

$$\partial_1 \Omega_j^{(2)} = \frac{i}{D_j}(\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(q)})_{q\leq l} + \frac{\omega_1''(k_j)}{2}(\partial_1^2 \Omega_j^{(1)}) - i(\partial_2 \Omega_j^{(1)})). \tag{3.2.16}$$

with the group velocity difference $D_j := \omega_1'(k_j) - \omega_1'(k_{3-j})$. In addition, we get a system of equations — and corresponding "non-resonance conditions" — from the cancelation of mixed and higher order harmonics (which we will discuss next).
Note that this modulation system is the counterpart to the one derived earlier in Sec. 2.2, namely (2.2.16)-(2.2.21).

Remark 3.2.3. *Since the envelope of pulse solutions is given by spatially localized solutions of the modulation equations, we will have to make sure that the Nonlinear Schrödinger equation* (3.2.11) *really supports spatially localized solutions. As mentioned before, for a Nonlinear Schrödinger equation*

$$i\partial_{\underline{T}} A = \nu_1 \partial_1^2 A + \nu_2 |A|^2 A, \quad \nu_j \in \mathbb{R}$$

the condition

$$\nu_1 \nu_2 > 0, \tag{3.2.17}$$

ensures that spatially localized solutions exist (cf. Rem. 1.1.2) and it is then called "focussing". Note that in the following chapters we will always add a factor $\gamma \in \mathbb{R}$ in front of the nonlinearity of each equation. By this, ν_2 is proportional to $\gamma \in \mathbb{R}$ and, thus, we can tune the sign of γ to force (3.2.17).

Mixed and higher order harmonics

The method that we employ to cancel any mixed and higher order harmonics is exactly as described in Sec. 2.2.4. For the sake of completeness, we illustrate it here again.
To simplify the exposition we will explain the procedure for canceling these terms only

- for the case $\mathcal{N} = \{1, 2\}$,
- for a one-pulse ansatz

$$V^{an} := (v^{an} + M^1_{mixed,\alpha}, \overline{v^{an}} + \overline{M^1_{mixed,\alpha}}, M^2_{mixed,\alpha}, \overline{M^2_{mixed,\alpha}})^T, \quad v^{an} := \varepsilon BE,$$

- and for a very simple nonlinearity given by

$$\underline{N}_j[V] := [\mathcal{F}^{-1}\tilde{N}[\mathcal{F}\{V\}]]_j := s_j(a_1 v_{1,1} + a_2 v_{1,2} + a_3 v_{2,1} + a_4 v_{2,2})^3,$$

  with $s_j, a_j \in \mathbb{C}$.

Note that even though the nonlinearity is more complex in general, the procedure is still the same, even though possibly more lengthy!

Let us assume for a moment that , $M^n_{mixed,\alpha} = 0, n = 1, 2$. Then

$$\underline{N}_1[V^{an}] = \varepsilon^3 s_1(3a_1^2 a_2 |B|^2 BE + a_1^3 B^3 E^3 + \ldots).$$

We have already taken care of all terms occurring at pure harmonics, i.e. we took from N_1 the term

$$3\varepsilon^3 s_1 a_1^2 a_2 |B|^2 BE$$

and used it to construct the Nonlinear Schrödinger equation (3.2.11).
Now to cancel, e.g. the term

$$\varepsilon^3 s_1 a_1^3 B^3 E^3$$

we will modify — exactly as explained in Sec. 2.2.4 — our ansatz to

$$v^{an} + \alpha\varepsilon^3 s_1 a_1^3 B^3 E^3$$

which gives the solvability condition

$$\alpha 3i\omega_1(k_1) = \alpha i\omega_1(3k_1) + s_1 a_1^3$$

and, hence, the equation for α given by

$$\alpha = -\frac{is_1 a_1^3}{3\omega_1(k_1) - \omega_1(3k_1)},$$

which is well–defined if the "non-resonance condition"

$$3\omega_1(k_1) \neq \omega_1(3k_1)$$

is fulfilled.
By the same reasoning we can cancel all these terms up to an order $\mathcal{O}(\varepsilon^6)$ from $\underline{N}_1$ and $\underline{N}_2$.
Hereafter, all the modifications

$$\alpha\varepsilon^3 s_1 a_1^3 B^3 E^3 + \ldots$$

are denoted by

$$M^1_{mixed,\alpha}.$$

Next let us turn our attention to $\underline{N}_3$. Note that we have not derived any conditions here. So we really have to come up with an ansatz for the third component of V^{an} that takes care of all the nonlinear terms. Again, if we first assume $M^2_{mixed,\alpha} = 0$ we get, similarly as for the first component,

$$\underline{N}_3[V^{an}] = \varepsilon^3 s_3(3a_1^2 a_2|B|^2 BE + a_1^3 B^3 E^3 + \ldots).$$

Now if we set the third component of V^{an} to

$$\varepsilon^3 s_3(3\alpha_1 a_1^2 a_2|B|^2 BE + \alpha_2 a_1^3 B^3 E^3 + \ldots)$$

we get at $\mathcal{O}(\varepsilon^3)$ the solvability condition

$$\begin{aligned} &3i\omega_1(k_1)s_3\alpha_1 a_1^2 a_2|B|^2 BE + 3i\omega_1(k_1)\alpha_2 a_1^3 B^3 E^3 + \ldots \\ &\quad = 3i\omega_2(k_1)s_3\alpha_1 a_1^2 a_2|B|^2 BE + 3i\omega_1(3k_1)\alpha_2 a_1^3 B^3 E^3 + \ldots \\ &\qquad + 3s_3\alpha_1 a_1^2 a_2|B|^2 BE + 3\alpha_2 a_1^3 B^3 E^3 + \ldots \end{aligned}$$

which then yields — when sorted for different harmonics E and E^3 — the solvability conditions

$$\begin{aligned} \alpha_1 &= -\frac{is_3 a_1^2 a_2}{\omega_1(k_1) - \omega_2(k_1)}, \\ \alpha_2 &= -\frac{is_3 a_1^3}{3\omega_1(k_1) - \omega_1(3k_1)}, \end{aligned}$$

which again are well-defined if the "non-resonance conditions"

$$\omega_1(k_1) \neq \omega_2(k_1), \quad \omega_1(k_1) \neq \omega_1(3k_1),$$

are met.
From now on, we denote

$$v_{n,1}, \quad n \geq 2,$$

by

$$M^n_{mixed,\alpha}, \quad n \geq 2.$$

In summary, the notation $M^n_{mixed,\alpha}, n \geq 1$, should emphasize that these terms are really only higher order modifications which play a minor role in the analysis of pulse interaction. They merely serve to diminish the residual.

3.2.3 Separation of internal and interaction dynamics

We have chosen the notation for the nonlinear terms of the modulation system following the reasoning from Ch. 2 where we achieved the "separation of internal and interaction dynamics" by a "clever sorting". Thus, our notation can only be understood from the view point of Sec. 2.2. Without bearing in mind the mechanisms described in that section, the recipe presented here will seem quite random.

The linear part of the extended modulation system (3.2.11)-(3.2.16) will always have the same form, regardless which equation is under consideration. The structure of the nonlinear terms, however, might be very different. Now the notation introduced should indicate the following recipe for treating the nonlinear terms when dealing with a specific equation:

- Compute $[\mathcal{F}^{-1}\tilde{N}[\mathcal{F}\{V\}]]_1$ up to order $\mathcal{O}(\varepsilon^4)$.
- Name the nonlinear terms occurring at $\mathcal{O}(\varepsilon^r), r = 3, 4$, which only involve the pure harmonic E_j

  $$\tilde{N}_{r-2,j}(B_j^{(q)}, B_{3-j}^{(q)})_{q \leq r-2}.$$

  This notation should emphasize the recursive structure of the modulation system and the fact that the nonlinear terms at $\mathcal{O}(\varepsilon^r)$ feed the equations of $B_j^{(r-2)}, j = 1, 2$.
- Further separate the nonlinear terms in two groups: Name those which exclusively feature products of $B_j^{(q)}$ the *internal dynamics terms*

  $$\tilde{N}^{internal}_{r-2,j}(B_j^{(q)})_{q \leq r-2},$$

  and name the rest the *interaction dynamics terms*

  $$\tilde{N}^{inter}_{r-2,j}(B_j^{(q)}, B_{3-j}^{(q)})_{q \leq r-2}.$$

- In order to derive the formulas for $\Omega_j^{(2)}$ and $\Psi_j^{(1)}$ further sort the interaction terms at order $\mathcal{O}(\varepsilon^4)$ for those proportional to $B_j^{(1)}$ - and name them

  $$\tilde{N}^{inter,\Omega_j^{(2)}}_{2,j}(B_j^{(q)}, B_{3-j}^{(q)})_{q \leq 2},$$

  and those proportional to $\partial_1 B_j^{(1)}$ - and name them

  $$\tilde{N}^{inter,\Psi_j^{(1)}}_{2,j}(B_j^{(q)}, B_{3-j}^{(q)})_{q \leq 2}.$$

Some remarks.

1. *In the next chapters we will discuss three equations, whose nonlinearities have a very similar structure, namely the one that we have already discussed in the example of Sec. 3.2.1. Hence, we will employ the recipe for sorting terms to* (3.2.6) *with specific values for* α_n, β_n.

2. *We do not have to compute the terms in* $[\mathcal{F}^{-1}\tilde{N}[\mathcal{F}\{V\}]]_1$ *at* $\mathcal{O}(\varepsilon^5)$ *since we will only employ qualitative estimates for the equations derived from the residual terms at that order. Hence, these terms do not have to be specified explicitly. All you have to ensure is that these terms meet the requirements to employ* (2.2.4).

3. *By deriving the above solvability conditions, we can formally make the residual* (3.2.1) *of order* $\mathcal{O}(\varepsilon^6)$.

3.3 Solution strategy for nonlinear wave equations with constant coefficients

In this section we present a scheme by which pulse interaction analysis can be carried out for a nonlinear wave equation with constant coefficients as, for example, the cubic Klein-Gordon equation (which we have already discussed in Ch. 2 and will revisit in Ch. 4) or the Maxwell-Lorentz system (see Ch. 5).

The analysis essentially consists of four steps:

1. **Transformation to "natural coordinates".** This step serves to transform the equation to the form (3.1.1) which enables us to use the formal calculations from Sec. 3.2.
   To this end, we first cast the equation into a first order system

$$\partial_t U = A[U] + N[U] \tag{3.3.1}$$

   and then use Fourier representation to get a system of ODEs

$$\partial_t \hat{U} = \hat{A}\hat{U} + \hat{N}[\hat{U}] \tag{3.3.2}$$

   parametrized by the spectral parameter $k \in \mathbb{R}$. At this point, we will try to diagonalize the linear part $\hat{A} = \hat{A}(k)$ in order to arrive at a system like (3.1.1). However, this step is in general not straightforward, since either the diagonalizing transformation is unbounded — see **CONDITION 2** below why this is needed — or the system is not even diagonalizable for all values of the spectral parameter $k \in \mathbb{R}$.
   After having overcome these difficulties (a nontrivial task as the next chapters will show) we can transform (3.3.2) into a diagonalizable system

$$\partial_t \hat{U} = \hat{A}_p\hat{U} + \hat{N}_p[\hat{U}] \tag{3.3.3}$$

   whose diagonalizing transformation $\tilde{Q}$ is bounded in k.

2. **Formal derivation of modulation equations.** After having performed the diagonalizing transformation, we can follow the computations of Sec. 3.2 line-by-line, i.e. we only have to fill in the nonlinear terms for our specific equation and rigorously show that (in some space suiting the modulation equations)

$$\|\widetilde{\mathrm{Res}}\{(\tilde{V}_j^{an})_{j\in\mathbb{N}}\}\|_\chi \leq C\varepsilon^{11/2},$$

if

$$(\tilde{V}_j^{an})_{j\in\mathbb{N}} := \mathcal{F}\{(V_n^{an})_{n\in\mathcal{N}}\},$$

where $(V_n^{an})_{n\in\mathcal{N}}$ is the two-pulse ansatz given by (3.2.10) and $\widetilde{\text{Res}}$ is the residual corresponding to the "system in natural coordinates" (3.1.1), i.e. (3.2.1).

3. Approximation result. To justify the extended two-pulse ansatz

$$U^{an} := \mathcal{F}^{-1}\tilde{Q}\mathcal{F}\{(V_n^{an})_{n\in\mathcal{N}}\}$$

for (3.3.3) over the long time scale of $\mathcal{O}(1/\varepsilon^2)$ (which is the natural time scale for pulses described by the Nonlinear Schrödinger equation) we will use a proof technique that is very similar to the one presented in Ch. 1 which is inspired by [KSM92]. The subsequent illustration of the proof will simplify the exposition in the next chapters, since the proofs of approximation theorems will only amount to showing that the conditions, that arise to make this proof work, are really fulfilled!

Using the representation $U =: U^{an} + R$ we derive the error equation

$$\partial_t R = A_p[R] + G[R, U^{an}] + \text{Res}\{U^{an}\}$$

with

$$G[R, U^{an}] := N_p[U^{an} + R] - N_p[U^{an}],$$

and with the residual

$$\text{Res}\{U^{an}\} := -\partial_t U^{an} + A_p[U^{an}] + N_p[U^{an}],$$

where $A_p[\cdot] := \mathcal{F}^{-1}\{\hat{A}_p\mathcal{F}\{\cdot\}\}$ and $N_p[\cdot] := \mathcal{F}^{-1}\{\hat{N}_p[\mathcal{F}\{\cdot\}]\}$.
If

CONDITION 1. The linear part is diagonalizable, i.e. it has a representation $A_p[U] = \mathcal{F}^{-1}(\hat{A}_p\hat{U}) = \mathcal{F}^{-1}(\tilde{Q}^{-1}\hat{D}\tilde{Q}\hat{U})$ with $\hat{D}$ a diagonal matrix.

is fulfilled, we have

$$\partial_t R = A_p[R] + G[R, U^{an}] + \mathcal{F}^{-1}\tilde{Q}\{\widetilde{\text{Res}}\{\mathcal{F}V^{an}\}\}.$$

Furthermore, if

CONDITION 2. The diagonalizing matrix and its inverse are uniformly bounded, i.e. for some $C > 0$ and for some matrix norm $\|\cdot\|$ we have

$$\sup_k \|\tilde{Q}(k)\| \le C, \quad \sup_k \|\tilde{Q}(k)^{-1}\| \le C.$$

is fulfilled, we can prove that $\hat{A}_p$ gives rise to a bounded, strongly continuous semigroup $(e^{\hat{A}_p t})_{t\ge 0}$ which can be estimated by

$$\|e^{\hat{A}_p t}\hat{U}\|_{\hat{\mathcal{X}}} = \|\hat{Q}e^{\hat{D}t}\hat{Q}^{-1}\hat{U}\|_{\hat{\mathcal{X}}} \le C\|\hat{U}\|_{\hat{\mathcal{X}}}. \tag{3.3.4}$$

This enables us then to use the variation of constant formula (with initial condition $R(t)|_{t=0} = 0$) to get

$$\begin{aligned}\|R\|_{\mathcal{X}} &\leq \int_0^t \|e^{A_p(t-s)} I[R, U^{an}](s)\|_{\mathcal{X}}\, ds \\ &\leq C \int_0^t \|e^{\hat{A}_p(t-s)} \mathcal{F}\{I[R, U^{an}]\}(s)\|_{\hat{\mathcal{X}}}\, ds \\ &\leq C \int_0^t \|\mathcal{F}\{I[R, U^{an}]\}(s)\|_{\hat{\mathcal{X}}}\, ds,\end{aligned}$$

where we used the notation $I[R, U^{an}] := G[R, U^{an}] + \mathcal{F}^{-1}\tilde{Q}\{\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\}$ and refrained from specifying the function spaces $\mathcal{X}, \hat{\mathcal{X}}$, since, of course, they depend on the specific equation under consideration.
From this point on we can essentially proceed as in the justification proof of Ch. 1 to get via a Gronwall-type inequality that

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|R\|_{\mathcal{X}} = \sup_{t\in[0,T_0/\varepsilon^2]} \|U - U^{an}\|_{\mathcal{X}} \leq C\varepsilon^{7/2},$$

for some $C > 0$ and $\varepsilon \in (0, \varepsilon_0), \varepsilon_0 > 0, \varepsilon_0 = \varepsilon_0(\underline{T}_0)$ (see [KSM92]).

4. **Description of pulse interaction.** Now the remaining task is to interpret what $U^{an} := \mathcal{F}^{-1}\tilde{Q}\mathcal{F}\{(V_n^{an})_{n\in\mathcal{N}}\}$ reveals about pulse interaction. The ultimate goal is to transfer the "separation of internal and interaction dynamics" (that was realized in the second step for the "system in natural coordinates") to the original equation.

Note that while the first and the last step involve some ingenious manipulations which take advantage of the specific structure of the underlying problem, the second and third step are basically the same regardless of the considered equation. This makes the presented method widely applicable!

3.4 Chapter summary

This chapter forms the center piece of the present work! It illustrates how to employ the "separation of internal and interaction dynamics" (that we have achieved in the first part of the work for a cubic Klein-Gordon equation) for a two-pulse of any system

$$\partial_t U = A(U) + N(U)$$

that can be brought via some change of variables $U = \mathcal{G}^{-1}\{(\tilde{V}_n)_{n\in\mathcal{N}}\}$ to its "natural coordinates form"

$$\partial_t \tilde{V}_n = \tilde{D}_n \tilde{V}_n + \tilde{N}_n[(\tilde{V}_n)_{n\in\mathcal{N}}],$$

for $\tilde{V}_n := (\tilde{v}_{n,1}, \tilde{v}_{n,2})^T, \tilde{v}_{n,l} = \tilde{v}_{n,l}(k,t) \in \mathbb{C}, n \in \mathcal{N}, l = 1, 2$, $\mathcal{N}$ some at most countable index set, with linear part

$$\tilde{D}_n = \tilde{D}_n(k) := \begin{pmatrix} i\omega_n(k) & 0 \\ 0 & -i\omega_n(k) \end{pmatrix},$$

where $\omega_n : \mathbb{R} \longrightarrow \mathbb{R}$ are nonlinear, piece-wise analytic functions and the nonlinear operator $\tilde{N}_n$ which is at least cubic and whose k-dependence is also piece-wise analytic.

The term "natural coordinates" is motivated by the fact that the quantities ω_n are the branches of the dispersion relation of the general system and the initial construction of pulses was achieved through an expansion of the dispersion relation in first place (remember (1.1.2)).

This chapter emphasizes that the derivation of modulation equations can be carried out in a much more transparent way by first bringing the equation to "natural coordinates"!

The formal result for interchannel interaction can be stated as follows: A two-pulse solution U of the above general system can be approximated by $U^{an} = \mathcal{GF}\{(V_n^{an})_{n\in\mathcal{N}}\}$ where $(V_n^{an})_{n\in\mathcal{N}}$ is the extended two-pulse ansatz

$$V_1^{an} = (v^{an}, \overline{v^{an}})^T, \quad V_n^{an} := (M_{mixed,\alpha}^n, \overline{M_{mixed,\alpha}^n})^T, n \geq 2,$$

with

$$v^{an}(x,t) = \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r B_j^{(r)}(\underline{X}_j, \underline{T}) e^{i\underline{Y}_j} + M_{mixed,\alpha}^1,$$

$$\underline{X}_j := \underline{X} + \varepsilon\omega_1'(k_j)t + \varepsilon^2\Psi_j^{(1)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T}),$$

$$\underline{Y}_j := k_j x + \omega_1(k_j)t + \sum_{l=1,2} \varepsilon^l \Omega_j^{(l)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T}),$$

$$\underline{X} := \varepsilon x, \quad \underline{T} := \varepsilon^2 t,$$

if its components fulfill the extended modulation system (3.2.11)-(3.2.16), i.e. the recursively solvable modulation system for internal dynamics

$$\begin{aligned}
\partial_2 B_j^{(1)} &= -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(1)} + \tilde{N}_{1,j}^{internal}(B_j^{(1)}),\\
\partial_2 B_j^{(2)} &= -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2},\\
\partial_2 B_j^{(3)} &= -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega_1^{(4)}(k_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\leq 3},
\end{aligned}$$

and the ordinary differential equations for the interaction quantities

$$\begin{aligned}
\partial_1\Omega_j^{(1)} &= \frac{i}{D_j}\tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)}),\\
\partial_1\Psi_j^{(1)} &= \frac{1}{D_j}(\tilde{N}_{2,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(q)})_{q\leq 2} + \omega_1''(k_j)(\partial_1\Omega_j^{(1)})),\\
\partial_1\Omega_j^{(2)} &= \frac{i}{D_j}(\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(q)})_{q\leq l} + \frac{\omega_1''(k_j)}{2}(\partial_1^2\Omega_j^{(1)}) - i(\partial_2\Omega_j^{(1)})).
\end{aligned}$$

with $D_j := \omega_1'(k_j) - \omega_1'(k_{3-j})$. The terms $M_{mixed,\alpha}^n$ are explained in Sec. 3.2.2, but they actually play a minor role for our analysis.
This modulation system is the counterpart of the one derived earlier in Ch. 2, i.e. (2.2.16)-(2.2.21). Hence, the "separation internal and interaction dynamics" is accomplished for the system in "natural coordinates".

Note that all following chapters will make use of the formal derivation presented here, i.e. the extended

modulation system (3.2.11)-(3.2.16) will reappear in the next three chapters, but with explicit nonlinear terms. Furthermore, we will show that the "separation of internal and interaction dynamics" can be also carried out for the original system.

To simplify the exposition in the next chapters, we presented the solution strategy that can be employed when dealing with nonlinear wave equations with constant coefficients. We will stick to this scheme in Ch. 4 and Ch. 5 and use a modified version of it for Ch. 6. The main steps are:

1. *Transformation to "natural coordinates"*, which brings the system to its "natural coordinates representation".

2. *Formal derivation of modulation equations*, which derives the extended modulation system (3.2.11)-(3.2.16), but with explicit nonlinear terms, and, hence, realizes the "separation of internal and interaction dynamics" in "natural coordinates"

3. *Approximation result*, which justifies the approximation for the long time scale of order $\mathcal{O}(1/\varepsilon^2)$.

4. *Description of pulse interaction*, which transfers the "separation of internal and interaction dynamics" from the system in "natural coordinates" (that was realized in the second step) to the original equation.

4 Pulse interaction for a cubic Klein-Gordon equation (revisited)

In this chapter we revisit the pulse interaction problem for two pulses traveling on different carrier waves in the setting of the cubic Klein-Gordon equation

$$\partial_t^2 u = \partial_x^2 u - u + u^3,$$

where $x, t, u = u(x,t) \in \mathbb{R}$. We have already discussed this in Sec. 2.2.
We reproduce the results from Ch. 2, but this time by means of the new method developed in Ch. 3. In this sense, this chapter links Part I (derivation in x-space) and Part II (derivation in "natural coordinates") of this work.
By reproducing the findings of Sec. 2.2 with the new approach we show that the derivation of modulation equations in x-space is equivalent in "natural coordinates" (see Sec. 4.4). However, the latter is more suited for equations with a more complicated dispersion relation as, for example, encountered in the case of nonlinear wave equations with periodic coefficients.

Note, that the solution strategy follows the one described in Sec. 3.3:

1. Transformation to "natural coordinates".
2. Formal derivation of modulation equations.
3. Approximation result.
4. Description of pulse interaction.

The chapter is organized according to this strategy. The last section is dedicated to "bridging" the x-space method and the one in "natural coordinates" through comparing the modulation systems (4.2.4)-(4.2.9) and (2.2.16)-(2.2.21) and showing that they are indeed equivalent.

4.1 Transformation to "natural coordinates"

Using the Fourier representation $u = \mathcal{F}^{-1}\{u\}$ yields

$$\partial_t^2 \hat{u} = -\omega^2 \hat{u} + \hat{u}^{*3}, \quad \omega = \omega(k) := \sqrt{k^2+1},$$

which can be converted into a first order system

$$\partial_t \hat{U} = \hat{A}\hat{U} + \begin{pmatrix} 0 \\ \hat{u}^{*3} \end{pmatrix}$$

with the definitions $\hat{U} := (\hat{u}, \partial_t \hat{u})^T$ and

$$\hat{A} = \hat{A}(k) := \begin{pmatrix} 0 & 1 \\ -\omega^2(k) & 0 \end{pmatrix}.$$

At this point one could simply introduce the transformation $\hat{Q}$ that diagonalizes $\hat{A}$. However, it turns out that by prefixing an additional transformation (which leaves the subspace of interest $\{\hat{U} \,|\, [\hat{U}]_2 = 0\}$ invariant!) we can get a unitary (even k-independent) diagonalizing transformation, which will significantly facilitate the analysis.
Note, that this idea has already been used before, for instance, in [BSTU06] by introducing the scaled variables $\hat{U}(k) := (\hat{u}(k), \frac{1}{\omega(k)}\partial_t \hat{u}(k))^T, k \in \mathbb{R}$. However, we want to give a deeper insight into this special choice by formalizing it as a supplementary transformation, since for more complicated settings this scaling might not be that obvious at first sight.

4.1.1 Polar decomposition

For every regular matrix $\hat{A}$ one can compute its unique polar decomposition $\hat{A} = PW$ with unitary matrix W and a positive definite Hermitian matrix P, such that the system can be written as

$$\partial_t \hat{U} = PW\hat{U} + \hat{N}[\hat{U}]$$

with

$$P := \begin{pmatrix} 1 & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad W := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Performing the change of variables $\tilde{U} := (\sqrt{P})^{-1}\hat{U}$ gives the system

$$\partial_t \tilde{U} = \hat{A}_p \tilde{U} + \hat{N}_p[\tilde{U}]$$

with

$$\hat{A}_p := \sqrt{P}W\sqrt{P} = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$$

and

$$\hat{N}_p[\tilde{U}] := (\sqrt{P})^{-1}\hat{N}[\sqrt{P}\hat{U}].$$

Note, that $\sqrt{P}W\sqrt{P}$ is a normal matrix and, hence, can be unitarily diagonalized! Furthermore, $\sqrt{P}$ leaves the subspace $\{\hat{U} \,|\, [\hat{U}]_2 = 0\}$ invariant.

By this preliminary transformation we have "prepared" the system for diagonalization, since the diagonalizing transformation has now better properties.

4.1.2 Diagonalizing transformation

Next we perform a change of coordinates $\tilde{V} := Q^*\tilde{U}$ where $\hat{A}_p = Q\tilde{D}Q^*$ with

$$Q := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix},$$

and arrive at a diagonalized system

$$\partial_t \tilde{V} = \tilde{D}\tilde{V} + \tilde{N}[\tilde{V}] \tag{4.1.1}$$

where

$$\tilde{D} := \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix}$$

and

$$\tilde{N}[\tilde{V}] := \frac{i}{4\omega}(\tilde{v}_1 + \tilde{v}_2)^{*3} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

with $\tilde{V} = (\tilde{v}_1, \tilde{v}_2)^T$.
With this, the first step of Sec. 3.3 is completed, since we arrived at a "system in natural coordinates" as (3.1.1) with $\mathcal{N} = \{1\}$.

4.2 Formal derivation of modulation equations

This section serves to diminish the residual

$$\widetilde{\mathrm{Res}}\{\tilde{V}\} := -\partial_t \tilde{V} + \tilde{D}\tilde{V} + \tilde{N}[\tilde{V}] \tag{4.2.1}$$

which corresponds to (4.1.1).
As explained in Sec. 3.2, we will have to use the spectral concentration property and inverse Fourier transformation in order to be able to perform pulse interaction analysis as in Ch. 2:
Taking a closer look at the example in Sec. 3.2.1, we see that our residual (4.2.1) is exactly (3.2.3) with $\alpha_{1/2}(k) := \pm\frac{i}{\omega(k)}$ and $\beta_{1/2} = 1$. Consequently, we will pursue the same strategy as for (3.2.3) and define the operators $\underline{D}, \underline{N}$ to obtain the x-space residual $\underline{\mathrm{Res}}$, for which we formally have (3.2.8), i.e.

$$\widetilde{\mathrm{Res}}\{(\mathcal{F}\{V_n^{an}\})_{n\in\mathcal{N}}\} = \mathcal{F}\underline{\mathrm{Res}}\{(V_n)_{n\in\mathcal{N}}\} + \mathcal{O}(\varepsilon^5). \tag{4.2.2}$$

To put this in words, due to the concentration in Fourier space we can expand the k-dependent coefficients of linear and nonlinear part, which is justified since ω and $\frac{1}{\omega}$ are analytic in k.
In summary, if we make an ansatz V^{an} such that

$$\|\underline{\mathrm{Res}}\{V^{an}\}\|_{H^s} \leq C\varepsilon^{11/2},$$

we immediately get

$$\|\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\|_{L^2(s)} \leq C\varepsilon^{11/2}.$$

The extended modulation system. Consider the extended two-pulse ansatz

$$V^{an} = (v^{an}, \overline{v^{an}})^T, \tag{4.2.3}$$

with

$$v^{an}(x,t) = \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r B_j^{(r)}(\underline{X}_j, \underline{T})e^{i\underline{Y}_j} + M_{mixed,\alpha},$$

$$\underline{X}_j := \underline{X} + \varepsilon\omega'(k_j)t + \varepsilon^2\Psi_j^{(1)}(\underline{X} + \varepsilon\omega'(k_{3-j})t, \underline{T}),$$

$$\underline{Y}_j := k_j x + \omega(k_j)t + \sum_{l=1,2}\varepsilon^l\Omega_j^{(l)}(\underline{X} + \varepsilon\omega'(k_{3-j})t, \underline{T}),$$

$$\underline{X} := \varepsilon x, \quad \underline{T} := \varepsilon^2 t,$$

where the mixed and higher order harmonic terms $M_{mixed,\alpha} = \mathcal{O}(\varepsilon^3)$ are according to Sec. 3.2.2. Plugging this ansatz into $\underline{\mathrm{Res}}\{V^{an}\}$ defined by (3.2.7) with $\alpha_{1/2}(k) := \pm\frac{i}{\omega(k)}$ and $\beta_{1/2} = 1$ gives the following extended modulation system:

The recursivley solvable system of *internal dynamics equations*

$$\partial_2 B_j^{(1)} = -i\frac{\omega''(k_j)}{2}\partial_1^2 B_j^{(1)} + \tilde{N}_{1,j}^{internal}(B_j^{(1)}), \tag{4.2.4}$$

$$\partial_2 B_j^{(2)} = -i\frac{\omega''(k_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega'''(k_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2}, \tag{4.2.5}$$

$$\partial_2 B_j^{(3)} = -i\frac{\omega''(k_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega'''(k_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega^{(4)}(k_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\leq 3}, \tag{4.2.6}$$

with the nonlinear terms

$$\begin{aligned}
\tilde{N}_{1,j}^{internal}(B_j^{(1)}) :=& \ \frac{3i}{4\omega(k_j)}|B_j^{(1)}|^2 B_j^{(1)},\\
\tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2} :=& \ \frac{3i}{4\omega(k_j)}(\overline{B_j^{(2)}}(B_j^{(1)})^2 + \frac{3i}{2\omega(k_j)}\overline{B_j^{(1)}}B_j^{(1)}B_j^{(2)}\\
&- \frac{3ik_j}{4\omega(k_j)^3}\partial_1(|B_j^{(1)}|^2 B_j^{(1)}),
\end{aligned}$$

the *interaction dynamics equations*

$$\partial_1\Omega_j^{(1)} = \frac{i}{D_j}\tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)}), \tag{4.2.7}$$

$$\partial_1\Psi_j^{(1)} = \frac{1}{D_j}(\tilde{N}_{2,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(q)})_{q\leq 2} + \omega''(k_j)(\partial_1\Omega_j^{(1)})), \tag{4.2.8}$$

$$\partial_1\Omega_j^{(2)} = \frac{i}{D_j}(\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(q)})_{q\leq l} + \frac{\omega''(k_j)}{2}(\partial_1^2\Omega_j^{(1)}) - i(\partial_2\Omega_j^{(1)})), \tag{4.2.9}$$

with $D_j := \omega'(k_j) - \omega'(k_{3-j})$, and the nonlinear terms

$$\begin{aligned}
\tilde{N}_{1,j}^{inter,\Omega_1^{(1)}}(B_{3-j}^{(1)})B_j^{(l)} :=& \ \frac{9i}{2\omega(k_j)}|B_{3-j}^{(1)}|^2 B_j^{(l)},\\
\tilde{N}_{2,j}^{inter,\Omega_1^{(2)}}(B_{3-j}^{(1)})_{q\leq 2}B_j^{(1)} :=& \ \frac{3i}{2\omega(k_j)}(\overline{B_{3-j}^{(1)}}B_j^{(1)}B_{3-j}^{(2)} + \overline{B_{3-j}^{(2)}}B_j^{(1)}B_{3-j}^{(1)})\\
&- \frac{3ik_j}{2\omega(k_j)^3}(\partial_1|B_{3-j}^{(1)}|^2)B_j^{(1)},
\end{aligned}$$

and the algebraic system of equations as described in Sec. 3.2.2 which cancels the mixed and higher order harmonic terms in the residual.

Lemma 4.2.1. *Consider the extended two-pulse ansatz* (4.2.3) *where*

$$B_j^{(r)} \in C\left([0,T_0]; H^{s_B-3(r-1)+l} \cap H^{s_B-3(r-1)}(l)\right), \quad s_B \geq 10, l \geq 2,$$

fulfill the extended modulation system (4.2.4)-(4.2.9)*. Then for all $T_0 > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\|_{(L^2(s))^2} \leq C\varepsilon^{11/2},$$

if $1 \leq s \leq s_B - 10$.

Proof. As already pointed out, it suffices to show the bound for $\underline{\text{Res}}\{V^{an}\}$, since the rest readily follows from (4.2.2) and Rem. 3.2.1. This puts us in the situation where we can use exactly the reasoning from the proof of Lemma 1.1.1:
By simply computing the residual for the extended ansatz and using the modulation system all terms up to order $\mathcal{O}(\varepsilon^5)$ cancel. The loss of regularity from $B_j^{(1)}$ to $B_j^{(3)}$ is due to recursive structure of the equations and the derivatives appearing on the right-hand sides. Note, that this loss is much more transparent for the modulation system (4.2.4)-(4.2.9) derived in "natural coordinates", since we do not have to dissolve temporal derivatives into spatial ones, as we have done that in the proof of Thm.1.1.1. For the remaining terms at order $\mathcal{O}(\varepsilon^6)$ like

$$\frac{\omega'''(k_j)}{6}\partial_1^3 B_j^{(3)}, \frac{\omega^{(4)}(k_j)}{4!}\partial_1^4 B_j^{(2)}, \dots$$

the $\mathcal{O}(1)$ boundedness in the H^s norm is guaranteed by choosing s high enough. Since the term causing most loss of regularity is

$$\frac{\omega^{(4)}(k_j)}{4!}\partial_1^4 B_j^{(3)}$$

we will have to choose $1 \leq s \leq s_B - 6 - 4 = s_B - 10$.
For the coupling terms we employ the estimate (2.2.4) from Ch. 2 to finally get

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\underline{\text{Res}}\{V^{an}\}\|_{(H^s)^2} \leq C\varepsilon^{11/2},$$

which concludes the proof. □

By this, we completed the second step of the solution strategy described in Sec. 3.3!

4.3 Approximation result

We are now ready to formulate the approximation theorem.

Theorem 4.3.1. *Given the conditions of Lemma 4.2.1 for all $T_0 > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|U - (\mathcal{F}^{-1}Q\mathcal{F})\{V^{an}\}\|_{C_b^{s-1}} \leq C\varepsilon^{7/2},$$

where $u = [U]_1$ is an exact two-pulse solution of the cubic Klein-Gordon equation (1.0.1).

Proof. Since both **CONDITION 1** and **CONDITION 2** are fulfilled, the proof follows exactly the third step of Sec. 3.3. Therefore, the only remaining task is to specify the function spaces $\chi, \hat{\chi}$:
We have

$$\|e^{\hat{A}_p(k)t}\hat{U}\|_{(L^2(s))^2} \leq \|\hat{U}\|_{(L^2(s))^2}$$

since

$$e^{\hat{A}_p(k)t} = Qe^{\hat{D}_p(k)t}Q^*$$

and Q, Q^* are isometries and so is $e^{\hat{D}_p(k)t}$ for every $t, k \in R$. Hence, by choosing

$$\chi := (H^s)^2, \quad \hat{\chi} := (L^2(s))^2,$$

we can readily apply the scheme from the third step of Sec. 3.3. □

Remark 4.3.2. *Note, that* $u = [U^{an}]_1 = [\mathcal{F}Q\mathcal{F}^{-1}\{V^{an}\}]_1$ *is - up to scaling constants - exactly our ansatz* (2.2.8)*, and hence we have really reproduced the result from Ch. 2 (cf. Sec. 4.4), in particular, the interpretation of the quantities* $\Omega_j^{(r)}, \Psi_j^{(1)}$ *is the same as in Sec. 2.2 where we gave a detailed explanation of their role.*

4.4 The bridge between x-space and spectral methods

To see that the extended modulation system (2.2.16)-(2.2.21) from Ch. 2 and the one derived here, i.e. (4.2.4)-(4.2.9), are indeed the same, we only need to

- compute the derivatives of $\omega = \omega(k)$ explicitly to see that

$$\omega''(k) = \frac{1 - (\omega'(k))^2}{\omega(k)}, \dots,$$

  which will make clear the relation between the coefficients of the two modulation systems,

- use at order $\mathcal{O}(\varepsilon^4)$ the identity

$$\partial_{\underline{XT}} A_j^{(1)} = \partial_1(i\nu_{j,1}\partial_1^2 A_j^{(1)} + i\nu_{j,2}|A_j^{(1)}|^2 A_j^{(1)}) = i\nu_{j,1}\partial_1^3 A_j^{(1)} + i\nu_{j,2}\partial_1(|A_j^{(1)}|^2 A_j^{(1)}),$$

  if $A_j^{(1)}$ solves the NLS equation

$$\partial_{\underline{T}} A_j^{(1)} = i\nu_{j,1}\partial_1^2 A_j^{(1)} + i\nu_{j,2}|A_j^{(1)}|^2 A_j^{(1)},$$

  and similarly for all the other terms as $\partial_{\underline{T}}^2 A_j^{(1)}, \partial_{\underline{XT}} A_j^{(2)}, \dots$.

Comparing the modulation systems again emphasizes the elegance of the derivation in "natural coordinates", since the equations have a much clearer structure!

4.5 Chapter summary

This chapter links the first and the second part of this work: We reproduced the findings of Ch. 2 by means of the new procedure illustrated in Ch. 3. This shows that the derivation of modulation equations in x-space as given in Part I of the present work and the derivation in "natural coordinates" as employed in Part II are indeed equivalent.
Moreover, since this chapter shows the results from chapter Ch. 2 from the point of view of Ch. 3, this readily paves the way of transferring them to equations exhibiting multibranch dispersion relations.

5 Pulse interaction for a Maxwell-Lorentz system

In this chapter we describe the interaction of pulses with different carrier waves in the context of a Maxwell-Lorentz system

$$\partial_t^2 u = \partial_x^2 u + \partial_t^2 p \tag{5.0.1}$$
$$\partial_t^2 p = -\omega_0^2 p - d^2 u + \gamma p^3 \tag{5.0.2}$$

with $x, t, u = u(x,t), p = p(x,t) \in \mathbb{R}$ and $\omega_0, d, \gamma \in \mathbb{R}$ some constants. Its dispersion relation (see middle panel of Fig.3.1) reads

$$k^2 = \omega^2 + d^2 \frac{\omega^2}{\omega_0^2 - \omega^2}. \tag{5.0.3}$$

In the following, we consider w.l.o.g. the case $d = \omega_0 = 1$, since the qualitative features of the dispersion relation are essentially the same for all $d, \omega_0 \in \mathbb{R}$. Note, that we do not set γ to a specific value, since it might have to be adjusted such that the Nonlinear Schrödinger equation, that we will derive for this system, is focussing (see Rem.3.2.3).

The study of pulse interaction for this system is mathematically very challenging, since

- the corresponding dispersion relation features a double zero eigenvalue and a corresponding Jordan block (which complicates the diagonalization to the "natural coordinates", see Sec. 5.1.1),
- and the interaction between branches has to be treated (which we have already explained in Sec.(3.2.2)).

Furthermore, this system is also interesting from the view point of applications, since it is a widely accepted model for light propagation in dispersive, nonlinear media (see [SB07],[JMR96], [Blo65]).

This chapter is one of the highlights of our work, since the Maxwell-Lorentz system indeed features most of the difficulties that can be encountered when dealing with pulse interaction analysis for a nonlinear wave equation. Hence, it is the perfect prototype equation for illustrating the success of our new approach!

As for the cubic Klein-Gordon case in the last chapter, we will stick to the scheme presented in Sec.3.3:

1. Transformation to "natural coordinates".
2. Formal derivation of modulation equations.
3. Approximation result.
4. Description of pulse interaction.

This chapter is organized along this strategy.

5.1 Transformation to "natural coordinates"

Casting the above equations into a first order system for $U = (u, \partial_t u, p, \partial_t p)^T$ and making use of Fourier representation we get a system for the corresponding Fourier coefficients

$$\partial_t \hat{U} = \hat{A}\hat{U} + \hat{N}[\hat{U}]$$

where

$$\hat{A} = \hat{A}(k) := \begin{pmatrix} 0 & 1 & 0 & 0 \\ -(k^2+1) & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{pmatrix}, \quad \hat{N}[\hat{U}] := \gamma[\hat{U}]_3^{*3} y, \quad y := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Let us take a closer look at the features of $\hat{A}$.

5.1.1 The system for small k - The problem with zero eigenvalue Jordon blocks

For $k \neq 0$ the matrix $\hat{A}(k)$ is diagonalizable, i.e.

$$\hat{A}(k) = \hat{Q}(k)\hat{D}(k)\hat{Q}(k)^{-1}.$$

The diagonal matrix is given by

$$\hat{D}(k) = diag[i\omega_1(k), -i\omega_1(k), i\omega_2(k), -i\omega_2(k)],$$

with

$$\omega_1(k) := \frac{1}{\sqrt{2}}\sqrt{k^2 + 2 + \sqrt{k^4+4}},$$
$$\omega_2(k) := \frac{1}{\sqrt{2}}\sqrt{k^2 + 2 - \sqrt{k^4+4}},$$

and the diagonalizing transformation reads

$$\hat{Q}(k) = \left(\; f_1(k) \;\middle|\; \overline{f_1(k)} \;\middle|\; f_2(k) \;\middle|\; \overline{f_2(k)} \; \right),$$

where

$$f_{1/2}(k) = \begin{pmatrix} \frac{1}{4}(\sqrt{k^4+4} \pm k^2) \\ \frac{i}{4\omega_{1/2}(k)}(\sqrt{k^4+4}(k^2+1) \pm (k^4+k^2+2)) \\ \pm\frac{1}{2} \\ \frac{i}{4\omega_{1/2}(k)}(\sqrt{k^4+4} \pm (k^2+2)) \end{pmatrix},$$

are the normalized eigenvectors corresponding to $i\omega_{1/2}(k)$.

However, in the limit $|k| \to 0$ the eigenvectors $f_2(k)$ and $\overline{f_2(k)}$ approach each other, thus, the transformation $\hat{Q}$ becomes singular in this limit. Note, that this is in conflict with **CONDITION 2** from Sec. 3.3.

For $k = 0$ the double zero eigenvalue results in a Jordan block, i.e.

$$\hat{A}(0) = \hat{Q}_0 \hat{J}_0 \hat{Q}_0^{-1}$$

with

$$\hat{J}_0 = \begin{pmatrix} i\omega(0) & 0 & 0 & 0 \\ 0 & -i\omega(0) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\hat{Q}_0 = \left(\begin{array}{c|c|c|c} \tilde{f}_1 & \overline{\tilde{f}_1} & \tilde{f}_2 & \tilde{f}_3 \end{array} \right)$$

with

$$\tilde{f}_1 = \begin{pmatrix} 1 \\ i\sqrt{2} \\ 1 \\ i\sqrt{2} \end{pmatrix}, \quad \tilde{f}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \tilde{f}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix},$$

where $\tilde{f}_1$ is the eigenvector corresponding to $\omega_1(0) = i\sqrt{2}$, $\tilde{f}_2$ is the eigenvector corresponding to $\omega_2(0) = 0$ and $\tilde{f}_3$ is the generalized eigenvector corresponding to the double zero eigenvalue. Evidently, at $k = 0$ **CONDITION 1** from Sec. 3.3 is violated. Furthermore, a zero eigenvalue Jordan block causes undesired growth:

Dynamics of zero eigenvalue Jordan blocks. The dynamics expected from a zero eigenvalue Jordan block can best be demonstrated by the first order system of ordinary differential equations reading

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

which is nothing but the second order problem

$$\partial_t^2 u_1 = 0.$$

This is, however, solved by

$$u_1(t) = u_1'(0)t + u_1(0) = u_2(0)t + u_1(0).$$

Since the solutions, that we are interested in, are bounded in time, one has to set $u_2(0) = u_1'(0) = 0$ at $k = 0$ to avoid growth coming from this Jordan block.

Restriction to a subspace at $k = 0$. To resolve the just explained difficulties, we use the following observation:
The change of variables $\hat{U} = \hat{Q}_0 \tilde{U}$ at $k = 0$ results in the system

$$\partial_t \tilde{U} = \begin{pmatrix} i\omega(0) & 0 & 0 & 0 \\ 0 & -i\omega(0) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tilde{U} - 2\gamma [\hat{Q}_0 \hat{U}]_3^{*3}|_{k=0} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Hence, we have at $k = 0$ the invariant subspace

$$\mathcal{M} := \{\tilde{U} \,|\, [\tilde{U}]_4 = [\hat{Q}_0^{-1} \hat{U}]_4 = 0\},$$

on which the linear part of the system is diagonalized.
The restriction to this subspace can be formalized by introducing $\hat{U} = \tilde{Q}_0\tilde{U}$ with

$$\tilde{Q}_0 = \left(\begin{array}{c|c|c|c} f_1(0) & \overline{f_1(0)} & f_2(0) & 0 \end{array} \right),$$

that is $\hat{Q}_0$ with the last column (the generalized eigenvector) set to zero, since

$$(\hat{Q}_0\tilde{U})|_{\mathcal{M}} = \hat{Q}_0 P_{\mathcal{M}}\tilde{U} = \tilde{Q}_0\tilde{U}$$

with the projection

$$P_{\mathcal{M}} := diag[1,1,1,0].$$

Now, in order to smoothly approach this subspace for $|k| \to 0$, we employ for small $k \neq 0$ the change of variables $\hat{U} = (\hat{Q}_0 T_0)\tilde{U}$ with

$$T_0(k) := diag[1,1,1,k],$$

such that we get for small $k \neq 0$ the system

$$\partial_t \tilde{U}(k) = \begin{pmatrix} i\omega(0) + iC_1k^2 & -iC_1k^2 & -\frac{1}{2}k^2 & 0 \\ iC_1k^2 & -i\omega(0) - iC_1k^2 & -\frac{1}{2}k^2 & 0 \\ 0 & 0 & 0 & k \\ iC_1k & -iC_1k & -\frac{1}{2}k & 0 \end{pmatrix} \tilde{U}(k) - 2\gamma[\hat{Q}_0 T_0 \hat{U}]_3^{*3}(k) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Clearly, this system is now diagonalizable and, since

$$\begin{pmatrix} i\omega(0) + iC_1k^2 & -iC_1k^2 & -\frac{1}{2}k^2 & 0 \\ iC_1k^2 & -i\omega(0) - iC_1k^2 & -\frac{1}{2}k^2 & 0 \\ 0 & 0 & 0 & k \\ iC_1k & -iC_1k & -\frac{1}{2}k & 0 \end{pmatrix} = \begin{pmatrix} i\omega(0) & 0 & 0 & 0 \\ 0 & -i\omega(0) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(k),$$

we have that the diagonalizing transformation $\tilde{Q}$ for this system has the structure

$$\tilde{Q}(k) = Id + \mathcal{O}(k),$$

where $Id := diag[1,1,1,1]$ is the identity matrix. In particular, we have that $\tilde{Q}(k)$ is bounded for $|k| \to 0$, so **CONDITION 2** is met.

Note, that $\hat{Q}_0 T_0(k)$ smoothly approaches $\hat{Q}_0 P_{\mathcal{M}}$ as $|k| \to 0$, and that, even though $T_0(k)^{-1}$ becomes singular in this limit, this does not create any problems, since this inverse only appears as

$$(\hat{Q}_0 T_0(k))^{-1} y = T_0(k)^{-1}\hat{Q}_0^{-1} y = T_0(k)^{-1} \begin{pmatrix} -2 \\ -2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 0 \\ 0 \end{pmatrix},$$

which is independent of k.

5.1.2 The system for $|k| \to \infty$

For $|k| \to \infty$ we have the problem that $\|f_1(k)\|_\infty \to \infty$ and $\|f_2(k)\|_\infty \to 0$, which again is in conflict with **CONDITION 2**. This can easily be fixed by introducing for large k the change of variables $\hat{U} = T_\infty \tilde{U}$ with

$$T_\infty(k) := diag[1, \sqrt{k^2+1}, 1, 1],$$

which results in the "balanced" system matrix

$$\hat{A}_{T_\infty}(k) := T_\infty(k)^{-1}\hat{A}_{T_\infty}T_\infty(k) := \begin{pmatrix} 0 & \sqrt{k^2+1} & 0 & 0 \\ -\sqrt{k^2+1} & 0 & -\frac{1}{\sqrt{k^2+1}} & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{pmatrix}$$

which converges for $|k| \to \infty$ to

$$\begin{pmatrix} 0 & k & 0 & 0 \\ -k & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{pmatrix}.$$

Since we have for the diagonalizing transformation $\tilde{Q}$ for $\hat{A}_{T_\infty}$ that in the limit $|k| \to \infty$

$$\tilde{Q}(k) \longrightarrow \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ -\frac{i}{2} & \frac{i}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{i}{2} & -\frac{i}{2} \end{pmatrix},$$

we arrived at a system whose linear part can be diagonalized by a bounded transformation also for large values of k and, so, **CONDITION 2** is fulfilled.

5.1.3 The diagonalizable system

Merging the results from the last sections, we can construct a piece-wise smooth, global transformation

$$T(k) = \begin{cases} \hat{Q}_0 & , k = 0 \\ \hat{Q}_0 T_0(k) & , k_0 < |k| \leq k_0 + 1 \\ (k_0 + 1 - k)\hat{Q}_0 + (k - k_0)Id & , k = 0 \\ Id & , 2 < |k| \leq k_\infty \\ (k_\infty + 1 - k)Id + (k - k_\infty)T_\infty(k) & , k_\infty < |k| \leq k_\infty + 1 \\ T_\infty(k) & , k_\infty + 1 < |k| \end{cases}$$

with fixed, but arbitrary $k_0 > 0, k_\infty \gg 0$, such that the change of variables $\tilde{U} = T^{-1}\hat{U}$ results in the system

$$\partial_t \tilde{U} = \hat{A}_p \tilde{U} + T^{-1}\hat{N}_p[TU] \tag{5.1.1}$$

whose linear part can be diagonalized by a transformation $\tilde{Q}$ that meets **CONDITION 2**. In fact, through the above construction we have $\tilde{Q}(0) = Id$ and $\tilde{Q}(k) = \hat{Q}(k)$ for all $k_0 < |k| \leq k_\infty$. Therefore, we will refrain from specifying $\tilde{Q}(k)$. For our pulse interaction analysis we will simply assume w.l.o.g. that the wave numbers of the interacting pulses are $k_1, k_2 \in (k_0, k_\infty]$. To put it in a different way, we can simply choose k_1, k_2 and then adjust k_0, k_∞ since they can be chosen are arbitrarily.
Note, that we will have to translate the restriction to the subspace $\mathcal{M}$ to x-space and include it as condition for our approximation result. This subspace is characterized in x-space by

$$\int_{\mathbb{R}} (\partial_t u)\, dx = \int_{\mathbb{R}} (\partial_t p)\, dx.$$

5.1.4 Diagonalization

With these preparations we can now diagonalize the system by setting $\tilde{V} := \tilde{Q}^{-1}\hat{U}, \tilde{V} =: (\tilde{V}_1, \tilde{V}_2)^T$, such that

$$\partial_t \tilde{V}_n = \tilde{D}_n \tilde{V}_n + \tilde{N}_n[\tilde{V}], \quad n = 1, 2, \tag{5.1.2}$$

with linear part

$$\tilde{D} = \tilde{D}_n(k) := \begin{pmatrix} i\omega_n(k) & 0 \\ 0 & -i\omega_n(k) \end{pmatrix}$$

and with

$$\tilde{N}_n[\tilde{V}] := [\gamma \tilde{Q}^{-1}[\tilde{Q}\tilde{V}]_3^{*3} y]_n.$$

Evidently, this is exactly the setting (3.1.1) with $\mathcal{N} = \{1, 2\}$, i.e. we have brought the system to its representation in "natural coordinates".

5.2 Formal derivation of modulation equations

This section serves to diminish the residual

$$\begin{aligned} \widetilde{\mathrm{Res}}\{\tilde{V}\} &:= (\widetilde{\mathrm{Res}}_n\{\tilde{V}\})_{n=1,2}, \quad \tilde{V} := (\tilde{V}_1, \tilde{V}_2)^T, \\ \widetilde{\mathrm{Res}}_n\{\tilde{V}\} &:= -\partial_t \tilde{V}_n + \tilde{D}_n \tilde{V}_n + \tilde{N}_n[\tilde{V}], \quad n = 1, 2, \end{aligned} \tag{5.2.1}$$

which corresponds to (5.1.2).
Following the strategy from Sec. 3.2.1, we will again introduce an auxiliary residual $\underline{\mathrm{Res}}$ that involves only x-space expressions and is, hence, amenable to the extended modulation ansatz (3.2.10) (which we cannot Fourier transform due to the nested functions).

As already explained in great detail in Ch. 3, we will only employ our two-pulse ansatz for the two-dimensional subsystem which corresponds to the pair of branches, to which the temporal frequencies of the pulses belong (see Fig. 3.2 for an illustration and Sec. 3.1 and Sec. 3.2.2 for an explaination). In the following we will assume w.l.o.g. that the pulses belong to the pair of branches given by $\omega_1, -\omega_1$ and, hence, the main part of the residual is really $\widetilde{\mathrm{Res}}_1\{\tilde{V}\}$. Taking a closer look at it, we see that it is exactly (3.2.3) with $\alpha_n(k) := \tilde{q}_{n1}^{-1}(k) + \tilde{q}_{n3}^{-1}(k), n = 1, 2$ and $\beta_n(k) := \tilde{q}_{3n}(k), n = 1, 2$ with the notation $\tilde{Q}(k) =: ((\tilde{q}_{lm}(k)))_{lm}$ and $\tilde{Q}(k)^{-1} =: ((\tilde{q}_{lm}^{-1}(k)))_{lm}$. Since $\tilde{Q}$ is piece-wise analytic, we can define exactly as in Sec. 3.2.1 the x-space residual

$$\underline{\mathrm{Res}}_1\{V\} := \underline{D}_1[V] + \underline{N}_1[V], \tag{5.2.2}$$

For the other part of the residual, namely $\widetilde{\mathrm{Res}}_2\{\tilde{V}\}$ the procedure is essentially the same, the expansions being around $3k_1, 2k_1 + k_2, \ldots$ since the ansatz for this part will only involve mixed and higher order harmonics (cf. Sec. 3.2.2). Denoting these expanded, truncated and inverse Fourier transformed operators by $\underline{D}_2$ and $\underline{N}_2$ we can analogously define

$$\underline{\mathrm{Res}}_2\{V\} := \underline{D}_2[V] + \underline{N}_2[V],$$

and, putting the two parts together,

$$\underline{\mathrm{Res}}\{V\} := (\underline{\mathrm{Res}}_n\{V\})_{n=1,2}. \tag{5.2.3}$$

In summary, in the following we will construct an ansatz V^{an} such that

$$\|\underline{\mathrm{Res}}\{V^{an}\}\|_{H^s} \leq C\varepsilon^{11/2},$$

which will immediately give us the desired bound

$$\|\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\|_{L^2(s)} \leq C\varepsilon^{11/2}.$$

The extended modulation system. Consider the extended two-pulse ansatz

$$V^{an} = (v^{an}, \overline{v^{an}}, M^2_{mixed,\alpha}, \overline{M^2_{mixed,\alpha}})^T, \tag{5.2.4}$$

with

$$\begin{aligned}
v^{an}(x,t) &= \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r B_j^{(r)}(\underline{X}_j, \underline{T}) e^{i\underline{Y}_j} + M^1_{mixed,\alpha},\\
\underline{X}_j &:= \underline{X} + \varepsilon\omega_1'(k_j)t + \varepsilon^2\Psi_j^{(1)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T}),\\
\underline{Y}_j &:= k_j x + \omega_1(k_j)t + \sum_{l=1,2}\varepsilon^l\Omega_j^{(l)}(\underline{X} + \varepsilon\omega_1'(k_{3-j})t, \underline{T})\\
\underline{X} &:= \varepsilon x, \quad \underline{T} := \varepsilon^2 t,
\end{aligned}$$

where the mixed and higher order harmonic terms $M^n_{mixed,\alpha} = \mathcal{O}(\varepsilon^3), n = 1, 2$, are according to Sec. 3.2.2. Plugging this ansatz into $\underline{\mathrm{Res}}\{V^{an}\}$ as defined in (5.2.3) gives the following extended modulation system:
The recursivley solvable system of *internal dynamics equations*

$$\partial_2 B_j^{(1)} = -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(1)} + \tilde{N}_{1,j}^{internal}(B_j^{(1)}), \tag{5.2.5}$$

$$\partial_2 B_j^{(2)} = -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2}, \tag{5.2.6}$$

$$\partial_2 B_j^{(3)} = -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega_1^{(4)}(k_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\leq 3}, \tag{5.2.7}$$

with the nonlinear terms

$$\begin{aligned}
\tilde{N}_{1,j}^{internal}(B_j^{(1)}) :=\ & 3s_j q_1^{[j]}(q_{1-}^{[j]})^2 |B_j^{(1)}|^2 B_j^{(1)},\\
\tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2} :=\ & 3s_j q_{1-}^{[j]}(q_1^{[j]})^2[\overline{B_j^{(2)}}(B_j^{(1)})^2 + 2\overline{B_j^{(1)}}B_j^{(1)}B_j^{(2)}]\\
&+ 3s_j' q_{1-}^{[j]}(q_1^{[j]})^2 - i\partial_1(|B_j^{(1)}|^2 B_j^{(1)})\\
&+ 6s_j q_{1-}^{[j]} q_1^{[j]} q_1'^{[j]} |B_j^{(1)}|^2(-i\partial_1 B_j^{(1)})\\
&+ 3s_j q_{1-}'^{[j]}(q_1^{[j]})^2\overline{(-i\partial_1 B_j^{(1)})}(B_j^{(1)})^2,
\end{aligned}$$

the *interaction dynamics equations*

$$\partial_1\Omega_j^{(1)} = \frac{i}{D_j}\tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)}), \tag{5.2.8}$$

$$\partial_1\Psi_j^{(1)} = \frac{1}{D_j}(\tilde{N}_{2,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(q)})_{q\leq 2} + \omega_1''(k_j)(\partial_1\Omega_j^{(1)})), \tag{5.2.9}$$

$$\partial_1\Omega_j^{(2)} = \frac{i}{D_j}(\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(q)})_{q\leq l} + \frac{\omega_1''(k_j)}{2}(\partial_1^2\Omega_j^{(1)}) - i(\partial_2\Omega_j^{(1)})), \tag{5.2.10}$$

with $D_j := \omega_1'(k_j) - \omega_1'(k_{3-j})$ and the nonlinear terms

$$\tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)})B_j^{(l)} := 6s_j q_1^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} |B_{3-j}^{(1)}|^2 B_j^{(l)},$$

$$\begin{aligned}\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(1)})_{q\leq 2}B_j^{(1)} :=\ & 6s_j q_1^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} (\overline{B_{3-j}^{(1)}} B_j^{(1)} B_{3-j}^{(2)} + \overline{B_{3-j}^{(2)}} B_j^{(1)} B_{3-j}^{(1)}) \\ & + 6s_j' q_1^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} (-i\partial_1 |B_{3-j}^{(1)}|^2) B_j^{(1)} \\ & + 6s_j q_1^{[j]} q_{1-}'^{[3-j]} q_1^{[3-j]} B_j^{(1)} (-i\partial_1 B_{3-j}^{(1)}) \overline{B_{3-j}^{(1)}} \\ & + 6s_j q_1^{[j]} q_{1-}^{[3-j]} q_1'^{[3-j]} B_j^{(1)} B_{3-j}^{(1)} (-i\partial_1 \overline{B_{3-j}^{(1)}}),\end{aligned}$$

$$\begin{aligned}\tilde{N}_{2,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(1)})_{q\leq 2}(B_j^{(1)}) :=\ & 6s_j q_1'^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} (-i\partial_1 B_j^{(1)}) B_{3-j}^{(1)} \overline{B_{3-j}^{(1)}} \\ & + 6s_j' q_1^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} \overline{B_{3-j}^{(1)}} B_j^{(1)} (-i\partial_1 B_j^{(1)}),\end{aligned}$$

where we used the notation

$$\tilde{Q}(k) := (\tilde{q}_{lm}(k))_{lm}, \quad \tilde{Q}(k)^{-1} y := (\tilde{s}_l(k))_l,$$
$$q_1^{[j]} := \tilde{q}_{31}(k_j), \quad q_{1-}^{[j]} := \tilde{q}_{32}(-k_j), \quad s_j := \gamma\tilde{s}_1(k_j),$$

the prime denotes a k–derivative, and, finally, the algebraic system of equations as described in Sec.3.2.2.

Lemma 5.2.1. *Let V^{an} be the two-pulse ansatz* (5.2.4) *where*

$$B_j^{(r)} \in C\left([0,T_0]; H^{s_B-3(r-1)+l} \cap H^{s_B-3(r-1)}(l)\right), \quad s_B \geq 10, l \geq 2,$$

fulfill the extended modulation system (5.2.5)-(5.2.10)*. Then for all $T_0 > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\|_{(L^2(s))^4} \leq C\varepsilon^{11/2},$$

where the residual is given by (5.2.1)*.*

Proof. The proof follows the proof of Lemma 4.2.1 line-by-line. It is merely a little more intricate due to the additional technicalities introduced by the branch interactions. However, we have already discussed how to treat those in (3.2.2). □

By this, we completed the second step of the solution strategy described in Sec.3.3!

5.3 Approximation result

We are now ready to formulate the approximation theorem.

Theorem 5.3.1. *Given the conditions of Lemma 5.2.1, for all $T_0 > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u - [(\mathcal{F}^{-1}Q\mathcal{F})\{V^{an}]_1\}\|_{C_b^{s-1}} \leq C\varepsilon^{7/2},$$

where u solves the Maxwell-Lorentz system (5.0.1)-(5.0.2) *restricted to the subspace*

$$\{(u,p) \,|\, \int_{\mathbb{R}} (\partial_t u)\, dx = \int_{\mathbb{R}} (\partial_t p)\, dx\}.$$

Proof. Since both **CONDITION 1** and **CONDITION 2** are fulfilled (after the preliminary transformation $T(k)$), the proof follows exactly the third step of Sec.3.3. So, the only remaining task is to specify the function spaces:

$$\mathcal{X} := (L^2(s))^4, \quad \hat{\mathcal{X}} := (H^s)^4.$$

□

5.4 Description of pulse interaction

So far, we referred to the quantities $\Omega_j^{(r)}, \Psi_j^{(1)}$ only as interaction quantities, since it is not clear at first sight what role they will play for the corresponding pulses in x-space, i.e. for $U^{an} = \mathcal{F}^{-1}\tilde{Q}\mathcal{F}\{V^{an}\}$. Surprisingly, it turns out that the interpretation coincides with the one given in Ch. 2. This can be seen as follows:

Let us take our extended two-pulse ansatz (5.2.4) and expand it around the "classical perturbation variables" (much as we have done it in Ch. 2 to analyze the accuracy r_{acc} we require to ensure the validity of the interaction effect formulas, see Rem. 2.2.3). The expansion yields

$$\begin{aligned} v^{an} = \sum_{j=1,2} e^{i\underline{Y}_j}(\quad & \varepsilon \quad \underbrace{[B_j^{(1)}]}_{=:\tilde{B}_j^{(1)}} \\ & +\varepsilon^2 \ \underbrace{[B_j^{(2)} + iB_j^{(1)}\Omega_j^{(1)}]}_{=:\tilde{B}_j^{(2)}} \\ & +\varepsilon^3 \ \underbrace{[B_j^{(3)} + iB_j^{(1)}\Omega_j^{(2)} + iB_j^{(2)}\Omega_j^{(1)} + (\partial_1 B_j^{(1)})\Psi_j^{(1)} - \frac{1}{2}B_j^{(1)}(\Omega_j^{(1)})^2]}_{=:\tilde{B}_j^{(3)}} \\ & +\mathcal{O}(\varepsilon^4)) + M_{mixed,\alpha}. \end{aligned}$$

If we perform the change of variables $U^{an} = \mathcal{F}^{-1}\tilde{Q}_{trunc}\mathcal{F}\{V^{an}\}$, where $\tilde{Q}_{trunc}$ contains only the expansion of the coefficients $\tilde{q}_{ln}$ of $\tilde{Q}$, we get

$$\begin{aligned} u^{an} = \sum_{j=1,2} e^{i\underline{Y}_j}(\quad & \varepsilon \ [\tilde{q}_{11}(k_j)\tilde{B}_j^{(1)}] \\ & +\varepsilon^2 \ [\tilde{q}_{11}(k_j)\tilde{B}_j^{(2)} + i\tilde{q}'_{11}(k_j)(\partial_1\tilde{B}_j^{(1)})] \\ & +\varepsilon^3 \ [\tilde{q}_{11}(k_j)\tilde{B}_j^{(3)} + i\tilde{q}'_{11}(k_j)(\partial_1\tilde{B}_j^{(2)}) - \tilde{q}''_{11}(k_1)(\partial_1^2\tilde{B}_j^{(1)})] \\ & +\mathcal{O}(\varepsilon^4)) + M_{mixed,\alpha}. \end{aligned}$$

If we now dissolve the definition of $\tilde{B}^{(r)}$ we get

$$\begin{aligned} u^{an} = \sum_{j=1,2} e^{i\underline{Y}_j}(\quad & \varepsilon \ [\tilde{q}_{11}(k_j)B_j^{(1)}] \\ & +\varepsilon^2 \ [\tilde{q}_{11}(k_j)[B_j^{(2)} + iB_j^{(1)}\Omega_j^{(1)}] + i\tilde{q}'_{11}(k_j)(\partial_1 B_j^{(1)})] \\ & +\varepsilon^3 \ [\tilde{q}_{11}(k_j)[B_j^{(3)} + iB_j^{(1)}\Omega_j^{(2)} + iB_j^{(2)}\Omega_j^{(1)} + (\partial_1 B_j^{(1)})\Psi_j^{(1)} - \frac{1}{2}B_j^{(1)}(\Omega_j^{(1)})^2] \\ & \quad + i\tilde{q}'_{11}(k_j)\partial_1[B_j^{(2)} + iB_j^{(1)}\Omega_j^{(1)}] - \tilde{q}''_{11}(k_1)(\partial_1^2 B_j^{(1)})] \\ & +\mathcal{O}(\varepsilon^4)) + M_{mixed,\alpha}. \end{aligned}$$

In conclusion, with the definitions

$$\begin{aligned} A_j^{(1)} &:= \tilde{q}_{11}(k_j)B_j^{(1)}, \\ A_j^{(2)} &:= \tilde{q}_{11}(k_j)B_j^{(2)} + i\tilde{q}'_{11}(k_j)(\partial_1 B_j^{(1)}), \\ A_j^{(3)} &:= \tilde{q}_{11}(k_j)B_j^{(3)} + i\tilde{q}'_{11}(k_j)(\partial_1 B_j^{(2)}) - \frac{1}{2}\tilde{q}''_{11}(k_j)(\partial_1^2 B_j^{(1)}), \\ \tilde{\Omega}_j^{(1)} &:= \Omega_j^{(1)} \\ \tilde{\Psi}_j^{(1)} &:= \Psi_j^{(1)}, \\ \tilde{\Omega}_j^{(2)} &:= \Omega_j^{(2)} + i\tilde{q}'_{11}(k_j)(\partial_1\Omega_j^{(1)}), \end{aligned}$$

we get

$$\begin{aligned} u^{an} = \sum_{j=1,2} e^{i\underline{Y}_j}(\quad & \varepsilon \ [A_j^{(1)}] + \varepsilon^2 \ [A_j^{(2)} + iA_j^{(1)}\tilde{\Omega}_j^{(1)}] \\ & + \varepsilon^3 \ [A_j^{(3)} + iA_j^{(1)}\tilde{\Omega}_j^{(2)} + iA_j^{(2)}\tilde{\Omega}_j^{(1)} + (\partial_1 A_j^{(1)})\tilde{\Psi}_j^{(1)} - \frac{1}{2}A_j^{(1)}(\tilde{\Omega}_j^{(1)})^2] \\ & + \mathcal{O}(\varepsilon^4)) + M_{mixed,\alpha}, \end{aligned} \tag{5.4.1}$$

which becomes — when undoing the expansion — the extended two-pulse ansatz

$$\begin{aligned} u^{an}(x,t) &= \sum_{r=1}^{3}\sum_{j=1,2} \varepsilon^r A_j^{(r)}(\underline{\tilde{X}}_j, \underline{T}) E_j + M_{mixed,\alpha} + c.c. \\ E_j &= e^{ik_j x + i\omega_1(k_j)t + i\sum_{r=1,2}\varepsilon^r \tilde{\Omega}_j^{(r)}(\underline{X}_{3-j},\underline{T})}, \quad \underline{X} = \varepsilon x, \quad \underline{T} = \varepsilon^2 t, \\ \underline{\tilde{X}}_j &= \varepsilon(x + \omega_1(k_j)t + \varepsilon\tilde{\Psi}_j^{(1)}(\underline{X}_{3-j},\underline{T})). \end{aligned}$$

Therefore,the interpretation of the quantities as given in Sec. 2.2 still holds!

5.5 Chapter Summary

We managed to give a separate description of the pure internal dynamics of the individual pulses and the pure interaction dynamics between them in the context of a system that exhibits a multibranch dispersion relation and features a double zero eigenvalue (which complicates the transformation to "natural coordinates"). We were able to establish the following result.

Two pulses traveling on different carrier waves obeying the Maxwell-Lorentz system

$$\begin{aligned} \partial_t^2 u &= \partial_x^2 u + \partial_t^2 p \\ \partial_t^2 p &= -\omega_0^2 p - d^2 u + \gamma p^3 \end{aligned}$$

with $x, t, u = u(x,t), p = p(x,t) \in \mathbb{R}$ and $\omega_0, d, \gamma \in \mathbb{R}$, restricted to the subspace of solutions fulfilling

$$\int_{\mathbb{R}} (\partial_t u)\, dx = \int_{\mathbb{R}} (\partial_t p)\, dx,$$

can be approximated on the long time scale of order $\mathcal{O}(1/\varepsilon^2)$ by the ansatz

$$\begin{aligned} u^{an} &= \sum_{r=1}^{3}\sum_{j=1,2} \varepsilon^r A_j^{(r)}(\underline{\tilde{X}}_j, \underline{T}) E_j + M_{mixed,\alpha} + c.c., \\ E_j &= e^{ik_j x + i\omega_1(k_j)t + i\sum_{r=1,2}\varepsilon^r \tilde{\Omega}_j^{(r)}(\underline{X}_{3-j},\underline{T})}, \quad \underline{T} = \varepsilon^2 t, \\ \underline{\tilde{X}}_j &= \varepsilon(x + \omega_1(k_j)t + \varepsilon\tilde{\Psi}_j^{(1)}(\underline{X}_{3-j},\underline{T})). \end{aligned}$$

and similarly for p^{an}, if the envelopes $A_j^{(r)}$ fulfill

$$A_j^{(r)} := \sum_{m=0}^{r-1} \frac{q_{11}^{(m)}(k_j)}{m!} B_j^{(m)},$$

and $B_j^{(r)}$ fulfill the recursively solvable modulation system for internal dynamics

$$\begin{aligned} \partial_2 B_j^{(1)} &= -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(1)} + 3 s_j q_1^{[j]} (q_{1-}^{[j]})^2 |B_j^{(1)}|^2 B_j^{(1)}, \\ \partial_2 B_j^{(2)} &= -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\le 2}, \\ \partial_2 B_j^{(3)} &= -i\frac{\omega_1''(k_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega_1'''(k_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega_1^{(4)}(k_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\le 3}, \end{aligned}$$

and if the interaction quantities solve the following equations:

The **carrier shift** $\tilde{\Omega}_j^{(1)}$ (see Fig.2.3) has to satisfy the ordinary differential equation

$$\partial_1 \tilde{\Omega}_j^{(1)} = \frac{i}{D_j} 6 s_j q_1^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} |B_{3-j}^{(1)}|^2.$$

The **envelope position shift** $\tilde{\Psi}_j^{(1)}$ (see Fig.2.4) has to satisfy the ordinary differential equation

$$\partial_1 \tilde{\Psi}_j^{(1)} = \frac{1}{D_j} (6 s_j q_1'^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} |B_{3-j}^{(1)}|^2 + 6 s_j' q_1^{[j]} q_{1-}^{[3-j]} q_1^{[3-j]} \overline{B_{3-j}^{(1)}} B_j^{(1)} + \omega_1''(k_j)(\partial_1 \Omega_j^{(1)})).$$

The **second order carrier shift and shape correction** are described by the real and imaginary part of $\tilde{\Omega}_j^{(2)}$ which is given by

$$\tilde{\Omega}_j^{(2)} := \Omega_j^{(2)} + i q_{11}'(k_j)(\partial_1 \tilde{\Omega}_j^{(1)})$$

where $\Omega_j^{(2)}$ fulfills

$$\partial_1 \Omega_j^{(2)} = \frac{i}{D_j} (\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}} (B_{3-j}^{(q)})_{q \le l} + \frac{\omega_1''(k_j)}{2} (\partial_1^2 \tilde{\Omega}_j^{(1)}) - i(\partial_2 \tilde{\Omega}_j^{(1)})).$$

Note, that the illustration of the interaction is exactly as for the cubic Klein-Gordon case!

This result would have been very hard (if not impossible) to achieve with the conventional x-space method as described in Ch. 2 and, therefore, is something rather unexpected. However, using the methods developed in Ch. 3 made the transfer fairly easy and transparent.

6 Pulse interaction for nonlinear wave equations with periodic coefficients

This final chapter demonstrates the "separation of internal and interaction dynamics" for a cubic Klein-Gordon equation with periodic coefficients

$$\partial_t^2 u = \partial_x^2 u + a(x)u + \gamma b(x)u^3 \tag{6.0.1}$$

where $x, t, u(x,t) \in \mathbb{R}$ and a, b are real, smooth, 2π-periodic functions with $a(x) > \tilde{a} > 0$ and $b(x) > \tilde{b} > 0$ for $x \in \mathbb{R}$ and some constants $\tilde{a}, \tilde{b} \in \mathbb{R}$. The corresponding dispersion relation is depicted in Fig. 3.1. It features infinitely many branches. This completes the discussion of mathematical challenges that can be encountered when dealing with pulse interaction in the setting of nonlinear wave equations.

The solution strategy for this setting resembles the one described in 3.3, but it really involves some new elements.

1. **Transformation to "natural coordinates".** The derivation of modulation equations in x-space can be quite hard to accomplish for equations with periodic coefficients. Hence, we transform again the equation to its "natural coordinates" representation (3.1.1), i.e.

   $$\partial_t \tilde{V}_n = \tilde{D}_n \tilde{V}_n + \tilde{N}_n[\tilde{V}], \quad n \in \mathcal{N},$$

   for $\tilde{V} = (\tilde{V}_n)_{n\in\mathcal{N}}$ with $\mathcal{N} = \mathbb{N}$. The transformation will differ from the ones used in the chapters before, since for equations with periodic coefficients Fourier analysis is replaced by Bloch analysis (see [Ea73],[SR80] for a good introduction), which is a little more difficult to handle.

2. **Formal derivation of modulation equations.** In order to get the residual

   $$\widetilde{\mathrm{Res}}\{\tilde{V}\} := \left(-\partial_t \tilde{V}_n + \tilde{D}_n \tilde{V}_n + \tilde{N}_n[\tilde{V}]\right)_{n\in\mathcal{N}} := \left(\widetilde{\mathrm{Res}}_n\{\tilde{V}\}\right)_{n\in\mathcal{N}},$$

   small, we need to use the reasoning from Sec. 3.2.1. However, we will have to add an additional step to our analysis before the expansion and truncation. This is due to the fact that the dispersion relation is periodic and, so, we cannot directly employ inverse Fourier transform (remember, this is needed since the Fourier transform of the extended ansatz (3.2.10) cannot be computed explicitly). Again, making use of the spectral concentration of pulses, we will cure this problem by introducing a cut–off function and thereby define a residual $\widetilde{\mathrm{Res}}^{cut}$. This will then put us into a position, where we can follow the usual path, i.e. following the example in Sec. 3.2.1 we expand and truncate all l-dependent[1] terms in the residual and then employ inverse Fourier transform, which yields the auxiliary residual $\underline{\mathrm{Res}}$ that solely features x-space expressions.
   Hence, we will have to make two things rigorous: First, we have to make sure, that the cut–off does not produce significant errors, i.e. that formally

   $$\widetilde{\mathrm{Res}}\{\tilde{V}^{an}\} = \widetilde{\mathrm{Res}}^{cut}\{\tilde{V}^{an}\} + \mathcal{O}(\varepsilon^5),$$

[1] The spectral parameter is usually called l instead of k in the context of Bloch analysis, we stick to this convention

where $\widetilde{\mathrm{Res}}^{cut}$ is created by cutting off all l-dependent coefficients respecting the spectral concentration of the occurring terms. Furthermore, we have to show that the error introduced by expansion and truncation can be neglected, i.e. that we have formally

$$\widetilde{\mathrm{Res}}\{\mathcal{F}\{V^{an}\}\} = \mathcal{F}\underline{\mathrm{Res}}\{V_n^{an}\} + \mathcal{O}(\varepsilon^5),$$

where $\underline{\mathrm{Res}}$ is defined according to 3.2.8. Note, that both reduction procedures really only make use of estimates like in Lemma 3.2.2.

3. **Approximation result.** The justification proof will be very similar to the one in Ch. 1. As opposed to the strategy from the last chapters, we will carry it out in x-space by using energy estimates and Gronwall's inequality.

4. **Description of pulse interaction.** The interpretation of U^{an} will turn out to be exactly as in the Maxwell-Lorentz case (cf. 5.4)!

Note, that this chapter is structured according to these four steps of the solution strategy.

6.1 Transformation to "natural coordinates"

This section merges the results from the work of K. Busch, G. Schneider, L. Tkeshelashvili and H. Uecker from 2006 [BSTU06] with the pulse interaction analysis developed in Ch. 3.

Applying Bloch representation for u, i.e.

$$u = u(x,t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \breve{u}(l,x,t)e^{ilx}\ dl =: \mathcal{B}^{-1}\{\breve{u}\}$$

we get

$$\partial_t^2\breve{u} = -L_l[\breve{u}] + \gamma b(x)\breve{u}^{\star 3}, \quad \breve{u}(l,x,t) = \breve{u}(l,x+2\pi,t),$$

with the linear operator

$$L_l[\breve{u}] = -[(\partial_x + il)^2 + a(x)]\breve{u}$$

and the convolution

$$(\breve{u} \star \breve{v})(l) := \int_{-\frac{1}{2}}^{\frac{1}{2}} \breve{u}(l-m)\breve{v}(m)\ dm$$

for the Bloch coefficients

$$\breve{u} = \breve{u}(l,x,t) := \sum_{n\in\mathbb{Z}} \hat{u}(l-n)e^{inx} =: \mathcal{B}\{u\}$$

where $\hat{u} = \mathcal{F}\{u\}$ is the Fourier transform.

Intermezzo: Features of the operator L_l and eigenfunction representation. For fixed l the operator L_l is self-adjoint and positive definite in the space $L^2([0,2\pi);\mathbb{C})$. Thus, for each fixed l there are countably many eigenvalues $\lambda_j(l) > 0$ with corresponding eigenfunctions $f_j(\cdot,l)$ which form a Schauder basis of $L^2([0,2\pi);\mathbb{C})$. Furthermore, $f_j(\cdot,l)$ are smooth in l (cf. [BSTU06]).

Lemma 6.1.1. *The mapping*

$$\begin{aligned}\mathcal{P} : L^2\left((-\frac{1}{2},\frac{1}{2}); H^s((0,2\pi))\right);\mathbb{C}) &\longrightarrow L^2\left((-\frac{1}{2},\frac{1}{2}); l^2(s)\right)\\ \breve{u} &\longmapsto (<\breve{u}(l,\cdot), f_j(l,\cdot)>)_{j\in\mathbb{N}}\end{aligned}$$

is an isomorphism.

This leads to the eigenfunction representation of $\breve{u}$ given by

$$\breve{u} = \breve{u}(l,x,t) = \sum_{j\in\mathbb{N}} \tilde{u}_j(l,t) f_j(l,x) =: \mathcal{P}^{-1}\{(\tilde{u}_j)_{j\in\mathbb{N}}\}$$

with

$$(\tilde{u}_j)_{j\in\mathbb{N}} = (\tilde{u}_j(l,t))_{j\in\mathbb{N}} = (<\breve{u}(l,\cdot,t), f_j(l,\cdot)>)_{j\in\mathbb{N}} =: \mathcal{P}\{\breve{u}\}$$

and, hence, to equations for $\tilde{u}_j$ given by

$$\partial_t^2 \tilde{u}_j = -\omega_j^2 \tilde{u}_j + \gamma < b(\cdot)\breve{u}^{\star 3}, f_j(l,\cdot)>, \quad \omega_j = \omega_j(l) := \sqrt{\lambda_j(l)},$$

with $j \in \mathbb{N}, l \in [-\frac{1}{2},\frac{1}{2})$.
At this point we will use the scaled (see 4.1.1 for a motivation) and arranged variables $\tilde{Z}_j := (\tilde{Z}_{j,1}, \tilde{Z}_{j,2})^T$ with

$$\begin{aligned}\tilde{Z}_{j,1}(l,t) &:= \tilde{u}_j(l,t),\\ \tilde{Z}_{j,2}(l,t) &:= \frac{1}{\omega_j(l)}\partial_t \tilde{u}_j(l,t),\end{aligned}$$

to arrive at

$$\partial_t \tilde{Z}_j = \tilde{A}_j \tilde{Z}_j + \hat{N}_j[(\tilde{Z}_j)_{j\in\mathbb{N}}],$$

with

$$\tilde{A}_j = \tilde{A}_j(l) := \begin{pmatrix} 0 & \omega_j(l) \\ -\omega_j(l) & 0 \end{pmatrix},$$

and the nonlinearity

$$\hat{N}_j[(\tilde{Z}_j)_{j\in\mathbb{N}}] = \hat{N}_j[(\tilde{Z}_j)_{j\in\mathbb{N}}](l) := \gamma\frac{1}{\omega_j(l)} < b(\cdot)(\mathcal{P}^{-1}\{(\tilde{Z}_{j,1})_{j\in\mathbb{N}}\})^{\star 3}, f_j(l,\cdot) > y$$

with the notation $y = (0,1)^T$.
Finally, the diagonalizing transformation $\tilde{V}_n := Q^* \tilde{Z}_n$ where

$$Q := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix},$$

brings our system to the form

$$\partial_t \tilde{V}_n = \tilde{D}_n \tilde{V}_n + \tilde{N}_n[(\tilde{V}_n)_{n\in\mathbb{N}}], \tag{6.1.1}$$

with linear part

$$\tilde{D}_n = \tilde{D}_n(l) := \begin{pmatrix} i\omega_n(l) & 0 \\ 0 & -i\omega_n(l) \end{pmatrix},$$

and nonlinear part

$$\tilde{N}_n[(\tilde{V}_n)_{n\in\mathbb{N}}] = \tilde{N}_n[(\tilde{V}_n)_{n\in\mathbb{N}}](l) := \gamma \frac{1}{\omega_j(l)} < b(\cdot)(\mathcal{P}^{-1}\{(Q\tilde{V}_n)_{n\in\mathbb{N}}\})^{\star 3}, f_j(l,\cdot) > Q^* y$$

which has exactly the structure of (3.1.1), i.e. we brought the system to its representation in "natural coordinates". Hence, it is possible to make use of the formal analysis performed in Sec. 3.2.

6.2 Formal derivation of modulation equations

Suppose that we manage to construct a two-pulse ansatz $\tilde{V}^{an} := (\tilde{V}_j^{an})_{j\in\mathbb{N}}$ such that

$$\|\widetilde{\mathrm{Res}}\{\tilde{V}^{an}\}\|_{L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2)} \leq C\varepsilon^{11/2},$$

with

$$\widetilde{\mathrm{Res}}\{\tilde{V}\} := \left(\partial_t \tilde{V}_n + \tilde{D}_n \tilde{V}_n + \tilde{N}_n[\tilde{V}]\right)_{n\in\mathcal{N}} := \left(\widetilde{\mathrm{Res}}_n\{\tilde{V}\}\right)_{n\in\mathcal{N}}. \tag{6.2.1}$$

If we set

$$u^{an} := \underbrace{\mathcal{B}^{-1}\mathcal{P}^{-1}\mathcal{T}^{-1}(Q}_{=:\mathcal{G}^{-1}} V_j^{an})_{j\in\mathbb{N}},$$

where we used the notation

$$\begin{aligned} \mathcal{T}\{(\tilde{u}_j)_{j\in\mathbb{N}}\} &:= (\tilde{Z}_j)_{j\in\mathbb{N}} = (\tilde{u}_j, \frac{1}{\omega_j(l)}\partial_t \tilde{u}_j)^T_{j\in\mathbb{N}}, \\ \mathcal{T}^{-1}\{(\tilde{Z}_j)_{j\in\mathbb{N}}\} &:= (\tilde{u}_j)_{j\in\mathbb{N}}, \end{aligned}$$

to formalize the introduction of the variable $\tilde{Z}_j$, which enables us to cast the second order equation into a first order system, then we can conclude that

$$\begin{aligned} \|\mathrm{Res}\{u^{an}\}\|_{H^s} &= \|\mathcal{G}^{-1}\mathcal{G}\mathrm{Res}\{u^{an}\}\|_{H^s} \\ &= \|\mathcal{G}^{-1}\mathcal{G}\mathrm{Res}\{\mathcal{G}^{-1}\tilde{V}^{an}\}\|_{H^s} \\ &= \|\mathcal{G}^{-1}\widetilde{\mathrm{Res}}\{\tilde{V}^{an}\}\|_{H^s} \\ &\leq |||\mathcal{G}^{-1}|||_{L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2)\longrightarrow H^s} \|\widetilde{\mathrm{Res}}\{\tilde{V}^{an}\}\|_{L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2)} \qquad (6.2.2) \\ &\leq C\varepsilon^{11/2} \qquad (6.2.3) \end{aligned}$$

if $\mathcal{G}^{-1}$ is bounded. To sum up, we can boil down the problem of diminishing $\mathrm{Res}\{u^{an}\}$ to

1. showing that $|||\mathcal{G}^{-1}|||_{L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2)\longrightarrow H^s} \leq C$, for some $C > 0$ and
2. diminishing the residual $\widetilde{\mathrm{Res}}\{\tilde{V}^{an}\}$ of the "system in natural coordinates".

6.2.1 Estimate for the diagonalizing transformation

The boundedness of the diagonalizing transformation that brings our equation to its "natural coordinates" representation is guaranteed by the following Lemma.

Lemma 6.2.1. *The mapping* $\mathcal{G}^{-1} : L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2) \longrightarrow H^s$ *defined by* $\mathcal{G}^{-1} := \mathcal{B}^{-1}\mathcal{P}^{-1}\mathcal{T}^{-1}Q$ *is bounded.*

Proof. Clearly, $\mathcal{T}^{-1}Q$ does not cause any difficulties since it is merely a unitary transformation followed by a projection. Furthermore, we have that the Bloch transform is an isometric isomorphism from $H^s(\mathbb{R};\mathbb{R})$ into $L^2([-\frac{1}{2},\frac{1}{2});H^s([0,2\pi);\mathbb{C}))$. Finally, the boundedness of $\mathcal{P}^{-1}$ follows from Lemma 6.1.1. □

6.2.2 Getting the residual small in natural coordinates

The remainder of this section is dedicated to diminishing the residual of the "system in natural coordinates" defined in (6.2.1), i.e.

$$\widetilde{\mathrm{Res}}_n\{\tilde{V}\} = -\partial_t\tilde{V}_n + \tilde{D}_n\tilde{V}_n + \tilde{N}_n[\tilde{V}], \quad \tilde{V} := (\tilde{V}_j)_{j\in\mathbb{N}}, n \in \mathbb{N}.$$

As already alluded to in the last chapters, we cannot directly employ the extended two-pulse ansatz for the residual (6.2.1), since we cannot explicitly Fourier transform it (due to the appearance of nested functions). In the last chapter, we could overcome this difficulty by making use of the spectral concentration property of pulse solutions and subsequent inverse Fourier transformation, which brought the residual to a form involving only x-space expressions, such that it was possible to employ our extended ansatz (3.2.10). However, for the periodic case we have that the dispersion relation is periodic and, hence, we first have to get rid of this periodicity of the l-dependent components of (6.2.1) before we can employ the inverse Fourier transform. To this end we introduce a cut–off, which is tailored according to the ansatz (we will see in a moment, what this means).

The cut–off. Suppose that

$$\tilde{v}(l,t) := \hat{B}(\underline{L}_1)e^{i\omega_1'(l_1)\underline{L}_1 t},$$

and that $\hat{B} \in L^2(s)$ (hence, $\tilde{v}$ has the structure of the Fourier transform of a one-pulse). Then we define

$$\omega_j^{cut,l_1}(l) := \chi^{l_1}\omega_j(l), \quad l \in \mathbb{R},$$

where we used the notation

$$\chi^{l_1} := \chi_{[l_1-\frac{\varepsilon}{2},l_1+\frac{\varepsilon}{2}]}$$

for the cut-off function to emphasize that the cut-off depends on the type of solution under consideration (or, to be precise, its spectral content). If we choose $s_B - s \geq 5$, this cut-off produces error terms beyond the accuracy we need and is therefore negligible. In detail,

$$\begin{aligned}
&\|(\omega_j(l) - \omega_j^{cut,l_1}(l))\tilde{v}(l)\|_{L^2(s)} \\
&\quad = \varepsilon^{1/2}\,\|(\omega_j(l_1+\varepsilon\underline{L}_1) - \omega_j^{cut}(l_1+\varepsilon\underline{L}_1))\hat{B}(\underline{L})\|_{L^2(s_B)} \\
&\quad \leq \varepsilon^{1/2}\left(\sup_{l\in[-\frac{1}{2},\frac{1}{2}]}|\omega_j(l)|\right)\sup_{\underline{L}_1\in\mathbb{R}}\left|(1-\chi_{[-\frac{1}{2},\frac{1}{2}]}(\underline{L}_1))(1+\underline{L}_1^2)^{\frac{s-s_B}{2}}\right|\,\|\hat{B}\|_{L^2(s_B)} \\
&\quad \leq \varepsilon^{1/2+(s_B-s)}.
\end{aligned} \tag{6.2.4}$$

Using the above estimate, we can conclude the following

Lemma 6.2.2. *Let*

$$\tilde{V}_1(l,t) = (\sum_{j=1,2} \hat{B}_j(\underline{L}_j) e^{i\omega_1'(l_j)\underline{L}_j t}, \sum_{j=1,2} \hat{B}_j(\underline{L}_{j-}) e^{-i\omega_1'(l_j)\underline{L}_{j-} t})^T, \quad \tilde{V}_n(l,t) = \mathcal{O}(\varepsilon^3), n \geq 2,$$

with $0 < \varepsilon \ll 1, l_1 \in \mathbb{R}, \underline{L}_j := \frac{l-l_j}{\varepsilon}, \underline{L}_{j-} := \frac{l+l_j}{\varepsilon}$ *where* $|l_j| \leq \frac{1}{10} - \delta_1$, *for some* $\delta_1 > 0$ *independent of* ε *and* $\hat{B}_j \in L^2(s)$. *Then for* $s_B - s \leq 5$ *we have*

$$\|\widetilde{\mathrm{Res}}_n\{\tilde{V}^{an}\} - \widetilde{\mathrm{Res}}_n^{cut}\{\tilde{V}^{an}\}\|_{L^2(s)} \leq C\varepsilon^{11/2},$$

for $n = 1, 2$, *where* $\widetilde{\mathrm{Res}}\{\tilde{V}\} := (\widetilde{\mathrm{Res}}_n\{\tilde{V}\})_{n\in\mathbb{N}}$ *is according to* (6.2.1) *and*

$$\widetilde{\mathrm{Res}}_n^{cut}\{\tilde{V}\} := -\partial_t \tilde{V}_n + \tilde{D}_n^{cut}\tilde{V}_n + \tilde{N}_n^{cut}[(\tilde{V}_n)_{n\in\mathbb{N}}],$$

with

$$\tilde{D}_n^{cut} = \tilde{D}_n^{cut}(l) := \begin{pmatrix} i\omega_n^{cut,l_1,l_2}(l) & 0 \\ 0 & -i\omega_n^{cut,l_1,l_2}(l) \end{pmatrix}, \quad \omega_n^{cut,l_1,l_2}(l) := \sum_{j=1,2} \chi_{[l_j-\frac{\varepsilon}{2},l_j+\frac{\varepsilon}{2}]}(l)\,\omega_n(l),$$

and

$$\tilde{N}_n^{cut}[(\tilde{V}_n)_{n\in\mathbb{N}}](l) := \chi^{(l_w)_w} \frac{\gamma}{\omega_n(l)} < b(\cdot)(\mathcal{P}^{-1}\{(Q\tilde{V}_n)_{n\in\mathbb{N}}\})^{*3}, \chi^{(l_w)_w} f_n(l,\cdot) > Q^* y, \quad l \in \mathbb{R},$$

where $*$ *is now the usual convolution (!) and where* $\chi^{(l_w)_w}$ *is a cut-off function, which "cuts out" the areas around the wave numbers generated by the nonlinearity.*

Proof. The proof only amounts to explicitly computing the residual and employing a cut–off function, depending on the wavenumber where the resp. term is concentrated around. The assumption $|l_j| \leq \frac{1}{10} - \delta_1$, might seem quite random. Nevertheless, it is an assumption we need to keep the technicalities on a lower level, since it ensures that all the higher order and mixed terms generated by the nonlinearity — and consequently all concentration points — are within the interval $[-\frac{1}{2}, \frac{1}{2})$ and we do not bother about keeping track, which branch the new frequencies really belong to. □

With this we are at a point where we can proceed as described in Sec. 3.2.1.

Expansion and truncation As explained in Sec. 3.2, we will have to use the spectral concentration property and inverse Fourier transformation in order to be able to perform pulse interaction analysis as in Ch. 2. Hence, we will again introduce the x-space residual $\underline{\mathrm{Res}}$ for the formal derivation of the modulation system.
Comparing our residual $(\widetilde{\mathrm{Res}}_n^{cut})_{n\in\mathbb{N}}$ to the residual from the example in Sec. 3.2.1, we see that each $\widetilde{\mathrm{Res}}_n^{cut}$ is very similar to the residual (3.2.3). The only difference is that the nonlinearity features also a projection on an eigenfunction. In detail, we have

$$[\widetilde{N}[\tilde{V}](l)]_n := \gamma \frac{i}{4\omega_n(l)} \left(\int_0^{2\pi} b(x) \left(f_1(\cdot,x)\tilde{v}_1(\cdot) + f_1(\cdot,x)\tilde{v}_2(\cdot) + \ldots\right)^{*3}(l)\, \overline{f_n(l,x)}\, dx \right) \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

However, this does not lead to major changes in the analysis that we have performed in 3.2.1.

Hence, if we assume that the pulses that we want to consider in the following have their spectral

content in the first pair of branches $\omega_1, -\omega_1$, we carry out the analysis for $\widetilde{\text{Res}}_1^{cut}$ exactly as we have done it in the example of Sec. 3.2.1, i.e. expand and truncate. In the nonlinearity we also perform the projection onto the eigenfunction.

For the rest of the residual, i.e. $\widetilde{\text{Res}}_n^{cut}, n \geq 2$, we perform the expansion and truncation around the wave numbers of all mixed and higher order harmonics (analogously to Sec. 5.2).

In summary, we introduce in the spirit of (3.2.7) the x-space residual

$$\underline{\text{Res}}\{V\} := (\underline{\text{Res}}_n\{V\})_{n\in\mathbb{N}} := (\underline{D}_n[V_n] + \underline{N}_n[V])_{n\in\mathbb{N}},$$

and argue by the formal relation (3.2.8), i.e.

$$\widetilde{\text{Res}}\{(\mathcal{F}\{V_n^{an}\})_{n\in\mathcal{N}}\} = \mathcal{F}\underline{\text{Res}}\{(V_n)_{n\in\mathcal{N}}\} + \mathcal{O}(\varepsilon^5),$$

that it really suffices to come up with an ansatz that makes $\underline{\text{Res}}\{V\}$ small.

We refrain from writing down all the technical details which lead to the definition of $\underline{\text{Res}}$. The computations are a straightforward extension of the ideas explained in the example of Sec. 3.2.1.

In summary, if we make an ansatz V^{an} such that

$$\|\underline{\text{Res}}_n\{V^{an}\}\|_{(H^s)^2} \leq C\varepsilon^{11/2}, \quad n \in \mathcal{N},$$

we immediately get

$$\|\widetilde{\text{Res}}\{\mathcal{F}V^{an}\}\|_{L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2)} \leq C\varepsilon^{11/2}.$$

The extended modulation system. Consider the extended two-pulse ansatz

$$V^{an} = (v^{an}, \overline{v^{an}})^T, \quad V_n^{an} = (M_{mixed,\alpha}^n, \overline{M_{mixed,\alpha}^n}), n \geq 2, \tag{6.2.5}$$

with

$$v^{an}(x,t) = \sum_{j=1,2}\sum_{r=1}^{3} \varepsilon^r B_j^{(r)}(\underline{X}_j, \underline{T})e^{i\underline{Y}_j} + M_{mixed,\alpha}^1,$$

$$\begin{aligned}
\underline{X}_j &:= \underline{X} + \varepsilon\omega'(l_j)t + \varepsilon^2\Psi_j^{(1)}(\underline{X} + \varepsilon\omega'(l_{3-j})t, \underline{T}),\\
\underline{Y}_j &:= l_j x + \omega(l_j)t + \sum_{l=1,2}\varepsilon^l\Omega_j^{(l)}(\underline{X} + \varepsilon\omega'(l_{3-j})t, \underline{T}),\\
\underline{X} &:= \varepsilon x, \quad \underline{T} := \varepsilon^2 t,
\end{aligned}$$

with the technical (but not really necessary) condition $|l_j| \leq \frac{1}{10} - \delta_1$ (see proof of Lemma 6.2.2), where the mixed and higher order harmonic terms $M_{mixed,\alpha}^n = \mathcal{O}(\varepsilon^3), n \geq 1$, are according to Sec. 3.2.2. Plugging this ansatz into $\underline{\text{Res}}\{V^{an}\}$ gives the following extended modulation system:
The recursively solvable modulation system of *internal dynamics equations*

$$\partial_2 B_j^{(1)} = -i\frac{\omega_1''(l_j)}{2}\partial_1^2 B_j^{(1)} + \tilde{N}_{1,j}^{internal}(B_j^{(1)}), \tag{6.2.6}$$

$$\partial_2 B_j^{(2)} = -i\frac{\omega_1''(l_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega_1'''(l_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\leq 2}, \tag{6.2.7}$$

$$\partial_2 B_j^{(3)} = -i\frac{\omega_1''(l_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega_1'''(l_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega_1^{(4)}(l_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\leq 3}, \tag{6.2.8}$$

and the *interaction dynamics equations*

$$\partial_1 \Omega_j^{(1)} = \frac{i}{D_j} \tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)}), \tag{6.2.9}$$

$$\partial_1 \Psi_j^{(1)} = \frac{1}{D_j} (\tilde{N}_{2,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(q)})_{q\leq 2} + \omega_1''(l_j)(\partial_1 \Omega_j^{(1)})), \tag{6.2.10}$$

$$\partial_1 \Omega_j^{(2)} = \frac{i}{D_j} (\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(q)})_{q\leq l} + \frac{\omega_1''(l_j)}{2}(\partial_1^2 \Omega_j^{(1)}) - i(\partial_2 \Omega_j^{(1)})), \tag{6.2.11}$$

with $D_j := \omega_1'(l_j) - \omega_1'(l_{3-j})$ and where the nonlinear terms are given by

$$\begin{aligned}
\tilde{N}_{1,j}^{internal}(B_j^{(1)}) &:= 3\left(s(l_j)\int_0^{2\pi} b(x)|f_1(l_1,x)|^4\,dx\right)|B_j^{(1)}|^2,\\
\tilde{N}_{1,j}^{inter,\Omega_j^{(1)}}(B_{3-j}^{(1)}) &:= 6\left(s(l_j)\int_0^{2\pi} b(x)|f_1(l_1,x)|^2|f_1(l_2,x)|^2\,dx\right)|B_{3-j}^{(1)}|^2,\\
\tilde{N}_{1,j}^{inter,\Psi_j^{(1)}}(B_{3-j}^{(1)}) &:= 6\left(s(l_j)\int_0^{2\pi} b(x)|f_1(l_2,x)|^2(\partial_l f_1(l_1,x))f_1(l_1,x)\,dx\right)|B_{3-j}^{(1)}|^2\\
&\quad + 6\left(s'(l_j)\int_0^{2\pi} b(x)|f_1(l_1,x)|^2|f_1(l_2,x)|^2\,dx\right)\overline{B_{3-j}^{(1)}}B_j^{(1)},
\end{aligned}$$

with $s(l) := -\gamma\frac{i}{4\omega_1(l)}$ where we used $f_1(-l_1,x) = \overline{f_1(l_1,x)}$. We refrain from specifying the other nonlinear terms. Furthermore, we get an algebraic system of equations as described in Sec. 3.2.2. Note, that we also get an infinite number of non-resonance conditions (as opposed to finitely many as in 3.2.2, cf. [BSTU06] for a discussion).

Lemma 6.2.3. *Consider the extended two-pulse ansatz* (6.2.5) *where*

$$B_j^{(r)} \in C\left([0,T_0]; H^{s_B-3(r-1)+w} \cap H^{s_B-3(r-1)}(w)\right), \quad s_B \geq 10, w \geq 2,$$

fulfill the extended modulation system (6.2.6)-(6.2.11). *Then for all $T_0 > 0$ there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have, if $1 \leq s \leq s_B - 10$, that*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\widetilde{\mathrm{Res}}\{\mathcal{F}V^{an}\}\|_{L^2([-\frac{1}{2},\frac{1}{2});(l^2(s))^2)} \leq C\varepsilon^{11/2},$$

and, thus,

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\mathrm{Res}\{u^{an}\|_{H^s} \leq C\varepsilon^{11/2}.$$

Since the proof follows the proof of the respective statement in Ch. 5, i.e. Thm. 5.2.1, line-by-line, we completed the second step of the solution strategy!

6.3 Approximation result

We are now ready to formulate the approximation theorem.

Theorem 6.3.1. *Given the conditions of Lemma 6.2.3 for all $T_0 > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u - \mathcal{G}^{-1}\{V^{an}\}\|_{C_b^{s-1}} \leq C\varepsilon^{7/2},$$

where u is an exact two-pulse solution of the cubic Klein-Gordon equation (6.0.1).

Proof. In contrast to the last chapters, the justification proof can be carried out in x-space. We can follow exactly the proof of Thm. 1.2.1 from Ch. 1 with the slight modification the energy E is replaced by

$$E_a(R) := \int_{\mathbb{R}} a(x)R^2 + (\partial_t R)^2 + (\partial_x R)^2 \, dx.$$

□

6.4 Description of pulse interaction

In order to extracr information about the interaction of pulses for the original equation, we can employ the exact same reasoning as for the Maxwell-Lorentz case in Sec. 5.4. Thus, we can conclude that with the definitions

$$\begin{aligned}
A_j^{(1)} &:= f_1(l_j, x)B_j^{(1)}, \\
A_j^{(2)} &:= f_1(l_j, x)B_j^{(2)} + if_1'(l_j, x)(\partial_1 B_j^{(1)}), \\
A_j^{(3)} &:= f_1(l_j, x)B_j^{(3)} + if_1'(l_j, x)(\partial_1 B_j^{(2)}) - \frac{1}{2}f_1''(l_j, x)(\partial_1^2 B_j^{(1)}), \\
\tilde{\Omega}_j^{(1)} &:= \Omega_j^{(1)} \\
\tilde{\Omega}_j^{(2)} &:= \Omega_j^{(2)} - if_1'(l_j, x)(\partial_1 \Omega_j^{(1)}) \\
\tilde{\Psi}_j^{(1)} &:= \Psi_j^{(1)},
\end{aligned}$$

we get again an extended two-pulse ansatz

$$\begin{aligned}
u^{an}(x,t) &= \sum_{j=1,2} \sum_{r=1}^{3} \varepsilon^r A_j^{(r)}(\underline{\tilde{X}}_j, \underline{T})E_j + M_{mixed,\alpha} + c.c. \\
E_j &= e^{il_j x + i\omega_1(l_j)t + i\sum_{r=1,2}\varepsilon^r \tilde{\Omega}_j^{(r)}(\underline{X}_{3-j}, \underline{T})}, \quad \underline{T} = \varepsilon^2 t, \\
\underline{\tilde{X}}_j &= \varepsilon(x + \omega_1'(l_j)t + \varepsilon\tilde{\Psi}_j^{(1)}(\underline{X}_{3-j}, \underline{T})).
\end{aligned}$$

So, again, we get the same qualitative interpretation of the interaction as described in Sec. 2.2 for the constant coefficient case. However, the equations are at little more complicated.

6.5 Chapter Summary

This chapter completes the discussion of pulse interaction for equations exhibiting a multibranch dispersion relation, by treating the case of a nonlinear wave equation with periodic coefficients whose dispersion relation has infinitely many branches. Let us briefly summarize the outcome:

Two pulses traveling on different carrier waves that satisfy the cubic Klein-Gordon equation with periodic coefficients

$$\partial_t^2 u = \partial_x^2 u + a(x)u + \gamma b(x)u^3$$

where $x, t, u(x,t) \in \mathbb{R}$ and a, b are real, smooth, 2π-periodic functions with $a(x) > \tilde{a} > 0$ and $b(x) > \tilde{b} > 0$ for $x \in \mathbb{R}$ and some constants $\tilde{a}, \tilde{b} \in \mathbb{R}$, can be approximated by

$$\begin{aligned} u^{an}(x,t) &= \sum_{r=1}^{3} \sum_{j=1,2} \varepsilon^r A_j^{(r)}(\underline{\tilde{X}}_j, \underline{T}) E_j + M_{mixed,\alpha} + c.c., \\ E_j &= e^{il_j x + i\omega_1(l_j)t + i\sum_{r=1,2} \varepsilon^r \tilde{\Omega}_j^{(r)}(\underline{X}_{3-j}, \underline{T})}, \quad \underline{T} = \varepsilon^2 t, \\ \underline{\tilde{X}}_j &= \varepsilon(x + \omega_1'(l_j)t + \varepsilon \tilde{\Psi}_j^{(1)}(\underline{X}_{3-j}, \underline{T})). \end{aligned}$$

if the envelopes $A_j^{(r)}$ fulfill

$$A_j^{(r)} := \sum_{m=0}^{r-1} \frac{\partial_1^{(m)} f_1(l_j, \cdot)}{m!} B_j^{(m)}$$

where the terms $B_j^{(r)}$ satisfy the recursively solvable modulation system for internal dynamics given by

$$\begin{aligned} \partial_2 B_j^{(1)} &= -i\frac{\omega_1''(l_j)}{2}\partial_1^2 B_j^{(1)} + 3\left(s(l_j)\int_0^{2\pi} b(x)|f_1(l_1,x)|^4\, dx\right)|B_j^{(1)}|^2, \\ \partial_2 B_j^{(2)} &= -i\frac{\omega_1''(l_j)}{2}\partial_1^2 B_j^{(2)} - \frac{\omega_1'''(l_j)}{6}\partial_1^3 B_j^{(1)} + \tilde{N}_{2,j}^{internal}(B_j^{(q)})_{q\le 2}, \\ \partial_2 B_j^{(3)} &= -i\frac{\omega_1''(l_j)}{2}\partial_1^2 B_j^{(3)} - \frac{\omega_1'''(l_j)}{6}\partial_1^3 B_j^{(2)} + \frac{\omega_1^{(4)}(l_j)}{4!}\partial_1^3 B_j^{(1)} + \tilde{N}_{3,j}^{internal}(B_j^{(q)})_{q\le 3}, \end{aligned}$$

and the interaction quantities solve the following equations:

The **carrier shift** $\tilde{\Omega}_j^{(1)}$ (see Fig. 2.3) has to satisfy the ordinary differential equation

$$\partial_1 \tilde{\Omega}_j^{(1)} = \frac{6i}{D_j}\left(s(l_j)\int_0^{2\pi} b(x)|f_1(l_1,x)|^2|f_1(l_2,x)|^2\, dx\right)|B_{3-j}^{(1)}|^2.$$

The **envelope position shift** $\tilde{\Psi}_j^{(1)}$ (see Fig. 2.4) has to satisfy the ordinary differential equation

$$\begin{aligned} \partial_1 \tilde{\Psi}_j^{(1)} = \frac{i}{D_j}[&6\left(s(l_j)\int_0^{2\pi} b(x)|f_1(l_2,x)|^2(\partial_l f_1(l_1,x))f_1(l_1,x)\, dx\right)|B_{3-j}^{(1)}|^2 \\ &+ 6\left(s'(l_j)\int_0^{2\pi} b(x)|f_1(l_1,x)|^2|f_1(l_2,x)|^2\, dx\right)\overline{B_{3-j}^{(1)}}B_j^{(1)} + \omega_1''(k_j)(\partial_1\Omega_j^{(1)})]. \end{aligned}$$

The **second order carrier shift and shape correction** are described by the real and imaginary part of $\tilde{\Omega}_j^{(2)}$ which is given by

$$\tilde{\Omega}_j^{(2)} := \Omega_j^{(2)} + iq_{11}'(l_j)(\partial_1\tilde{\Omega}_j^{(1)})$$

where $\Omega_j^{(2)}$ has to fulfill

$$\partial_1\Omega_j^{(2)} = \frac{i}{D_j}(\tilde{N}_{2,j}^{inter,\Omega_j^{(2)}}(B_{3-j}^{(q)})_{q\le l} + \frac{\omega_1''(k_j)}{2}(\partial_1^2\tilde{\Omega}_j^{(1)}) - i(\partial_2\tilde{\Omega}_j^{(1)})).$$

In conclusion, the interaction behavior of two pulses with different velocities is again exactly as described in great detail in Sec. 2.2. As already explained for the Maxwell-Lorentz system, this would have been very difficult (if possible at all) to reveal with conventional x-space methods as described in Ch. 2!

7 Applications in nonlinear fiber-optics and photonics

This final chapter indicates the consequences of our results for applications in photonics, a research area which is concerned with harnessing light to transport and process information. The fascinating thing about photonic devices is the fact that they provide the possibility to control and manipulate light, which usually seems a quite unmanageable force.

7.1 Wavelength-Division multiplexing

The first application that we want to discuss belongs to the field of optical communication (for a good introduction see [Has95], [Mol04], [Agr01]):
The transport of digital information through optical fibers can be realized by modulating a carrier wave as depicted in Fig. 7.1. To ensure high bit rates while avoiding bit errors one has to tailor carefully the shape and spacing of the pulses, since effects like dispersion and nonlinear interaction might distort the pulse train.
The Maxwell-Lorentz system that we have discussed in Ch. 5 is a widely accepted model for light propagation in a nonlinear material such as an optical fiber (see [SB07],[JMR96], [Blo65]).
In the present work we have analyzed specific solutions for this system — *multipulses*. This is motivated as follows:

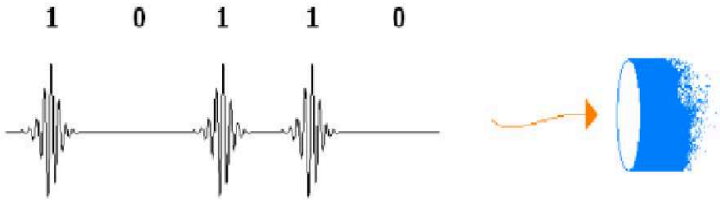

Figure 7.1: Transport of the digital code 10110 by a modulated optical carrier wave through an optical fiber

The performance of a single fiber can be improved by simultaneously using several carrier waves, which are in this context often referred to as “channels” (see Fig. 7.2). This technique is called “Wavelength-Division Multiplexing" (WDM). Again, due to the nonlinearity, pulses belonging to different channels interact, leading to further difficulties in adjusting the spacing between pulses.
The two types of disturbances that a WDM system can suffer from are

- **Intrachannel interaction** which describes the coupling effects between pulses belonging to the *same* channel and
- **Interchannel interaction** which describes the coupling between pulses of *different* channels.

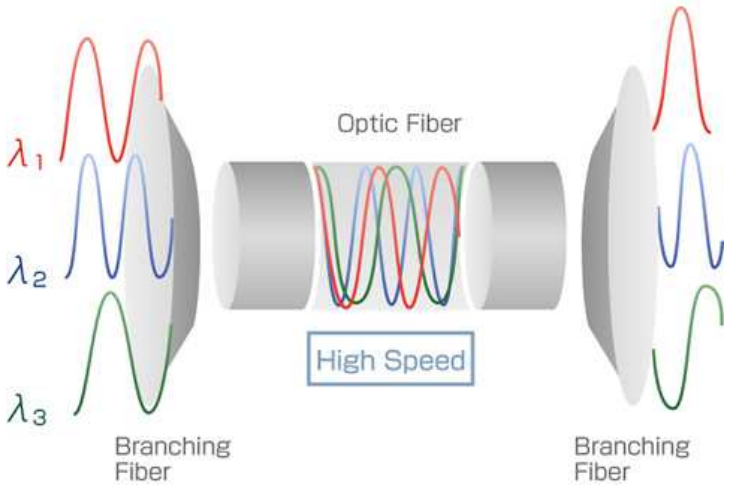

Figure 7.2: Multiplexing of several carrier waves characterized by their wavelengths (source [KVH])

In Ch. 2 we gave a detailed analysis of both interaction types and merged them to describe the dynamics of a multipulse, which would correspond to the setting encountered for a WDM system. Note that a multipulse as depicted in Fig. 2.6 describes the worst case scenario, where only $'1's$ are being sent.
Concerning intrachannel interaction, we have managed to derive a condition involving the spacing of the pulses and their localization strength which ensures that the pulse train remains undistorted for a very long time period.
For the interaction between pulses of different channels we derived explicit formulas for the interaction effects of leading orders, namely a shift of the carrier and envelope and a shape deformation. Furthermore, we improved the bound on the envelope shift of $\mathcal{O}(1)$ (that one gets with a pure qualitative analysis as illustrated in Sec. 2.2.1) to $\mathcal{O}(\varepsilon)$. This indicates that the interaction is indeed very weak and the pulse spacing could be chosen smaller as first anticipated.
In order to further optimize the WDM system, one might wish to set $N = \mathcal{O}(1/\varepsilon^{\alpha_N})$ with α_N large, i.e. to have a large number of different channels. However, as already pointed out in Sec. 2.3.1 this would lead to an increased envelope shift, and, so, the spacing between pulses would have to be chosen larger. Hence, we see that there is a trade-off between the wish to increase the number of channels and the wish to pack the pulses as dense as possible.

Both types of interactions have already been analyzed before, but in a different way (see [Agr01], [Has95],[Mol04]).

7.2 Detection of standing light pulses

The second application that could benefit from our results comes from the field of "Photonics", a relatively young area of research that has gained immensely in importance over the last decade.

One of the main objectives of "Photonics" is the creation of devices that provide the possibility to guide and manipulate light ("photons") such that it can be used to process and store information — the ultimate goal being ultrahigh-speed all-optical signal processing on a single photonic chip (see [CUD] for a very nice animation).

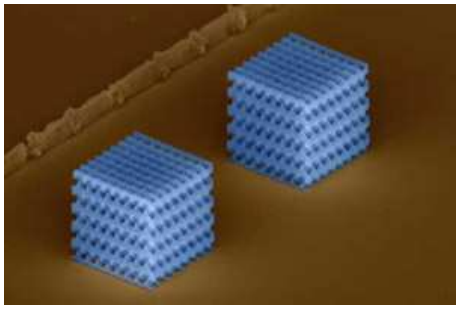

Figure 7.3: "Photonic Crystal" (source [KIT])

However, it is hard to find materials in nature that allow a complete control of light. Therefore, many efforts have been made to artificially design new materials that are suited for photonic devices. The most promising among these new materials are so-called "Photonic Crystals" which consist of a periodic arrangement of materials with different dielectric properties (see [BvFL$^+$07] and [Joa08] for a good introduction). One of the most intriguing features of "Photonic Crystals" is that they are believed to support "standing light pulses", i.e. a pulse with group velocity zero. To be precise, the theoretical prediction of this phenomenon has been already confirmed by numerical simulations, but it has, to this day, not been observed in experiments. Hence, the question arises how one can detect such "standing pulses". It has been proposed in [TPB04] and [TNPB06] that this detection can be realized by making another pulse collide with the standing one. If the material has nonlinear features, the collision should leave its traces on the pulses. The hope is that, on the one hand, the interaction is weak enough to ensure that the moving pulse will not be destroyed, but, on the other hand, strong enough to ensure that the parameters of the moving pulse have changed to an amount that can actually be measured. This is exactly the scenario that we have described for a two-pulse solution in Ch. 6, where we treated a nonlinear wave equation with periodic coefficients. As a consequence, the methods presented in this work can be employed to derive a modulation system like (6.2.6)-(6.2.11) which could be used to engineer a detection mechanism for "standing light pulses"!

"... The optical soliton in fibers presents a beautiful example in which an abstract mathematical concept has produced a large impact on the real world of high technologies."

[Has95]

Bibliography

[Ada07] John J. F. Adams, Robert A. ; Fournier. *Sobolev spaces*. Pure and applied mathematics ; 140. Academic Press, Amsterdam, 2. ed., repr. edition, 2007.

[Agr01] Govind P. Agrawal. *Nonlinear fiber optics*. Optics and photonics. Academic Press, San Diego [u.a.], 3. ed. edition, 2001.

[Blo65] Nicolaas Bloembergen. *Nonlinear optics : a lecture note and reprint volume*. Frontiers in physics ; [21]. Benjamin, New York [u.a.], 1965.

[Bou99] Jean Bourgain. *Global solutions of nonlinear Schrödinger equations*. Colloquium publications / American Mathematical Society ; 46. American Mathematical Society, Providence, R.I., 1999.

[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.

[BvFL+07] K. Busch, G. von Freymann, S. Linden, S. F. Mingaleev, L. Tkeshelashvili, and M. Wegener. Periodic nanostructures for photonics. *Physics Reports*, 444:101, 6 2007.

[CBCSU08] Martina Chirilus-Bruckner, Christopher Chong, Guido Schneider, and Hannes 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.

[CKS95] Walter Craig, Thomas Kappeler, and Walter Strauss. Microlocal dispersive smoothing for the Schrödinger equation. *Comm. Pure Appl. Math.*, 48(8):769–860, 1995.

[CUD] CUDOS (Centre for Ultrahigh bandwidth Devices for Optical Systems). http://www.physics.usyd.edu.au/cudos/index.html.

[Has95] Yuji Hasegawa, Akira ; Kodama. *Solitons in optical communications*. Oxford series in optical and imaging sciences ; 7. Clarendon Press, Oxford, 1995. : £ 60.00.

[JMR96] J. L. Joly, G. Metivier, and J. Rauch. Global solvability of the anharmonic oscillator model from nonlinear optics. *SIAM J. Math. Anal.*, 27(4):905–913, 1996.

[Joa08] John D. Joannopoulos, editor. *Photonic crystals : molding the flow of light*. Princeton Univ. Press, Princeton, NJ, 2. ed. edition, 2008.

[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.

[KIT] KIT (Karlsruhe Institute of Technology). http://www.kit.edu.

[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.

[KVH] KVH Co., Ltd. http://www.kvh.co.jp/en.

[Mol04] Alan C. Moloney, Jerome V. ; Newell. *Nonlinear optics*. The advenced book program. Westview, Boulder, Colo. [u.a.], 2004.

[OY74] Masayuki Oikawa and Nobuo Yajima. A perturbation approach to nonlinear systems. ii. interaction of nonlinear modulated waves. *Journal of the Physical Society of Japan*, 37(2):486–496, 1974.

[Pac98] B. G. Pachpatte. *Inequalities for differential and integral equations*, volume 197 of *Mathematics in Science and Engineering*. Academic Press Inc., San Diego, CA, 1998.

[Pol04] Valentin F. Poljanin, Andrej D. ; Zajcev. *Handbook of nonlinear partial differential equations*. Chapman & Hall/CRC, Boca Raton, Fla. [u.a.], 2004. Includes bibliographical references and index.

[SB07] M. P Sørensen and M. Brio. The maxwell-lorentz model for optical pulses. *The European Physical Journal - Special Topics*, 147(1):254–265, 2007.

[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.

[Sul99] Pierre-Louis Sulem, Catherine ; Sulem. *Nonlinear Schrödinger equation : self-focusing and wave collapse*. Applied mathematical sciences ; 139. Springer, New York, 1999. Includes bibliographical references and index.

[TNPB06] L. Tkeshelashvili, J. Niegemann, S. Pereira, and K. Busch. Nonlinear wave interaction in photonic band gap materials. *Photonics and Nanostructures - Fundamentals and Applications*, 4:75, 2006.

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