Commun. Math. Phys. 278, 329–384 (2008)
Digital Object Identifier (DOI) 10.1007/s00220-007-0406-0

Communications in
**Mathematical
Physics**

Wavepacket Preservation Under Nonlinear Evolution

A. Babin, A. Figotin

Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA.
E-mail: ababine@uci.edu; afigotin@uci.edu

Received: 27 July 2006 / Accepted: 24 August 2007
Published online: 8 January 2008 –

Abstract: We study nonlinear systems of hyperbolic PDE's in $\mathbb{R}^d$, the hyperbolicity is understood in a wider sense, namely multiple roots of the characteristic equation are allowed and dispersive equations are permitted. They describe wave propagation in dispersive nonlinear media such as, for example, electromagnetic waves in nonlinear photonic crystals. The initial data is assumed to be a finite sum of wavepackets referred to as a *multi-wavepacket*. The wavepackets and the medium nonlinearity are characterized by two principal small parameters β and ϱ where: (i) $\frac{1}{\beta}$ is a factor describing spatial extension of involved wavepackets; (ii) $\frac{1}{\varrho}$ is a factor describing the relative magnitude of the linear part of the evolution equation compared to its nonlinearity. A key element in our approach is a proper definition of a wavepacket. Remarkably, the introduced definition has a flexibility sufficient for a wavepacket to preserve its defining properties under a general nonlinear evolution for long times. In particular, the corresponding wave vectors and the band numbers of involved wavepackets are "conserved quantities". We also prove that the evolution of a multi-wavepacket is described with high accuracy by a properly constructed system of envelope equations with a universal nonlinearity. The universal nonlinearity is obtained by a time averaging applied to the original nonlinearity, in simpler cases the averaged system turns into a system of Nonlinear Schrodinger equations.

1. Introduction

The underlying physical subject of this work is propagation of a *multi-wavepacket* (a finite system of wavepackets) in a spatially dispersive and nonlinear medium, and we are particularly interested in electromagnetic waves propagation in nonlinear photonic crystals, see [4–7,55,56,58] and references therein, with the nonlinear optics constitutive relations, [12,15, Sects. 1,2, 42,48]. The mathematical subject of interest is the following general nonlinear evolutionary system

$$\partial_\tau \mathbf{U} = -\frac{\mathrm{i}}{\varrho}\mathbf{L}(-\mathrm{i}\nabla)\mathbf{U} + \mathbf{F}(\mathbf{U}),\ \ \mathbf{U}(\mathbf{r},\tau)|_{\tau=0} = \mathbf{h}(\mathbf{r}),\ \ \mathbf{r}\in\mathbb{R}^d, \tag{1}$$

where (i) $\mathbf{U} = \mathbf{U}(\mathbf{r},\tau)$, $\mathbf{r}\in\mathbb{R}^d$, $\mathbf{U}\in\mathbb{C}^{2J}$ is a $2J$ dimensional vector; (ii) $\mathbf{L}(-\mathrm{i}\nabla)$ is a linear self-adjoint differential (pseudodifferential) operator with constant coefficients with the symbol $\mathbf{L}(\mathbf{k})$, which is a Hermitian $2J\times 2J$ matrix; (iii) $\mathbf{F}$ is a polynomial nonlinearity such that $\mathbf{F}(\mathbf{0}) = \mathbf{0}$, $\mathbf{F}'(\mathbf{0}) = \mathbf{0}$ and $\mathbf{F}(\mathbf{U})$ is translation-invariant, i.e. if $T_{\mathbf{a}}\mathbf{U}(\mathbf{r}) = \mathbf{U}(\mathbf{r}+\mathbf{a})$ for $\mathbf{a}\in\mathbb{R}^d$ then $\mathbf{F}(T_{\mathbf{a}}\mathbf{U}) = T_{\mathbf{a}}\mathbf{F}(\mathbf{U})$; (iv) $\mathbf{h} = \mathbf{h}(\mathbf{r})$ is assumed to be the sum of a finite number of wavepackets $\mathbf{h}_l$, $l = 1,\ldots,N$; (v) $\varrho > 0$ is a *small parameter*. In the case of nonlinear photonic crystals the components of the vector field $\mathbf{U}(\mathbf{r})$ are the modal amplitudes of the electromagnetic field and the nonlinearity $\mathbf{F}(\mathbf{U})$ is constructed from the nonlinear medium polarization in the adiabatic approximation, [15, Sects. 2.4.2]. The systems of the form (1) also describe as a particular case well-known equations, namely: complexification of the Nonlinear Schrodinger equation; coupled envelope equations which arise in nonlinear birefringent optical media, [41, Sect. 2i]; nonlinear Klein-Gordon and Sine-Gordon equations [61, Sect. 14.1,43, Sect. 5.8.3,44, Sect. 9.6]. Such equations appear in a number of physical problems: elementary particles, dislocations in crystals, propagation of Bloch's domain walls in the theory of ferromagnetism, self-induced transparency in nonlinear optics, the propagation of magnetic flux quanta in long Josephson transmission lines. Significance and importance of wavepacket solutions from both physical and mathematical points of view is discussed in [4–7,41, Sect. 2, 55,58].

There are numerous problems involving small parameters only in the initial data which can be reduced to the form (1), for instance, problems with high frequency initial data or small initial data with consequent evolution on long time intervals (see Sect. 3 for details).

We study the nonlinear evolution equation (1) on a finite time interval

$$0\le\tau\le\tau_*,\ \text{ where } \tau_* > 0 \text{ is a fixed number.} \tag{2}$$

The time τ_* may depend on the L^∞ norm of the initial data $\mathbf{h}$ but, importantly, τ_* *does not depend on* ϱ. *We consider classes of initial data such that wave evolution governed by* (1) *is significantly nonlinear on the time interval* $[0,\tau_*]$ *and the effect of the nonlinearity* $F(\mathbf{U})$ *does not vanish as* $\varrho\to 0$.

Since both the linear operator $\mathbf{L}(-\mathrm{i}\nabla)$ and the nonlinearity $\mathbf{F}(\mathbf{U})$ are translation invariant, it is natural and convenient to recast the evolution equation (1) by applying to it the Fourier transform with respect to the space variables $\mathbf{r}$, namely

$$\partial_\tau\hat{\mathbf{U}}(\mathbf{k}) = -\frac{\mathrm{i}}{\varrho}\mathbf{L}(\mathbf{k})\hat{\mathbf{U}}(\mathbf{k}) + \hat{\mathbf{F}}\left(\hat{\mathbf{U}}\right)(\mathbf{k}),\ \ \hat{\mathbf{U}}(\mathbf{k})\Big|_{\tau=0} = \hat{\mathbf{h}}(\mathbf{k}), \tag{3}$$

where $\hat{\mathbf{U}}(\mathbf{k})$ is the Fourier transform of $\mathbf{U}(\mathbf{r})$, i.e.

$$\hat{\mathbf{U}}(\mathbf{k}) = \int_{\mathbb{R}^d}\mathbf{U}(\mathbf{r})\,\mathrm{e}^{-\mathrm{i}\mathbf{r}\cdot\mathbf{k}}\mathrm{d}\mathbf{r},\ \ \mathbf{U}(\mathbf{r}) = (2\pi)^{-d}\int_{\mathbb{R}^d}\hat{\mathbf{U}}(\mathbf{k})\,\mathrm{e}^{\mathrm{i}\mathbf{r}\cdot\mathbf{k}}\mathrm{d}\mathbf{r},\ \ \text{where } \mathbf{r},\mathbf{k}\in\mathbb{R}^d, \tag{4}$$

and $\hat{\mathbf{F}}$ is the Fourier form of the nonlinear operator $\mathbf{F}(\mathbf{U})$ involving convolutions.

The nonlinear evolution equations (1), (3) are commonly interpreted as describing wave propagation in a nonlinear medium. We assume that the linear part $\mathbf{L}(\mathbf{k})$ is a $2J\times 2J$ Hermitian matrix with eigenvalues $\omega_{n,\zeta}(\mathbf{k})$ and eigenvectors $\mathbf{g}_{n,\zeta}(\mathbf{k})$ satisfying

$$\mathbf{L}(\mathbf{k})\,\mathbf{g}_{n,\zeta}(\mathbf{k}) = \omega_{n,\zeta}(\mathbf{k})\,\mathbf{g}_{n,\zeta}(\mathbf{k}),\ \ \zeta=\pm,\ \omega_{n,+}(\mathbf{k})\ge 0,\ \omega_{n,-}(\mathbf{k})\le 0,\ n = 1,\ldots,J, \tag{5}$$

where $\omega_{n,\zeta}(\mathbf{k})$ are real-valued, continuous for all non-singular $\mathbf{k}$ functions, and vectors $\mathbf{g}_{n,\zeta}(\mathbf{k}) \in \mathbb{C}^{2J}$ have unit length in the standard Euclidean norm. The functions $\omega_{n,\zeta}(\mathbf{k})$, $n = 1, \ldots, J$, are called *dispersion relations* between the frequency ω and the *wave-vector* $\mathbf{k}$ with n being the *band number*. We assume that the eigenvalues are naturally ordered by

$$\omega_{J,+}(\mathbf{k}) \geq \ldots \geq \omega_{1,+}(\mathbf{k}) \geq 0 \geq \omega_{1,-}(\mathbf{k}) \geq \ldots \geq \omega_{J,-}(\mathbf{k}), \tag{6}$$

and for *almost every* $\mathbf{k}$ (with respect to the standard Lebesgue measure) the eigenvalues are distinct and, consequently, the above inequalities become strict. Importantly, we also assume the following *diagonal symmetry* condition

$$\omega_{n,-\zeta}(-\mathbf{k}) = -\omega_{n,\zeta}(\mathbf{k}), \ \zeta = \pm, \ n = 1, \ldots, J, \tag{7}$$

which is naturally present in many physical problems (see also Remark 14 below), and is a fundamental condition imposed on the matrix $\mathbf{L}(\mathbf{k})$. In addition to that in many examples we also have

$$\mathbf{g}_{n,\zeta}(\mathbf{k}) = \mathbf{g}^{*}_{n,-\zeta}(-\mathbf{k}), \ \text{where } z^{*} \text{ is complex conjugate to } z. \tag{8}$$

Very often we will use the following abbreviation:

$$\omega_{n,+}(\mathbf{k}) = \omega_n(\mathbf{k}). \tag{9}$$

From (7) we obtain

$$\omega_{n,-}(\mathbf{k}) = -\omega_n(-\mathbf{k}), \ \omega_{n,\zeta}(\mathbf{k}) = \zeta\omega_n(\zeta\mathbf{k}), \ \zeta = \pm. \tag{10}$$

We also will often use the orthogonal projection $\Pi_{n,\zeta}(\mathbf{k})$ in $\mathbb{C}^{2J}$ onto the complex line defined by the eigenvector $\mathbf{g}_{n,\zeta}(\mathbf{k})$, namely

$$\Pi_{n,\zeta}(\mathbf{k})\,\hat{\mathbf{u}}(\mathbf{k}) = \tilde{u}_{n,\zeta}(\mathbf{k})\,\mathbf{g}_{n,\zeta}(\mathbf{k}) = \hat{\mathbf{u}}_{n,\zeta}(\mathbf{k}), \ n = 1, \ldots, J, \ \zeta = \pm. \tag{11}$$

As it is indicated by the title of this paper we study the nonlinear problem (1) for initial data $\hat{\mathbf{h}}$ in the form of a properly defined *wavepacket* or, more generally, a sum of wavepackets which we refer to as *multi-wavepacket*. The simplest example of a wavepacket $\mathbf{w}$ is provided by the following formula:

$$\mathbf{w}(\mathbf{r}, \beta) = \Phi_{+}(\beta\mathbf{r})\,\mathrm{e}^{\mathrm{i}\mathbf{k}_{*}\cdot\mathbf{r}}\mathbf{g}_{n,+}(\mathbf{k}_{*}), \ \mathbf{r} \in \mathbb{R}^{d}, \tag{12}$$

where $\mathbf{k}_{*} \in \mathbb{R}^{d}$ is a *wavepacket wave vector*, n is *band number*, and $\beta > 0$ is a small parameter. We refer to the pair $(n, \mathbf{k}_{*})$ in (12) as the *wavepacket nk-pair*. Observe that the space extension of the wavepacket $\mathbf{w}(\mathbf{r}, \beta)$ is proportional to β^{-1} and it is large for small β. Notice also that if $\beta \to 0$ the wavepacket $\mathbf{w}(\mathbf{r}, \beta)$ as in (12) tends, up to a constant factor, to the elementary eigenmode $\mathrm{e}^{\mathrm{i}\mathbf{k}_{*}\cdot\mathbf{r}}\mathbf{g}_{n,\zeta}(\mathbf{k}_{*})$ of the operator $\mathbf{L}(-\mathrm{i}\nabla)$ with the corresponding eigenvalue $\omega_{n,\zeta}(\mathbf{k}_{*})$. We refer to wavepackets of the simple form (12) as *simple wavepackets* to underline the very special way the parameter β enters its representation. The function $\Phi_{\zeta}(\mathbf{r})$, which we call *the wavepacket envelope*, describes its shape and it can be any scalar complex-valued regular enough function, for example a function from Schwartz space. Importantly, as $\beta \to 0$ the L^{∞} *norm of a wavepacket* (12) *remains constant, and, hence, nonlinear effects in* (1) *remain strong*.

Evolution of wavepackets in problems which can be reduced to the form (1) were studied for a variety of equations in numerous physical and mathematical papers, mostly by asymptotic expansions with respect to a single small parameter similar to β, see 10,13,18,20,23,29,30,38,47,50,51] and references therein. We are interested in general properties of evolutionary systems of the form (1) with wavepacket initial data which hold for a wide class of nonlinearities and all values of the space dimensions d of the number $2J$ of the system components. Our approach is not based on asymptotic expansions but involves the two small parameters β and ϱ with mild constraints on their relative smallness. The constraints can be expressed either in the form of certain inequalities or equalities, and a possible simple form of such a constraint can be a power law

$$\beta = C\varrho^{\varkappa} \text{ where } C > 0 \text{ and } \varkappa > 0 \text{ are arbitrary constants.} \tag{13}$$

Of course, general features of wavepacket evolution are independent of particular values of the constant C. In addition to that, some fundamental properties such as wavepacket invariance, are also totally independent of the particular choice of the values of $\varkappa$ in (13), whereas other properties are independent of $\varkappa$ as it varies in certain intervals. For instance, dispersion effects are dominant for $\varkappa < 1/2$, whereas the wavepacket superposition principle of 7] holds for $\varkappa < 1$.

The qualitative picture of wavepacket evolution dependence on small β and ϱ is as follows. The parameter β enters problem (1) through the multi-wavepacket initial data $\mathbf{h}(\mathbf{r}, \beta)$, whereas ϱ enters it through the factor $\frac{1}{\varrho}$ before the linear part. Evidently the factor $\frac{1}{\varrho}$ determines the relative magnitude of the linear part compared to the nonlinearity and since $\frac{1}{\varrho}$ is large, one expects the linear part to provide an important input into solutions properties. This input includes, in particular, *a key role of eigenmodes and eigenfrequencies* (*dispersion relations*) in expressing the nonlinear evolution. Importantly, in many cases of interest though $\frac{1}{\varrho}$ is large, *nonlinear phenomena are significant* and this is the case when $\beta \le C\varrho^{1/2}$. More precisely, if $\beta \le C\varrho^{1/2}$ then, as in the case of finite-dimensional nonlinear ODE evolutionary systems, the large values of $\frac{1}{\varrho}$ lead to a well defined solution factorization into the fast (high frequency) and the slow (low frequency) components. The interplay between the fast and slow components is also similar to the ODE case, namely, the nonlinear evolution is associated primarily with the slow component governed by a nonlinear equation obtained from the original one by a certain canonical time averaging procedure. Our further analysis of the above mentioned interplay shows the following. Firstly, the *linear superposition principle holds*, 7], that is if $\varkappa < 1$ is as in (13) and the initial data is a sum of generic wavepackets then the solution is the sum of the solutions for single involved wavepackets with precision $\frac{\varrho}{\beta^{1+\epsilon}}$ with arbitrary small ϵ. Secondly, properly defined *wavepackets and their linear combinations are preserved* under the nonlinear evolution (1), which is a subject of this paper.

In the light of the above discussion we introduce the *slow variable* $\hat{\mathbf{u}}(\mathbf{k}, \tau)$ by the formula

$$\hat{\mathbf{U}}(\mathbf{k}, \tau) = \mathrm{e}^{-\frac{\mathrm{i}\tau}{\varrho}\mathbf{L}(\mathbf{k})}\hat{\mathbf{u}}(\mathbf{k}, \tau), \tag{14}$$

and recast Eq. (3) for it as follows:

$$\partial_\tau \hat{\mathbf{u}} = \mathrm{e}^{\frac{\mathrm{i}\tau}{\varrho}\mathbf{L}}\hat{\mathbf{F}}\left(\mathrm{e}^{\frac{-\mathrm{i}\tau}{\varrho}\mathbf{L}}\hat{\mathbf{u}}\right), \quad \hat{\mathbf{u}}\big|_{\tau=0} = \hat{\mathbf{h}}. \tag{15}$$

Then we obtain an integral form of (15) by integrating it with respect to τ:

$$\hat{\mathbf{u}} = \mathcal{F}\left(\hat{\mathbf{u}}\right) + \hat{\mathbf{h}}, \quad \mathcal{F}\left(\hat{\mathbf{u}}\right) = \mathcal{F}\left(\varrho\right)\left(\hat{\mathbf{u}}\right) = \int_0^\tau \mathrm{e}^{\frac{\mathrm{i}\tau'}{\varrho}\mathbf{L}}\hat{\mathbf{F}}\left(\mathrm{e}^{\frac{-\mathrm{i}\tau'}{\varrho}\mathbf{L}}\hat{\mathbf{u}}\left(\tau'\right)\right)\mathrm{d}\tau' \tag{16}$$

with an explicitly defined nonlinear polynomial integral operator $\mathcal{F}(\varrho)$, which depends on the parameter ϱ. This operator is bounded uniformly with respect to ϱ in the Banach space $E = C\left([0, \tau_*], L^1\right)$ of functions $\hat{\mathbf{v}}(\mathbf{k}, \tau)$, $0 \leq \tau \leq \tau_*$, with the norm

$$\left\| \hat{\mathbf{v}}(\mathbf{k}, \tau) \right\|_E = \left\| \hat{\mathbf{v}}(\mathbf{k}, \tau) \right\|_{C([0,\tau_*],L^1)} = \sup_{0 \leq \tau \leq \tau_*} \int_{\mathbb{R}^d} \left| \hat{\mathbf{v}}(\mathbf{k}, \tau) \right| \, \mathrm{d}\mathbf{k}, \tag{17}$$

where L^1 is the Lebesgue space of functions $\hat{\mathbf{v}}(\mathbf{k})$ with the standard norm

$$\left\| \hat{\mathbf{v}}(\cdot) \right\|_{L^1} = \int_{\mathbb{R}^d} \left| \hat{\mathbf{v}}(\mathbf{k}) \right| \, \mathrm{d}\mathbf{k}. \tag{18}$$

Sometimes we use more general weighted spaces $L^{1,a}$ with the norm

$$\left\| \hat{\mathbf{v}} \right\|_{L^{1,a}} = \int_{\mathbb{R}^d} (1 + |\mathbf{k}|)^a \left| \hat{\mathbf{v}}(\mathbf{k}) \right| \, \mathrm{d}\mathbf{k}, \ a \geq 0. \tag{19}$$

A rather elementary existence and uniqueness theorem (Theorem 29) implies that for a small and, importantly, independent of ϱ constant $\tau_* > 0$ this equation has a unique solution

$$\hat{\mathbf{u}}(\tau) = \mathcal{G}\left(\mathcal{F}(\varrho), \hat{\mathbf{h}}\right)(\tau), \ \tau \in [0, \tau_*], \ \hat{\mathbf{u}} \in C^1\left([0, \tau_*], L^1\right), \tag{20}$$

where $\mathcal{G}$ denotes the *solution operator* for Eq. (16), the operator depends on operator $\mathcal{F}(\varrho)$, which itself depends on the parameter ϱ. If $\hat{\mathbf{u}}(\mathbf{k}, \tau)$ is a solution to Eq. (16) we call the function $\mathbf{U}(\mathbf{r}, \tau)$ defined by (14), (4) an *F-solution* to Eq. (1). We denote by $\hat{L}^1$ the space of functions $\mathbf{V}(\mathbf{r})$ such that their Fourier transform $\hat{\mathbf{V}}(\mathbf{k})$ belongs to L^1, and define $\|\mathbf{V}\|_{\hat{L}^1} = \left\| \hat{\mathbf{V}} \right\|_{L^1}$. Since

$$\|\mathbf{V}\|_{L^\infty} \leq (2\pi)^{-d} \left\| \hat{\mathbf{V}} \right\|_{L^1} \text{ and } \hat{L}^1 \subset L^\infty, \tag{21}$$

F-solutions to (1) belong to $C^1\left([0, \tau_*], \hat{L}^1\right) \subset C^1\left([0, \tau_*], L^\infty\right)$.

We would like to define wavepackets in a form which explicitly allows them to be real valued. This is accomplished based on the symmetry (7) of the dispersion relations by introduction of a *doublet wavepacket*

$$\mathbf{w}(\mathbf{r}, \beta) = \Phi_+(\beta\mathbf{r}) \, \mathrm{e}^{\mathrm{i}\mathbf{k}_*\cdot\mathbf{r}} \mathbf{g}_{n,+}(\mathbf{k}_*) + \Phi_-(\beta\mathbf{r}) \, \mathrm{e}^{-\mathrm{i}\mathbf{k}_*\cdot\mathbf{r}} \mathbf{g}_{n,-}(-\mathbf{k}_*). \tag{22}$$

Such a wavepacket is real if $\Phi_-(\mathbf{r})$, $\mathbf{g}_{n,-}(-\mathbf{k}_*)$ is complex conjugate to $\Phi_+(\mathbf{r})$, $\mathbf{g}_{n,+}(\mathbf{k}_*)$, i.e. if

$$\Phi_-(\mathbf{r}) = \Phi_+^*(\mathbf{r}), \ \mathbf{g}_{n,+}(\mathbf{k}_*) = \mathbf{g}_{n,-}(-\mathbf{k}_*)^*. \tag{23}$$

Considering wavepackets with nk-pair $(n, \mathbf{k}_*)$ we usually mean doublet ones as in (22), but sometimes Φ_+ or Φ_- may be zero producing (12).

To identify characteristic properties of a wavepacket suitable for our needs, let us look at the Fourier transform $\hat{\mathbf{w}}(\mathbf{k}, \beta)$ of an elementary wavepacket $\mathbf{w}(\mathbf{r}, \beta)$ defined by (12), that is

$$\hat{\mathbf{w}}(\mathbf{k}, \beta) = \beta^{-d} \hat{\Phi}\left(\beta^{-1}(\mathbf{k} - \mathbf{k}_*)\right) \mathbf{g}_{n,\zeta}(\mathbf{k}_*). \tag{24}$$

We call such $\hat{\mathbf{w}}(\mathbf{k}, \beta)$ a wavepacket too, obviously it possesses the following properties: (i) its L^1 norm is bounded (in fact, constant), uniformly in $\beta \to 0$; (ii) for every $\epsilon > 0$

the value $\hat{\mathbf{w}}(\mathbf{k}, \beta) \to 0$ for every $\mathbf{k}$ outside a $\beta^{1-\epsilon}$-neighborhood of $\mathbf{k}_*$, and the convergence is faster than any power of β if Φ is a Schwartz function. To explicitly interpret the last property we introduce a *cutoff function* $\Psi(\boldsymbol{\eta})$,

$$\Psi(\boldsymbol{\eta}) = 1 \text{ for } |\boldsymbol{\eta}| \le 1, \ \Psi(\boldsymbol{\eta}) = 0 \text{ for } |\boldsymbol{\eta}| > 1, \tag{25}$$

together with its shifted/rescaled modification

$$\Psi(\mathbf{k}; \mathbf{k}_*) = \Psi\left(\mathbf{k}; \mathbf{k}_*, \beta^{1-\epsilon}\right) = \Psi\left(\beta^{-(1-\epsilon)}(\mathbf{k} - \mathbf{k}_*)\right). \tag{26}$$

If in an elementary wavepacket $\mathbf{w}(\mathbf{r}, \beta)$ defined by (24) $\Phi_\zeta(\mathbf{r})$ is a Schwartz function then

$$\left\|\left(1 - \Psi\left(\cdot, \mathbf{k}_*, \beta^{1-\epsilon}\right)\right)\hat{\mathbf{w}}(\cdot, \beta)\right\| \le C_{\epsilon,s}\beta^s, \ 0 < \beta \le 1,$$

which holds for arbitrarily small $\epsilon > 0$ and arbitrarily large $s > 0$. Based on the above discussion we give the following definition of a wavepacket which is a minor variation of 7, Def. 8].

Definition 1 (Single-band wavepacket). *Let* $0 < \epsilon < 1$ *be a fixed number. For a given band number* $n \in \{1, \ldots, J\}$ *and a wavevector* $\mathbf{k}_* \in \mathbb{R}^d$*, a function* $\hat{\mathbf{h}}(\beta, \mathbf{k})$ *is called a* **wavepacket with** nk**-pair** $(n, \mathbf{k}_*)$ *and the degree of regularity* $s > 0$ *if there exists such* $\beta_0 > 0$ *that for* $\beta < \beta_0$ *the following conditions are satisfied:* (i) $\hat{\mathbf{h}}(\beta, \mathbf{k})$ *is* L^1*-bounded uniformly in* β*, i.e.*

$$\left\|\hat{\mathbf{h}}(\beta, \cdot)\right\|_{L^1} \le C, \ 0 < \beta < \beta_0 \text{ for some } C > 0; \tag{27}$$

(ii) $\hat{\mathbf{h}}(\beta, \mathbf{k})$ *has the following structure:*

$$\hat{\mathbf{h}}(\beta, \mathbf{k}) = \hat{\mathbf{h}}_-(\beta, \mathbf{k}) + \hat{\mathbf{h}}_+(\beta, \mathbf{k}) + \hat{D}_h, \ 0 < \beta < \beta_0, \text{ where} \tag{28}$$

$$\hat{\mathbf{h}}_\zeta(\beta, \mathbf{k}) = \Psi\left(\mathbf{k}, \zeta\mathbf{k}_*, \beta^{1-\epsilon}\right)\Pi_{n,\zeta}(\mathbf{k})\hat{\mathbf{h}}_\zeta(\beta, \mathbf{k}), \ \zeta = \pm, \tag{29}$$

with $\Psi\left(\cdot, \zeta\mathbf{k}_*, \beta^{1-\epsilon}\right)$ *defined by* (26) *and* $\hat{D}_h$ *satisfying the following* **tail estimate***:*

$$\left\|\hat{D}_h\right\|_{L^1} \le C'\beta^s, \ 0 < \beta < \beta_0 \text{ for some } C' > 0. \tag{30}$$

The inverse Fourier transform $\mathbf{h}(\beta, \mathbf{r})$ *of a wavepacket* $\hat{\mathbf{h}}(\beta, \mathbf{k})$ *is also called a wavepacket.*

Point (ii) of the above definition means that the wavepacket $\hat{\mathbf{h}}(\beta, \mathbf{k})$ is composed of two functions $\hat{\mathbf{h}}_\zeta(\beta, \mathbf{k})$, $\zeta = \pm$, which take values the in the n^{th} band eigenspace of $\mathbf{L}(\mathbf{k})$ and are localized near $\zeta\mathbf{k}_*$, where $(n, \mathbf{k}_*)$ is the nk-pair of the wavepacket. The number β_0 usually is small and may depend on a wavepacket.

Evidently, if a wavepacket has the degree of regularity s, it also has a smaller degree of regularity $s' \le s$ with the same ϵ. Observe that the degree of regularity s is related to the smoothness of $\Phi_\zeta(\mathbf{r})$ in (12) so that the higher the smoothness is the higher $\frac{s}{\epsilon}$ can be taken. Namely, if $\hat{\Phi}_\zeta \in L^{1,a}$ then one can take any $\frac{s}{\epsilon} < a$, see Lemma 52 below. For example, if in the elementary wavepacket $\mathbf{w}(\mathbf{r}, \beta)$ defined by (12) $\Phi_\zeta(\mathbf{r})$ is a Schwartz function then it has arbitrarily large degree of regularity.

Remarkably it turns out that wavepackets satisfying Definition 1 preserve their defining properties under nonlinear evolution. It is remarkable, in particular, since it is well-known that determination of classes of solutions which preserve their form under generic nonlinear evolution usually leads to infinite expansions, such as multi-scale expansions, power expansions, modal expansions, etc. with serious difficulties in establishing the convergence. Such expansions often are formally invariant, but they involve infinitely many rather complex terms and establishing the convergence is a very hard problem indeed if there is any convergence at all. Our Definition 1 of a wavepacket involves only a finite number of terms and its invariance is provided by the flexible *tail term* $\hat{D}_h$. We also find remarkable the very simplicity of the definition which nevetherless allows for a sufficiently detailed analysis of the dynamics, including, in particular, rigorously justified NLS-type approximations of wavepacket dynamics presented in the following sections.

Our special interest is in waves that are finite sums of wavepackets and we refer to them as *multi-wavepackets*.

Definition 2 (Multi-wavepacket). *Let S be a set of nk-pairs:*

$$S=\{(n_l,\mathbf{k}_{*l}),\ l=1,\ldots,N\}\subset\Sigma=\{1,\ldots,J\}\times\mathbb{R}^d,\ (n_l,\mathbf{k}_{*l})\neq(n_{l'},\mathbf{k}_{*l'})\ \textit{for}\ l\neq l',\tag{31}$$

and $N=|S|$ be their number. Let K_S be a set consisting of all different wavevectors $\mathbf{k}_{*l}$ involved in S with $|K_S|\leq N$ being the number of its elements. K_S is called **wavepacket *k*-spectrum** *and without loss of genericity we assume the indexing of elements in S to be such that*

$$K_S=\{\mathbf{k}_{*i},\ i=1,\ldots,|K_S|\},\ \textit{i.e.}\ l_i=i\ \textit{for}\ 1\leq i\leq|K_S|.\tag{32}$$

A function $\hat{\mathbf{h}}(\beta)=\hat{\mathbf{h}}(\beta,\mathbf{k})$ is called a **multi-wavepacket** *with* ***nk*-spectrum** *S if it is a finite sum of wavepackets, namely*

$$\hat{\mathbf{h}}(\beta,\mathbf{k})=\sum_{l=1}^{N}\hat{\mathbf{h}}_l(\beta,\mathbf{k}),\ 0<\beta<\beta_0\ \textit{for some}\ \beta_0>0,\tag{33}$$

where $\hat{\mathbf{h}}_l$, $l=1,\ldots,N$, is a wavepacket with nk -pair $(n_l,\mathbf{k}_{*l})\in S$ as in Definition 1.

Note that if $\hat{\mathbf{h}}(\beta,\mathbf{k})$ is a wavepacket then $\hat{\mathbf{h}}(\beta,\mathbf{k})+O(\beta^s)$ is a wavepacket as well with the same nk-spectrum, and the same is true for multi-wavepackets. Hence, we can introduce a multi-wavepackets equivalence relation "$\simeq$" of degree s by

$$\hat{\mathbf{h}}_1(\beta,\mathbf{k})\simeq\hat{\mathbf{h}}_2(\beta,\mathbf{k})\ \text{if}\ \left\|\hat{\mathbf{h}}_1(\beta,\mathbf{k})-\hat{\mathbf{h}}_2(\beta,\mathbf{k})\right\|_{L^1}\leq C\beta^s\ \text{for some constant}\ C>0.\tag{34}$$

Observe also that zero functions are (trivial) wavepackets for any given (n,k)-spectrum. A wavepacket with any pair (n,k) is equivalent to zero if its L^1 norm is bounded by β^s, and such trivial components of two equivalent wavepackets are excluded; the remaining sets of elements $(n_l,\mathbf{k}_{*l})$ of spectra of two equivalent wavepackets must coincide.

Let us turn now to the abstract nonlinear problem (16) where (i) $\mathcal{F}=\mathcal{F}(\varrho)$ depends on ϱ and (ii) the initial data $\hat{\mathbf{h}}=\hat{\mathbf{h}}(\beta)$ is a multi-wavepacket depending on β. We would like to state our first theorem on multi-wavepacket preservation under the evolution (16) for $\beta,\varrho\to0$, which holds, as it turns out, provided its nk-spectrum S

satisfies a certain natural condition called *resonance invariance*. This condition is intimately related to the so-called *phase and frequency matching* conditions for stronger nonlinear interactions, and its concise formulation is as follows. We define for given dispersion relations $\{\omega_n(\mathbf{k})\}$ and any finite set $S \subset \{1, \ldots, J\} \times \mathbb{R}^d$ another finite set $\mathcal{R}(S) \subset \{1, \ldots, J\} \times \mathbb{R}^d$, where $\mathcal{R}$ is a certain algebraic operation described in Definition 18 below. It turns out that for any S always $S \subseteq \mathcal{R}(S)$ but *if, in fact,* $\mathcal{R}(S) = S$ *we call* S *resonance invariant*. The condition of resonance invariance is instrumental for the multi-wavepacket preservation, and there are examples showing that if it fails, i.e. $\mathcal{R}(S) \neq S$, the wavepacket preservation does not hold. *Importantly, the resonance invariance* $\mathcal{R}(S) = S$ *allows resonances inside the multi-wavepacket, that includes, in particular, resonances associated with the second and the third harmonic generations, resonant four-wave interaction, etc.*

Theorem 3 (Multi-wavepacket preservation). *Suppose that the nonlinear evolution is governed by* (16) *and the initial data* $\hat{\mathbf{h}} = \hat{\mathbf{h}}(\beta, \mathbf{k})$ *is a multi-wavepacket with nk-spectrum* S *and the regularity degree* s*, and assume* S *to be resonance invariant* (*see Definition* 18 *below*)*. Let dependence between parametrs* ϱ *and* β *be any function* $\varrho = \rho(\beta)$ *satisfying*

$$0 < \rho(\beta) \leq C\beta^s, \text{ for some constant } C > 0, \tag{35}$$

and let us set $\varrho = \rho(\beta)$*. Then the solution* $\hat{\mathbf{u}}(\tau, \beta) = \mathcal{G}\left(\mathcal{F}(\rho(\beta)), \hat{\mathbf{h}}(\beta)\right)(\tau)$ *to* (16) *for any* $\tau \in [0, \tau_*]$ *is a multi-wavepacket with nk-spectrum* S *and the regularity degree* s*, i.e.*

$$\hat{\mathbf{u}}(\tau, \beta; \mathbf{k}) = \sum_{l=1}^{N} \hat{\mathbf{u}}_l(\tau, \beta; \mathbf{k}), \text{ where } \hat{\mathbf{u}}_l \text{ is wavepacket with nk-pair } (n_l, \mathbf{k}_{*l}) \in S. \tag{36}$$

The time interval length $\tau_* > 0$ *depends only on* L^1*-norms of* $\hat{\mathbf{h}}_l(\beta, \mathbf{k})$ *and* N*. The presentation* (36) *is unique up to the equivalence* (34)*.*

The above statement can be interpreted as follows. Modes in *nk*-spectrum S are always resonance coupled with modes in $\mathcal{R}(S)$ through the nonlinear interactions, but if $\mathcal{R}(S) = S$ then (i) all resonance interactions occur inside S and (ii) only a small vicinity of S is involved in nonlinear interactions leading to the multi-wavepacket preservation.

Many nonlinear evolution problems with small initial data can be readily reduced by elementary rescaling to the system (1) with a large parameter $\frac{1}{\varrho}$ before its linear part. For example, suppose that $\mathbf{F}(\mathbf{V})$ is a homogeneous nonlinearity of degree m ($m = 3$ for a cubic one) and that the nonlinear evolution is governed by

$$\partial_t \mathbf{V} = -\mathrm{i}\mathbf{L}(-\mathrm{i}\nabla)\mathbf{V} + \mathbf{F}(\mathbf{V}), \quad \mathbf{V}(\mathbf{r}, t)|_{t=0} = \varrho^{1/(m-1)}\mathbf{h}(\mathbf{r}), \quad \mathbf{r} \in \mathbb{R}^d, \tag{37}$$

considered for small ϱ on the large time interval $0 \leq t \leq \frac{\tau_*}{\varrho}$ with a fixed $\tau_* > 0$. Then the following simple change of variables:

$$\mathbf{V}(t) = \varrho^{1/(m-1)}\mathbf{U}(\tau), \quad \tau = t\varrho \tag{38}$$

transforms the problem (37) into the equivalent problem (1). In this case the inequality (35) describes a constraint between the spatial extension $\frac{1}{\beta}$ and the amplitude factor $\varrho^{1/(m-1)} = \rho(\beta)^{1/(m-1)}$ of the initial data. Observe that Eq. (37) does not have any small parameters and both small parameters ϱ and β enter the problem through its initial data. Theorem 3 can be restated for problem (37) as follows:

Corollary 4 (Multi-wavepacket preservation). *Let* $\mathbf{V}(\mathbf{r}, t)$ *be a solution to the nonlinear system* (37), $\rho(\beta)$ *is as in* (35) *and we set* $\varrho = \rho(\beta)$. *Then if the initial data is such that* $\varrho^{-1/(m-1)}\hat{\mathbf{V}}(\mathbf{k}, 0) = \hat{\mathbf{h}}(\mathbf{k})$ *is a multi-wavepacket, then* $\varrho^{-1/(m-1)}\hat{\mathbf{V}}(\mathbf{k}, t)$ *remains as a multi-wavepacket with the same nk-spectrum and the degree of regularity for all times* $t \in \left[0, \frac{\tau_*}{\varrho}\right]$.

The statements of Theorems 3 and Corollary 4 directly follow from the following general theorem which makes no assumptions on the relations between $\beta, \varrho \to 0$.

Theorem 5 (Multi-wavepacket approximation). *Let the initial data* $\hat{\mathbf{h}}$ *in the integral equation* (16) *be a multi-wavepacket* $\hat{\mathbf{h}}(\beta, \mathbf{k})$ *with nk-spectrum S as in* (31)*, the regularity degree s and with the parameter* $\epsilon > 0$ *as in Definition 1. Assume that S is resonance invariant in the sense of Definition* 18 *below. Let the cutoff function* $\Psi(\mathbf{k}, \mathbf{k}_*)$ *and the eigenvector projectors* $\Pi_{n,\pm}(\mathbf{k})$ *be defined by* (26) *and* (11) *respectively. For a solution* $\hat{\mathbf{u}}$ *of* (16) *we set*

$$\hat{\mathbf{u}}_l(\tau, \beta; \mathbf{k}) = \left[\sum_{\zeta=\pm} \Psi(\mathbf{k}, \zeta\mathbf{k}_{*l}) \Pi_{n_l,\zeta}(\mathbf{k})\right] \hat{\mathbf{u}}(\tau, \beta; \mathbf{k}), \ l = 1, \ldots, N. \tag{39}$$

Then every such $\hat{\mathbf{u}}_l(\mathbf{k}; \tau, \beta)$ *is a wavepacket and*

$$\sup_{0 \le \tau \le \tau_*} \left\| \hat{\mathbf{u}}(\tau, \beta; \mathbf{k}) - \sum_{l=1}^{N} \hat{\mathbf{u}}_l(\tau, \beta; \mathbf{k}) \right\|_{L^1} \le C_1 \varrho + C_2 \beta^s, \tag{40}$$

where the constant C_1 *does not depend on* ϵ, s *and* β*, and the constant* C_2 *does not depend on* β.

It is interesting to note that the statement of Theorem 5 can be extended to the special limit case $\beta = 0$, $\mathbf{k}_{*l} = 0$. In this case the initial data of (1) are constants in $\mathbf{r}$ and we can consider solutions $\mathbf{U}$ (1) which do not depend on $\mathbf{r}$. Then $\nabla \mathbf{U} = \mathbf{0}$, the linear operator $\mathbf{L}(-\mathrm{i}\nabla)$ reduces to the multiplication by a matrix $\mathbf{L}_0 = \mathbf{L}(0)$ and the system (1) turns into a system of ordinary differential equations (ODE). Notice that (i) the structure of the eigenvalues (7) implies that the linear part is time-reversible; (ii) the nonlinear part can be an arbitrary polynomial. The extension of Theorem 5 to this case (see Theorem 11) reads that in a generic, non-resonant situation if initial data are bounded and a set of eigenmodes of the matrix $\mathbf{L}_0$ is excited at $\tau = 0$, then in the course of evolution on a time interval $[0, \tau_*]$ where τ_* depends on magnitude of initial data: (i) all remaining modes remain unexcited with accuracy proportional to ϱ, and (ii) only the originally excited modes can significantly evolve with this level of accuracy. For finite-dimensional systems governed by ODE's such a statement can be derived from the classical time-averaging principle and the time-averaged equations remain nonlinear. For infinitely-dimensional systems governed by PDE and with the linear operator having a continuous spectrum, as in Theorem 5, the analysis is more complex but the time-averaging still plays an important role yielding an accurate approximation governed by a certain universal nonlinear PDE.

We would like to point out also that though Theorem 3 is a simple corollary of the more general Theorem 5, *it is important that the statement* (40) *can be formulated as multi-wavepacket invariance*. That, in particular, allows to take values $\hat{\mathbf{u}}(\tau_*)$ as new wavepacket initial data for (1) and extend the wavepacket invariance of a solution to the next time interval $\tau_* \leq \tau \leq \tau_{*1}$. This observation allows to extend the wavepacket invariance to larger values of τ (up to blow-up time or infinity) if some additional information about solutions with wavepacket initial data is available. In particular, the following theorem holds.

Theorem 6. *Assume that all conditions of Theorem 3 are satisfied and, in addition to that, solutions* $\hat{\mathbf{u}}(\tau,\beta)$ *of* (16) *with the multi-wavepacket initial data* $\hat{\mathbf{h}}(\beta)$ *and* $\varrho = \rho(\beta)$ *exist on an interval* $0 \leq \tau < \tau_0$, $\tau_0 \leq \infty$, *and the estimate* $\left\|\hat{\mathbf{u}}(\cdot,\beta)\right\|_{C([0,\tau_1],L^1)} \leq R(\tau_1)$ *holds for any* $\tau_1 < \tau_0$, *where* $R(\tau_1)$ *does not depend on* $\beta \leq \beta_0$. *Then the solution* $\hat{\mathbf{u}}(\tau,\beta) = \mathcal{G}\left(\mathcal{F}(\rho(\beta)), \hat{\mathbf{h}}(\beta)\right)(\tau)$ *to* (16) *for any* $\tau < \tau_0$ *is a multi-wavepacket with nk-spectrum* S *and the regularity degree* s, *that is* (36) *holds.*

The derivation of the above statement from Theorem 3 is straightforward with the following key points. The interval τ_* in Theorem 3 depends only on the L^1- norm of initial data and the solution $\hat{\mathbf{u}}(\tau,\beta)$ is assumed to be bounded in L^1 by $R(\tau) \leq R(T)$ for $0 \leq \tau \leq T$ for any $T < \tau_0$. Therefore, we can apply Theorem 3 consecutively on intervals $[n\tau_*, (n+1)\tau_*]$ for all integers n such that $0 \leq n\tau_* \leq T$ and conclude that if $\hat{\mathbf{u}}(\tau,\beta)$ is a wavepacket for $\tau = n\tau_*$ it remains to be a wavepacket for $\tau \in [n\tau_*, (n+1)\tau_*]$. Note that parameters β_0 and C' in Definition 1 may depend on a wavepacket and be different for different wavepackets. Importantly, τ_* in the statement of Theorem 5 does not depend on β_0 and C'. Since for any fixed $T < \tau_0$ we can apply Theorem 3 a finite number of times the solution $\hat{\mathbf{u}}(\tau)$ is a wavepacket on the interval $[0,T]$ if $T < \tau_0$ (with some parameters $\beta_0(T) > 0$ and $C'(T) < \infty$).

Note that the wavepacket form of solutions can be used to obtain long-time estimates of solutions. Namely, very often behavior of every single wavepacket is well approximated by its own nonlinear Schrodinger equation (NLS), see 17,34,18,23,30, 31,47,50,51,53] and references therein, see also Sect. 6. Many features of the dynamics governed by NLS-type equations are well-understood, see 14,16,32,49,57,59] and references therein. These results can be used to obtain long-time estimates for every single wavepacket (as, for example, in 31]) and, with the help of the superposition principle, for the multiwavepacket solution.

The wavepacket representation (36) from Theorem 3 can be used for more detailed analysis of dynamics of wavepackets $\hat{\mathbf{u}}_l(\tau,\beta)$ and interaction between them. The following theorem illustrates that by describing wavepacket interaction based on a system with a weakly universal nonlinearity similar to so-called coupled modes systems or NLS.

Theorem 7 (NLS-type approximation). *Let the conditions of Theorem 5 hold and, in addition to that, the initial data* $\hat{\mathbf{h}}_l(\mathbf{k})$ *are of the form* $\hat{\mathbf{h}}_l = \hat{\mathbf{h}}_{l,+} + \hat{\mathbf{h}}_{l,-} + \hat{D}_l$, *where*

$$\hat{\mathbf{h}}_{l,\zeta}(\mathbf{k}) = \beta^{-d}\hat{H}_{l,\zeta}\left(\beta^{-1}(\mathbf{k}-\zeta\mathbf{k}_{*l})\right)\mathbf{g}_{n_l,\zeta}(\mathbf{k}) \quad \textit{for } |\mathbf{k}-\mathbf{k}_{*l}| \leq \beta^{1-\epsilon},\ \zeta = \pm,$$

$\hat{D}_l$ *satisfies* (30), *and every function* $\hat{H}_{l,\zeta}(\boldsymbol{\eta})$, *which may depend on* β, *is defined for all* $\boldsymbol{\eta}$ *and is bounded in* $L^{1,a}$ *with* $a > \frac{s}{\epsilon}$ *uniformly in* β. *Then one can write a nonlinear system of differential equations for* $2N$ *scalar envelope functions* $z_{l,\zeta}(\tau,\mathbf{r})$ *with the initial data* $H_{l,\zeta}$, *a linear part of the system has order* $\mu \leq 3$ *and the nonlinearity is weakly universal as in* (238) *and has order* $\nu \leq 1$. *Let* $\hat{z}_{l,\zeta}(\tau,\mathbf{k})$, $l = 1,\ldots,N$, *be the Fourier transform of a solution to this system. Then there exist* $\beta_0 > 0$ *and a constant*

C which does not depend on β, ϱ such that for $\beta \leq \beta_0$ the solution $\hat{\mathbf{u}}$ of (16) with initial data $\hat{\mathbf{h}}$ can be approximated as follows:

$$\sum_{l=1}^{N} \left\| \hat{\mathbf{u}}_l(\tau, \beta) - \beta^{-d} \hat{z}_{l,\zeta} \left(\tau, \beta^{-1} (\cdot - \mathbf{k}_{*l}) \right) \mathbf{g}_{n_l,\zeta} \right\|_E \leq C \left[\varrho + \frac{\beta^{(\mu+1)(1-\epsilon)}}{\varrho} + \beta^{(\nu+1)(1-\epsilon)} + \beta^s \right]. \quad (41)$$

The above-mentioned system with a weakly universal nonlinearity is constructed based on Eq. (1) and nk-spectrum S with the help of time averaging (70) described below. Note that in the simplest case when $\mu = 2$, $\nu = 0$, $N = 1$ (and J is arbitrary) the resulting system with a universal nonlinearity is equivalent to the classical Nonlinear Schrodinger equation (NLS). If $N = 2$ and $\mathbf{k}_{*1} = -\mathbf{k}_{*2}$ we obtain the well-known coupled modes system for counterpropagating waves. This theorem applied to particular systems implies approximation theorems similar to results of (i) 30,53,6,23] on NLS approximation; (ii) 6,24,47,52] on coupled mode approximation; (iii) 54] on three-wave approximations. Note also that (41) implies that if $\varrho = \beta^{\varkappa'}$ with $1 < \varkappa' < 2$, then both the first order hyperbolic equations ($\mu = 1$, $\nu = 0$) and the second-order NLS ($\mu = 2$, $\nu = 0$) provide an approximation for a solution $\hat{\mathbf{u}}$ of (16), but NLS provides a better approximation $O\left(\beta^{(1-\epsilon)}\right)$ compared with $O\left(\beta^{2(1-\epsilon)-\varkappa'}\right)$ for first order hyperbolic equations.

Observe that in the form (22) for a simple wavepacket we require $\mathbf{g}_{n,\pm}(\mathbf{k}_*)$ to be an eigenvector of the Hermitian matrix $\mathbf{L}(\mathbf{k}_*)$, and one can wonder if $\mathbf{g}_{n,\pm}(\mathbf{k}_*)$ can be replaced with an arbitrary pair of vectors $\mathbf{g}_{\pm}$ in the case $J > 1$. The answer is affirmative, since one can always expand any $\mathbf{g}$ with respect to the basis $\mathbf{g}_{n,\pm}(\mathbf{k})$ using $\Pi_{n,\pm}(\mathbf{k})$, but the result will be a multi-wavepacket with up to $2J$ components rather than a single wavepacket.

The rest of the paper is organized as follows. In the next section we illustrate important points of parameter dependence and wavepacket preservation based on examples. In Sect. 3 we formulate conditions of wavepacket preservation including the key resonance invariance condition. In Sect. 4 we provide examples of different forms of equations and systems which involve small or large parameters and can be written in the form of (1) after a rescaling. In Sect. 5 we introduce and discuss integrated modal forms of the evolution equation. In Sect. 6 we introduce and study the wavepacket interaction system in its relation to the original system. In Sect. 7 we approximate the wavepacket interaction system by a certain minimal wavepacket interaction system, which in the simplest cases turns into the NLS or the coupled modes system.

2. Preliminary Discussion and Examples

Observe that the multi-wavepacket preservation as described in Theorems 3-7 states in different forms that (i) its modal composition is essentially preserved; (ii) its nk-spectrum (the set of nk-pairs $\{\mathbf{k}_{*l}, n_l\}$) remains the same at all times; (iii) no new modes are excited with good accuracy as a result of the nonlinear evolution. The preservation of multi-wavepackets as they evolve shows also that only the nonlinear interactions between small neighborhoods of points $(\mathbf{k}_{*l}, n_l)$ are essential and contribute constructively to the nonlinear dynamics, whereas the amplitudes of modes with wavevectors $\mathbf{k}$ outside those neighborhoods is vanishingly small as $\beta, \varrho \to 0$. The latter is quite remarkable since the coupling term $\hat{\mathbf{F}}\left(\hat{\mathbf{U}}\right)(\mathbf{k})$ in (3) for such $\mathbf{k}$ is not small. A qualitative explanation to that, confirmed by rigorous analysis, is based on a fact that the contribution of this term to the solution is a time integral involving highly oscillatory functions that becomes vanishingly small as $\beta, \varrho \to 0$. This mechanism is similar to the classical averaging mechanism for systems of ordinary differential equations described, for instance, in 11]; the relevance of the averaging mechanism for long-wave asymptotics for hyperbolic systems of PDE is well-known, see 30].

We would like to relate now the multi-wavepacket preservation property to the linear superposition for wavepackets established in 7]. According to that principle if the initial state $\mathbf{h} = \sum \mathbf{h}_l$, with $\mathbf{h}_l$, $l = 1, \ldots, N$ being "generic" wavepackets, then the solution $\hat{\mathbf{u}}(\tau) = \mathcal{G}(\mathbf{h})(\tau)$ to the evolution equation (15) equals with high accuracy the sum of individual solutions $\mathbf{u}_l$ of N equations with respective initial data $\mathbf{h}_l$. Namely, if $\beta, \varrho > 0$ satisfy the following relation:

$$\beta, \varrho \to 0,\ \beta \geq C_1\varrho \text{ with some } C_1 > 0, \tag{42}$$

then for all times $0 \leq \tau \leq \tau_*$ we have

$$\mathcal{G}\left(\sum_{l=1}^{N} \mathbf{w}_l\right)(\tau) = \sum_{l=1}^{N} \mathcal{G}(\mathbf{w}_l)(\tau) + \mathbf{D}(\tau), \tag{43}$$

$$\|\mathbf{D}(\tau)\|_E = \sup_{0\leq\tau\leq\tau_*} \|\mathbf{D}(\tau)\|_{L^\infty} \leq C_\epsilon \frac{\varrho}{\beta^{1+\epsilon}} + C\beta \text{ for any } \epsilon > 0. \tag{44}$$

The *linear superposition principle* is formulated in 7] for $\beta = C_2\varrho^{1/2}$, but, in fact, the provided proofs of (43), (44) remain valid as long as (42) holds. Obviously, the bound $\beta \geq C_1\varrho$ in (42) determines when (44) becomes trivial. This bound is sharp and examples below show that when $\beta \sim \varrho$ the remainder $\mathbf{D}(\tau)$ in (43) does not tend to zero when $\beta \to 0$.

Both the multi-wavepacket preservation and the linear superposition apply to sums of generic wavepackets. It is important to notice though that the multi-wavepacket preservation holds for any dependence between ϱ and β which satisfy (35), that is $\varrho(\beta) \leq C\beta^q$ with arbitrary small q whereas the linear superposition holds if $\varrho(\beta) \leq C\beta$. Thus, the bounds (42) on β determine the range of its values for which both multi-wavepacket preservation and linear superposition hold simultaneously (provided some genericity conditions are satisfied). In this range wavepacket preservation provides additional information on behavior of solutions with single wavepacket initial data, namely that the solution remains a single wavepacket. Obviously, the linear superposition principle does not follow from multi-wavepacket invariance. Below we use simple examples and models to discuss different ranges of parameters ϱ and β where wavepacket preservation is valid but the solutions of equations exhibit different behavior.

2.1. A model with explicit solutions and the effect of large group velocity. Here we introduce a simple model for our general system (1) with elementary solutions which makes explicit that in the limit $\varrho \to 0$ nonlinear effects do not vanish, in particular the blow-up time does not tend to infinity. This example also shows that on the time scale where τ is of order 1 solutions undergo significant nonlinear evolution. The influence of ϱ on solutions through the group velocity in this example can be seen explicitly. The model is the following system of two coupled nonlinear first order hyperbolic equations for variables $u_1(x,\tau)$, $u_2(x,\tau)$ with one-dimensional spatial variable x:

$$\partial_\tau u_1 = -\frac{c_1}{\varrho}\partial_x u_1 + F_1(u_1,u_2), \tag{45}$$

$$\partial_\tau u_2 = -\frac{c_2}{\varrho}\partial_x u_2 + F_2(u_1,u_2), \quad c_1 \neq c_2, \quad u_1|_{\tau=0} = h_1(x), \quad u_2|_{\tau=0} = h_2(x), \tag{46}$$

where the initial data h_1, h_2 in (46) are of wavepacket form:

$$h_1(x) = \Phi_1(\beta x)\cos k_{1*}x, \quad h_2(x) = \Phi_2(\beta x)\cos k_{2*}x, \quad |k_{1*}| \neq |k_{2*}|. \tag{47}$$

We take the nonlinearity to be quadratic and of the following simple form:

$$F_1(u_1,u_2) = u_1^2 + a_1u_1u_2, \quad F_2(u_1,u_2) = u_2^2 + a_2u_1u_2. \tag{48}$$

The system (45)–(47) allows for an explicit form of solutions with one-wavepacket initial data, describing a wave propagating with a constant speed controlled by the linear part and with a shape evolution controlled by the nonlinearity. This simplest case is compared then with the case of two-wavepacket initial data, for which an explicit solution is not available.

In the case when $h_2 = 0$ the second equation has trivial solution $u_2 = 0$ and the system (45)–(46) reduces to a single equation (45). The solution to this equation has the form of a traveling wave $v_1\left(x - \frac{c_1}{\varrho}\tau, \tau\right)$, where $v_1(y,\tau)$ is a solution of the ordinary differential equation

$$\partial_\tau v_1 = F_1(v_1,0), \quad v_1(y,0) = h_1(y). \tag{49}$$

The explicit formula in the case (49) yields

$$v_1(x,\tau) = \frac{h_1\left(x - \frac{c_1\tau}{\varrho}\right)}{1 - \tau h_1\left(x - \frac{c_1\tau}{\varrho}\right)} = \frac{\Phi_1\left(\beta\left(x - \frac{c_1\tau}{\varrho}\right)\right)\cos k_{1*}\beta\left(x - \frac{c_1\tau}{\varrho}\right)}{1 - \tau\Phi_1\left(\beta\left(x - \frac{c_1\tau}{\varrho}\right)\right)\cos k_{1*}\beta\left(x - \frac{c_1\tau}{\varrho}\right)} \tag{50}$$

for a time interval $0 \leq \tau < \tau_0$, where $\tau_0 = \frac{1}{\sup_y |h_1(y)|}$ is the *blow-up time*. Obviously, the blow-up time does not depend on ϱ. Consequently, the wave propagates with the velocity $\frac{c_1}{\varrho}$ with its shape evolution being controlled by the nonlinearity. Similarly, when $h_1 = 0$ the first equation has the trivial solution $u_1 = 0$ and the system (45)–(46) reduces to a single equation (46) which has a solution in the form of a traveling wave $v_2\left(x - \frac{c_2}{\varrho}\tau, \tau\right)$ propagating with the velocity $\frac{c_2}{\varrho}$. Observe that for the simple model (45)–(47) the group velocity coincides with the velocity of a traveling wave.

The above model is not exactly solvable if both initial conditions h_1 and h_2 do not vanish. But one can still see the way ϱ influences the nonlinear dynamics quite explicitly by applying the superposition principle from 6]. Indeed, let us assume that h_1 and h_2 are two nonzero initial wavepackets. Then the approximate superposition principle is applicable (in order to put the system in the framework of 6] we use the 4-component extension (115) and set $\varrho = \beta^{\varkappa'}$, $\varkappa' > 1$). According to the principle the exact solution

(u_1, u_2) is approximated by $\left(v_1\left(x - \frac{c_1}{\varrho}\tau, \tau\right), v_2\left(x - \frac{c_2}{\varrho}\tau, \tau\right)\right)$, which is explicitly given by (50) with the accuracy $O\left(\frac{\varrho}{\beta^{1+\epsilon}}\right) = O\left(\beta^{\varkappa'-1-\epsilon}\right)$ with arbitrary small ϵ if $c_1 \neq c_2$. As it as shown in 6] the validity of such an approximate presentation is due to the large difference $\frac{c_1 - c_2}{\varrho}$ of the group velocities of two wavepackets.

2.2. Dispersive effects and nonlinearity. Based on an elementary example of the Nonlinear Schrodinger equation (NLS),

$$\partial_\tau u = -\frac{\mathrm{i}}{\varrho}\left[\gamma_0 u + \mathrm{i}\gamma_1 \partial_x u + \gamma_2 \partial_x^2 u\right] + b_1 |u|^2 u,\ u = u(x, \tau),\ x \in \mathbb{R} \qquad (51)$$

with the initial data in the form of a wavepacket $u|_{\tau=0} = \Phi(\beta x)\,\mathrm{e}^{\mathrm{i}k_* x}$, we would like to explain here why we are interested mostly in the case

$$\frac{\varrho}{\beta^2} \geq C > 0, \qquad (52)$$

when the dispersion *is not dominant*. To make the dependence of u on β and ϱ explicit we change the variables

$$u(x) = v(\beta x)\,\mathrm{e}^{\mathrm{i}k_* x},\ \ \beta x = z, \qquad (53)$$

and obtain the equation

$$\partial_\tau v = -\frac{\mathrm{i}}{\varrho}\left[\gamma_0' v + \mathrm{i}\beta\gamma_1' \partial_z v + \gamma_2 \beta^2 \partial_z^2 v_1\right] + b|v|^2 v,\ \ v|_{\tau=0} = \Phi(z), \qquad (54)$$

where $\gamma_1' = \gamma_1/\beta + 2\gamma_2 k_*$. Changing variables once more,

$$v(z, \tau) = \mathrm{e}^{-\frac{\mathrm{i}\tau}{\varrho}\gamma_0'} w\left(z + \frac{\beta}{\varrho}\gamma_1'\tau, \tau\right),\quad z + \frac{\beta}{\varrho}\gamma_1'\tau = y, \qquad (55)$$

we obtain for the envelope w the following standard NLS equation:

$$\partial_\tau w = -\frac{\mathrm{i}\beta^2}{\varrho}\gamma_2 \partial_y^2 w + b|w|^2 w,\quad w|_{\tau=0} = \Phi(y),\ 0 \leq \tau \leq \tau_*, \qquad (56)$$

with initial data independent of the parameters β, ϱ. The behavior of the solution w to Eq. (56) on the time interval $0 \leq \tau \leq \tau_*$ is determined by the dispersion parameter $\frac{\beta^2}{\varrho}$, and evidently linear dispersive effects become significant when $\frac{\varrho}{\beta^2}$ is not too large. If $\frac{\beta^2}{\varrho} \to \infty$ and $\beta \to 0$, *the solution tends to zero at every fixed* $\tau = \tau_0 > 0$. Indeed, if we take $\varrho = \beta^{\varkappa'}$, $\varkappa' > 2$, and make another change of variables $\tau = t\beta^{\varkappa'-2}$, $w = \beta^{1-\varkappa'/2} W$, Eq. (56) reduces to the following problem with small initial data:

$$\partial_t W = -\mathrm{i}\gamma_2 \partial_y^2 W + b|W|^2 W,\quad W|_{t=0} = \beta^{\varkappa'/2-1}\Phi(y). \qquad (57)$$

For small enough β the solution W to this problem exists for all t and $W(t) \to 0$ as $t \to \infty$ (see 16]). In particular, for $t = \tau_0 \beta^{2-\varkappa'}$ we have $w(\tau_0) \to 0$ when $\beta \to 0$.

In the general case, the solution dependence on small β, ϱ is as follows. The dependence on large $\frac{1}{\varrho}$ in (51) is completely described by the change of variables (55), yielding a wave which (i) moves as a whole with a large group velocity $\frac{-\gamma_1'}{\varrho}$; (ii) has a slowly evolving shape as described by v and w in (53), (55), (56).

The above observations show that for small $\frac{\varrho}{\beta^2}$ the dispersive effects dominate and control the nonlinear ones. Keeping that in mind and being interested in stronger nonlinear effects we focus primarily on the case (52), i.e. $\frac{\varrho}{\beta^2} \geq C > 0$, for which there are two scenarios of the nonlinear evolution. In the first scenario, when $\frac{\beta^2}{\varrho} \to 0$, the linear dispersion produces only a small correction to the solution of the equation $\partial_\tau w = b\,|w|^2\,w$ with that nonlinear equation governing the nonlinear dynamics of the envelope w for τ_* being smaller than the blow-up time. In the second scenario, when $\beta^2 \sim \varrho$, Eq. (56) becomes independent of β, ϱ and describes the evolution of the envelope w governed by an interplay between the dispersion and the nonlinearity. The case $\beta^2 \sim \varrho$ can be also characterized as one where dispersive effects do occur but they don't dominate nonlinear effects, and, as it is well known, the dispersion can exactly balance the nonlinearity yielding solitons.

2.3. A coupled modes system. Here we illustrate statements of the general theorem on the wavepacket preservation and the approximate superposition principle by a simple but still nontrivial example. Let us consider a system of two coupled NLS type equations for variables $u_1(x,\tau)$, $u_2(x,\tau)$ with one-dimensional spatial variable x,

$$\partial_\tau u_1 = -\frac{\mathrm{i}}{\varrho}\left[\gamma_{01} + \mathrm{i}\gamma_{11}\partial_x + \gamma_{21}\partial_x^2\right]u_1 + \left(b_{11}|u_1|^2 + b_{12}|u_2|^2\right)u_1 + c_{12}|u_2|^2 u_2, \tag{58}$$

$$\partial_\tau u_2 = -\frac{\mathrm{i}}{\varrho}\left[\gamma_{02} + \mathrm{i}\gamma_{12}\partial_x + \gamma_{22}\partial_x^2\right]u_2 + \left(b_{21}|u_1|^2 + b_{22}|u_2|^2\right)u_2 + c_{22}|u_1|^2 u_1, \tag{59}$$

$$u_1|_{\tau=0} = h_1(x) = \Phi_1(\beta x)\,\mathrm{e}^{\mathrm{i}k_{*1}x},\quad u_2|_{\tau=0} = h_2(x) = \Phi_2(\beta x)\,\mathrm{e}^{\mathrm{i}k_{*2}x}, \tag{60}$$

where γ_{ij} are real and b_{ij} are complex coefficients and the initial data in (60) are in the form of wavepackets with $\Phi_j(y)$ being Schwartz functions. Notice that if in the coupled modes system (58)–(60) $h_2 = 0$ and $c_{12} = c_{22} = 0$, then it has trivial solution $u_2 = 0$, and reduces to a single NLS equation of the form (51). The dependence of the solution $\{u_1, u_2\}$ on the large $\frac{1}{\varrho}$ is captured by the change of variables (55). Namely, u_1 is a wave with a slowly varying envelope described by v_1 which moves with large velocity $\frac{-\gamma_{11}'}{\varrho}$. The dependence on β is of the form $v_1(y,\tau) = w_1(\beta y,\tau)$ (see the following subsection for details). Similarly we can consider the case when $h_1 = 0$ for which the first equation has trivial solution $u_1 = 0$, so the system (58)–(59) reduces to a single equation (59) with the solution represented by a wave having large spacial extension proportional to $\frac{1}{\beta}$ and moving with the large velocity $\frac{-\gamma_{12}'}{\varrho}$.

2.3.1. The superposition principle. Let us assume here that $h_1 \neq 0$, $h_2 \neq 0$, $c_{12} \neq 0$, $c_{22} \neq 0$ and $\beta = \varrho^{\varkappa}$, $0 < \varkappa < 1$. Applying the superposition principle we obtain for generic k_{*1}, k_{*2} the following representation of the exact solution:

$$u_1(x,\tau) = v_1(x,\tau)\,\mathrm{e}^{\mathrm{i}k_{*1}x} + D_1,\quad u_2(x,\tau) = v_2(x,\tau)\,\mathrm{e}^{\mathrm{i}k_{*2}x} + D_2,$$

where $v_1(x,\tau)$ is a solution of the NLS equation (58) with $b_{12}=c_{12}=0$, with $v_2(x,\tau)$ being a solution to a similar decoupled NLS equation for $b_{22}=c_{22}=0$, and D_1 and D_2 are small terms satisfying

$$\sup_{0\le\tau\le\tau_*}\|D_1(\cdot,\tau)\|_{L^\infty}+\sup_{0\le\tau\le\tau_*}\|D_2(\cdot,\tau)\|_{L^\infty}\le C\beta^{\varkappa'-1-\epsilon}+C\beta,\ \varkappa'=\varkappa^{-1}. \tag{61}$$

We would like to emphasize here that the coupling terms $b_{12}|u_2|^2u_1+c_{12}|u_2|^2u_2$ and $b_{21}|u_1|^2u_2+c_{22}|u_2|^2u_2$ in Eq. (58)–(59) are not small whereas their ultimate contributions to the solutions are small. One can explain/interpret that phenomenon as being due to the destructive wave interference and mismatch of group velocities.

2.3.2. Wavepacket preservation. Here we assume that $h_1\ne0, h_2=0, c_{12}\ne0, c_{22}\ne0$ and $\varrho=\beta^{\varkappa'},\ 0<\varkappa'\le2$. According to the wavepacket preservation we have

$$u_1(x,\tau)=v_1(x,\tau)\,\mathrm{e}^{\mathrm{i}k_{*1}x}+D_1,\ u_2(x,\tau)=D_1,$$

where $v_1(x,\tau)$ is a solution of (58) with $b_{12}=0$, $c_{12}=0$, and D_1 and D_2 are small terms satisfying

$$\sup_{0\le\tau\le\tau_*}\|D_1(\cdot,\tau)\|_{L^\infty}+\sup_{0\le\tau\le\tau_*}\|D_2(\cdot,\tau)\|_{L^\infty}\le C\varrho.$$

Notice once more (see the above section) an interesting phenomenon: Eq. (59) for $u_2(x,\tau)$ has a coupling term $b_{21}|u_1|^2u_2+c_{22}|u_1|^2u_1$ which does not become small as $\beta,\varrho\to0$, but, remarkably, its ultimate contribution to the solution is small.

2.3.3. Limitations of the superposition principle. Now we provide an example based on the system (58)–(60) with $c_{12}=c_{22}=0$ showing that the above estimate (61) in the superposition principle is sharp in the sense that $\beta^{\varkappa'-1-\epsilon}$ cannot be replaced by $\beta^{\varkappa'-1+\epsilon}$ with $\varkappa'\ge1$. We set here $\varkappa'=1$ and $\varrho=\beta$. After the change of variables (53) for u_1,u_2 followed by yet another change of variables $\beta x=z$, $v_1=\mathrm{e}^{-\mathrm{i}\tau\frac{\gamma'_{01}}{\beta}}w_1$, $v_2=\mathrm{e}^{-\mathrm{i}\tau\frac{\gamma'_{01}}{\beta}}w_2$, we obtain from (58)–(60) the following system:

$$\begin{aligned}
\partial_\tau w_1&=-\mathrm{i}\left[\mathrm{i}\gamma'_{11}\partial_z w_1+\beta\gamma_{21}\partial_z^2w_1\right]+\left(b_{11}|w_1|^2+b_{12}|w_2|^2\right)w_1,\\
\partial_\tau w_2&=-\mathrm{i}\left[\mathrm{i}\gamma'_{12}\partial_z w_2+\beta\gamma_{22}\partial_z^2w_2\right]+\left(b_{21}|w_1|^2+b_{22}|w_2|^2\right)w_2,\\
w_1|_{\tau=0}&=\Phi_1(z),\ w_2|_{\tau=0}=\Phi_2(z).
\end{aligned}$$

This system has a regular dependence on β as $\beta\to0$ with the solution converging in L^∞ to the solution of the system with $\beta=0$. If we set now in the last system $b_{12}=b_{21}=0$ it turns into a system of two decoupled equations. Notice then that the difference between the solutions of the decoupled system and the original one does not tend to zero as $\beta\to0$, implying that the superposition principle does not hold when $\varrho=\beta$.

2.4. Wavepacket interaction system with a universal nonlinearity. We will prove in the following sections that the dynamics of a multi-wavepacket with a universally resonance invariant nk-spectrum for a general system can be approximated with the accuracy $O(\varrho)$ by substituting the nonlinearity with a properly constructed *universal* or weakly universal one. Here we provide an example of a system, called *wavepacket interaction* *system*, with a universal nonlinearity and show that its dynamics preserves simple wavepackets as in (12). It is shown later that universal nonlinearities are related to universally invariant multi-wavepackets in the sense of Definition 18.

Wavepacket interaction system with universal nonlinearity has the form similar to NLS, namely

$$\partial_\tau u_{j,\zeta} = \frac{1}{\varrho}\left[-\mathrm{i}\zeta\gamma_{0,j} + \gamma_{1,j}\cdot\nabla_{\mathbf{r}} u_{j,\zeta} - \mathrm{i}\zeta\nabla_{\mathbf{r}}\cdot\gamma_{2,j}\nabla_{\mathbf{r}} u_{j,\zeta}\right] + F_{j,\zeta}(\vec{u}), \quad \mathbf{r}\in\mathbb{R}^d, \tag{62}$$

$$\vec{u} = (u_{1+}, u_{1-}, \ldots, u_{N+}, u_{N-}), \quad j = 1, \ldots, N, \quad \zeta = \pm, \tag{63}$$

$$u_{j,\zeta}\big|_{\tau=0} = h_{j,\zeta}, h_{j,\zeta}(\mathbf{r}) = \Phi_j(\beta\mathbf{r})\,\mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_{*j}\cdot\mathbf{r}}, \tag{64}$$

where for every j coefficient $\gamma_{0,j}\in\mathbb{R}$, $\gamma_{1,j}\in\mathbb{R}^d$ is a vector, $\gamma_{2,j}$ is a symmetric $d\times d$ matrix, $\gamma_{1,j}\cdot\nabla_{\mathbf{r}}$ is a first order scalar differential operator, $\nabla_{\mathbf{r}}\cdot\gamma_{2,j}\nabla_{\mathbf{r}}$ is the second order scalar differential operator,and the *universal polynomial nonlinearities* $F_{j,\zeta}$ have the following form:

$$F_{j,\zeta}(\vec{u}) = \sum_{\nu=1}^{\nu_F}\sum_{|\vec{\nu}|=\nu} b_{\vec{\nu},j,\zeta}\prod\nolimits_{l=1}^{N}\left(u_{l,+}u_{l,-}\right)^{\nu_l} u_{j,\zeta},$$
$$\text{where } \vec{\nu} = (\nu_1, \ldots, \nu_N), \quad j = 1, \ldots, N, \ \zeta = \pm. \tag{65}$$

Remark 8. Notice that if we set $h_{j,-} = h^*_{j,+}$, $b_{\vec{\nu},j,+} = b^*_{\vec{\nu},j,-} = b_{\vec{\nu},j}$ and $u_{j,+} = u^*_{j,-} = u_j$ then $u_{l,+}u_{l,-} = \left|u_{l,+}\right|^2$ and $F_{j,+}(\vec{u})$ turns into

$$F_j(u_1, \ldots, u_N) = \sum_{\nu=1}^{\nu_F}\sum_{|\vec{\nu}|=\nu} b_{\vec{\nu},j}\prod\nolimits_{l=1}^{N}|u_l|^{2\nu_l} u_j, \tag{66}$$

and equations of (62) with $\zeta = +$ turn into

$$\partial_\tau u_j = \frac{1}{\varrho}\left[-\mathrm{i}\gamma_{0,j} + \gamma_{1j}\cdot\nabla_{\mathbf{r}} u_j - \mathrm{i}\nabla_{\mathbf{r}}\cdot\gamma_{2,j}\nabla_{\mathbf{r}} u_j\right] + F_j(u_1, \ldots, u_N),$$
$$u_j\big|_{\tau=0} = h_{j,+}, \quad j = 1, \ldots, N, \ \zeta = \pm. \tag{67}$$

Obviously, a solution of (67) defines a solution $u_{j,+} = u_j$, $u_{j,-} = u^*_j$ of the system (62). In the simplest case $N = 1$, $d = 1$ (67) takes the form of classical NLS: $\partial_\tau u = \frac{\gamma_1}{\varrho}\partial_x u - \mathrm{i}\frac{\gamma_2}{\varrho}\partial_x^2 u + b\,|u|^2\,u$.

Note that the *universal nonlinearity* $F_{j,\zeta}$ has a characteristic property

$$F_{j,\zeta}\left(\mathrm{e}^{\mathrm{i}\phi_1 t}u_{1,+}, \mathrm{e}^{-\mathrm{i}\phi_1 t}u_{1,-}, \ldots, \mathrm{e}^{\mathrm{i}\phi_N t}u_{N,+}, \mathrm{e}^{-\mathrm{i}\phi_N t}u_{N,-}\right)$$
$$= \mathrm{e}^{\mathrm{i}\zeta\phi_j t}F_{j,\zeta}(u_{1+}, u_{1-}, \ldots, u_{N+}, u_{N-}), \tag{68}$$

holding for arbitrary set values ϕ_i. We also consider more general nonlinearities F for which (68) holds for a fixed set of frequencies $\phi_l = \omega_{n_l}(\mathbf{k}_{*l})$, and call them *weakly* *universal*. We introduce now the *averaging operator* A_T acting on polynomial functions $F : \left(\mathbb{C}^2\right)^N \to \left(\mathbb{C}^2\right)^N$ by

$$(A_T F)_{j,\zeta} = \left(A_{T,\vec{\phi}} F\right)_{j,\zeta} =$$
$$\frac{1}{T}\int_0^T \mathrm{e}^{-\mathrm{i}\zeta\phi_j t} F_{j,\zeta}\left(\mathrm{e}^{\mathrm{i}\phi_1 t}u_{1,+}, \mathrm{e}^{-\mathrm{i}\phi_1 t}u_{1,-}, \dots, \mathrm{e}^{\mathrm{i}\phi_N t}u_{N,+}, \mathrm{e}^{-\mathrm{i}\phi_N t}u_{N,-}\right)\mathrm{d}t, \tag{69}$$

where $\vec{\phi} = (\phi_1, \dots, \phi_N)$. The operator $A_{T,\vec{\phi}}$ depends on the frequency vector $\vec{\phi} = (\phi_1, \dots, \phi_N)$. If F is a universal polynomial nonlinearity, then $\left(A_{T,\vec{\phi}} F\right)_{j,\zeta} = F_{j,\zeta}$ for any choice of frequencies $\phi_1, \dots, \phi_N$. Note that averaging

$$G_{\mathrm{av},j,\zeta}(\vec{u}) = \lim_{T\to\infty}\left(A_{T,\vec{\phi}} G\right)_{j,\zeta}(\vec{u}) \tag{70}$$

is defined for any polynomial nonlinearity $G : \left(\mathbb{C}^2\right)^N \to \left(\mathbb{C}^2\right)^N$. If $\vec{\phi}$ is generic, then $G_{\mathrm{av},j,\zeta}(\vec{u})$ is always a universal nonlinearity. In a general case $G_{\mathrm{av},j,\zeta}$ for given frequencies $\vec{\phi}$ one obtains a weakly universal nonlinearity which might be not universal.

Systems with universal nonlinearities have interesting properties which we describe in the following proposition and remark.

Proposition 9. *Let $\varrho = \beta$ and $\gamma_{2,j} = 0$. Then evolution governed by the first order system with a universal nonlinearity* (62) *preserves simple wavepackets as defined by* (12).

Proof. Let $\vec{u}(\tau)$ be a solution of (62) for $0 \le \tau \le \tau_*$. Using the property (68) we change variables

$$u_{j,\zeta} = \mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_{*j}\cdot\mathbf{r}}\mathrm{e}^{-\mathrm{i}\frac{\zeta\gamma_{0,j}}{\varrho}\tau}\mathrm{e}^{-\mathrm{i}\frac{\gamma'_{0j,\zeta}}{\beta}\tau}v_{j,\zeta},\quad \gamma'_{0j,\zeta} = -\zeta\boldsymbol{\gamma}_{1j}\cdot\mathbf{k}_{*j} \tag{71}$$

and obtain from (62)

$$\partial_\tau v_{j,\zeta} = \frac{1}{\beta}\boldsymbol{\gamma}_{1j}\cdot\nabla_{\mathbf{r}}v_{j,\zeta} + F_{j,\zeta}(\vec{v}),\quad v_{j,\zeta}\big|_{\tau=0} = \Phi_{j,\zeta}(\beta\mathbf{r}). \tag{72}$$

Changing variables

$$v_{j,\zeta}(\mathbf{r},\tau) = w_{j,\zeta}(\beta\mathbf{r},\tau),\quad \beta\mathbf{r} = \mathbf{z}, \tag{73}$$

we obtain from (72) that w_j is a solution of the following system of differential equations:

$$\partial_\tau w_{j,\zeta} = \gamma_{1j}\cdot\nabla_{\mathbf{z}}w_{j,\zeta} + F_{j,\zeta}(\vec{w}),\quad w_{j,\zeta}\big|_{\tau=0} = \Phi_{j,\zeta}(\mathbf{z}), \tag{74}$$

which does not depend on β. Then using (73) and (71) we observe that every component u_l of the solution to (62) has the form of a simple wavepacket for every $\tau \in [0, \tau_*]$, with an envelope $\hat{w}_j(\tau)$. □

Remark 10. Equations (62) with universal nonlinearities allow special solutions in the form of $u_{j,\zeta} = \mathrm{e}^{\mathrm{i}\mathbf{k}_{*j}\cdot\mathbf{r}}\mathrm{e}^{-\mathrm{i}\frac{\gamma'_{0j}}{\beta}\tau} v_{j,\zeta}(\tau)$, where $v_{j,\zeta}(\tau)$ do not depend on $\mathbf{r}$. If the initial data in (72) are constants, $\Phi_{j,\zeta}(\beta\mathbf{r}) = \Phi_{j,\zeta}(0)$, then (72) turns into a system of ODE. This implies that every linear subspace of pure modal functions with the basis $v_j \mathrm{e}^{\mathrm{i}\mathbf{k}_{*j}\cdot\mathbf{r}}$, $v_{j,-}\mathrm{e}^{-\mathrm{i}\mathbf{k}_{*j}\cdot\mathbf{r}}$, $j = 1, .., N$ is invariant with respect to nonlinear equations (62). Another class of special solutions of (62) are time-harmonic solutions of the form $u_{j,\zeta}(\mathbf{r},\tau) = \mathrm{e}^{-\mathrm{i}\zeta\omega_j\tau} v_{j,\zeta}(\mathbf{r})$, where $v_{j,\zeta}$ solve a nonlinear eigenvalue problem; for universal nonlinearities ω_j can be considered as an unknown nonlinear eigenvalue. Existence of such special solutions is a special property of universal and weakly universal nonlinearities. It is remarkable that original nonlinear equations might not have time harmonic solutions whereas equations with universal nonlinearities which approximate evolution of wavepackets (see Theorem 7) admit such solutions.

2.5. Invariance of excited modes for finite-dimensional ODE's. Here we discuss the resonance invariance conditions imposed in Theorem 5 in a simpler case of finite-dimensional ODE's. In this case one can also see the rise of universal nonlinearities in the process of time averaging. As we already discussed in the introduction, a PDE system (1) when restricted to constant functions turns into the following system of ODE's:

$$\partial_\tau \mathbf{U} = -\frac{\mathrm{i}}{\varrho}\mathbf{L}^0\mathbf{U} + \mathbf{F}(\mathbf{U}), \quad \mathbf{U}(\tau)|_{\tau=0} = \mathbf{h}, \quad \mathbf{h} \in \mathbb{C}^{2J}, \quad \mathbf{U} \in \mathbb{C}^{2J}, \tag{75}$$

where $\mathbf{F}(\mathbf{U})$ is a polynomial, $\mathbf{U} = \left(U_{1,+}, U_{1,-}, \ldots, U_{J,+}, U_{J,-}\right) \in \mathbb{C}^{2J}$. We assume that the eigenvalues $\omega_{n,\zeta}(\mathbf{0}) = \omega^0_{n,\zeta}$ of the Hermitian matrix $\mathbf{L}^0 = \mathbf{L}(\mathbf{k})|_{\mathbf{k}=0}$ are distinct $\omega^0_{j,+} \neq \omega^0_{i,+}$ for $j \neq i$ and the symmetry conditions (7) take the form $\omega^0_{n,-\zeta} = -\omega^0_{n,\zeta}$. We also assume that the eigenvectors of $\mathbf{L}^0$ coincide with the coordinate orts in $\mathbb{C}^{2J}$. The following limit case of Theorem 5 with $\beta = 0$ shows that solutions to this system have the property to preserve the set of initially excited modes.

Theorem 11. *Let the initial data* $\mathbf{h} = \left(h_{1,+}, h_{1,-}, \ldots, h_{J,+}, h_{J,-}\right) \in \mathbb{C}^{2J}$ *in* (75) *have non-zero components* $h_{j,\zeta}$ *only for a subset* B *of indices* $j \in \{1, \ldots, J\}$*, and let* $B' = \{1, \ldots, J\} \setminus B$ *be its complementary set. Assume that* B *is* **resonance invariant** *in the sense that the resonance equation*

$$\omega^0_{n',\zeta} - \sum_{j=1}^{m} \omega^0_{n_j,\zeta^{(j)}} = 0, \quad \textit{where } n_j \in B, \; \zeta^{(j)} \in \{+,-\} \tag{76}$$

does not have solutions if $n' \in B'$ *(compare with Definition* 18 *in the special case when all* $\mathbf{k}_{*l} = 0$*). Then under the nonlinear evolution of* (75) *modes with indices* $n' \in B'$ *remain essentially unexcited in the following sense:*

$$\sup_{0\leq\tau\leq\tau_*} |U_{n'}(\tau)| \leq C\varrho \textit{ for all } n' \in B'. \tag{77}$$

Note that $\mathbf{F}(\mathbf{U})$ provides a nonlinear coupling between modes $U_{n_j,\zeta^{(j)}}$ with $n_j \in B$ and $U_{n',\zeta}$ with $n' \in B'$, but the resulting interaction is not $O(1)$ on a fixed time interval $[0, \tau_*]$ as one might expect, but rather of order $O(\upsilon)$ as (77) shows. One way to prove

Theorem 11 is to follow the proofs of Theorems 35 and 37 with obvious modifications and simplifications. In particular, instead of (15) one has to consider the following system with oscillatory coefficients:

$$\partial_\tau \mathbf{u} = \mathrm{e}^{\frac{\mathrm{i}\tau}{\varrho}\mathbf{L}^0}\mathbf{F}\left(\mathrm{e}^{\frac{-\mathrm{i}\tau}{\varrho}\mathbf{L}^0}\mathbf{u}\right), \quad \mathbf{u}(\tau)|_{\tau=0} = \mathbf{h}. \tag{78}$$

Alternatively, Theorem 11 can be derived directly from the classical time averaging principle. Indeed, the time averaging of (78) yields the following averaged system:

$$\partial_\tau \mathbf{v} = \mathbf{F}_{\mathrm{av}}(\mathbf{v}), \quad \mathbf{v}(\tau)|_{\tau=0} = \mathbf{h},$$

where $\mathbf{F}_{\mathrm{av}}$ is defined as in (69), (70) with the frequencies $\phi_j = \omega^0_{j,+}$. From the Krylov-Bogolyubov averaging theorem (see 11,37]) one obtains

$$|\mathbf{v}(\tau) - \mathbf{u}(\tau)| \leq C\varrho, \quad 0 \leq \tau \leq \tau_*.$$

A straightforward examination shows that if B is resonance invariant and $j \in B'$ then the polynomial components $F_{\mathrm{av},j,\zeta}(\mathbf{v})$ factorize into $F_{\mathrm{av},j,\zeta}(\mathbf{v}) = \sum_{j'\in B',\zeta'} F^1_{\mathrm{av},j',\zeta'}(\mathbf{v})\, v_{j',\zeta'}$, implying (77) since $v_{j,\zeta}(0) = 0$ for $j \in B'$.

A stronger *universal resonance invariance* condition in Definition 18 also takes a simpler form in the ODE case. Indeed, let us collect the terms in (76) at different $\omega^0_{j,+}$ as in (101), namely

$$\omega^0_{n',\zeta} - \sum_{j=1}^{m} \omega^0_{n_j,\zeta^{(j)}} = \sum_{i=1}^{J} \delta_i \omega^0_{i,+}, \text{ where } \delta_i \text{ are integers,} \tag{79}$$

Similarly to Definition 18 we call B universally resonance invariant if every solution to the resonance equation (76) must have $n' \in B$ and every coefficient δ_i in (79) for the solution is zero, i.e. $\delta_i = 0$, $i = 1, \ldots, J$. Obviously, if all $\omega^0_{n,+}$ are rationally independent then it is universally resonance invariant.

Now let us look at how universal nonlinearities arise under time averaging. Observe that if the entire set $\{1, \ldots, J\}$ is universally resonance invariant and $F_{j,\zeta}(\mathbf{v})$ are arbitrary polynomials, then the polynomials $F_{\mathrm{av},j,\zeta}(\mathbf{v})$ are obtained by discarding the "resonant" terms in $\mathrm{e}^{\frac{\mathrm{i}\tau}{\varrho}\mathbf{L}^0}\mathbf{F}\left(\mathrm{e}^{\frac{-\mathrm{i}\tau}{\varrho}\mathbf{L}^0}\mathbf{u}\right)$ yielding the universal form (65), (66). For example, if $\mathbf{F}$ is an arbitrary cubic nonlinearity in $\mathbb{C}^{2N}$ then the time averaging yields NLS-like nonlinearity $\mathbf{F}_{\mathrm{av}}$ with components

$$F_{\mathrm{av},j,\zeta}\left(u_{1,+}, u_{1,-}, \ldots, u_{N,+}, u_{N,-}\right) = \sum_{l=1}^{N} b_{l,j,\zeta} u_{l,+} u_{l,-} u_{j,\zeta}.$$

When B is resonance invariant but not universally resonance invariant the averaging produces a weakly universal nonlinearity. A nonlinearity which is weakly universal but not universal may include additional terms, for example the cubic nonlinearity in the classical four-wave interaction system where it is assumed that $\omega^0_{2,-} + \omega^0_{3,+} + \omega^0_{4,+} = \omega^0_{1,+}$ (see 46] p. 201) in the equation for $u_{1,+}$ in addition to NLS-like terms involves the product $u_{2,-}u_{3,+}u_{4,+}$.

3. Conditions and Definitions

In this section we formulate and discuss definitions and conditions under which we study the nonlinear evolutionary system (1) through its modal, Fourier form (3). Most of the conditions and definitions are naturally formulated for the modal form (3), and this is one of the reasons we use it as the basic form.

3.1. Linear part. The basic properties of the linear part $\mathbf{L}(\mathbf{k})$ of the system (3), which is a $2J \times 2J$ Hermitian matrix with eigenvalues $\omega_{n,\zeta}(\mathbf{k})$, has been already discussed in the Introduction. To account for all needed properties of $\mathbf{L}(\mathbf{k})$ we define the singular set of points $\mathbf{k}$.

Definition 12 (Band-crossing points). *We call* $\mathbf{k}_0$ *a* **band-crossing point** *for* $\mathbf{L}(\mathbf{k})$ *if* $\omega_{n+1,\zeta}(\mathbf{k}_0)=\omega_{n,\zeta}(\mathbf{k}_0)$ *for some* n, ζ *or* $\mathbf{L}(\mathbf{k})$ *is not continuous at* $\mathbf{k}_0$ *or if* $\omega_{1,\pm}(\mathbf{k}_0)=0$, *we denote the set of such points by* σ_{bc}.

In the next condition we collect all constraints imposed on the linear operator $\mathbf{L}(\mathbf{k})$.

Condition 13 (Linear part). *The linear part* $\mathbf{L}(\mathbf{k})$ *of the system* (3) *is a* $2J \times 2J$ *Hermitian matrix with eigenvalues* $\omega_{n,\zeta}(\mathbf{k})$ *and corresponding eigenvectors* $\mathbf{g}_{n,\zeta}(\mathbf{k})$ *satisfying for* $\mathbf{k} \notin \sigma_{bc}$ *the basic relations* (5)–(7). *In addition to that we assume:*

(i) *the set of band-crossing points* σ_{bc} *is a closed, nowhere dense set in* $\mathbb{R}^d$ *and has zero Lebesgue measure;*
(ii) *the entries of the Hermitian matrix* $\mathbf{L}(\mathbf{k})$ *are infinitely differentiable in* $\mathbf{k}$ *for all* $\mathbf{k} \notin \sigma_{bc}$ *that readily implies via the spectral theory, 35], infinite differentiability of all eigenvalues* $\omega_n(\mathbf{k})$ *in* $\mathbf{k}$ *for all* $\mathbf{k} \notin \sigma$*;*
(iii) $\mathbf{L}(\mathbf{k})$ *satisfies the polynomial bound*

$$\|\mathbf{L}(\mathbf{k})\| \leq C\left(1+|\mathbf{k}|^{p}\right), \quad \mathbf{k} \in \mathbb{R}^d, \quad \textit{for some } C>0 \textit{ and } p>0. \tag{80}$$

Remark 14 (Dispersion relations symmetry). The symmetry condition (7) on the dispersion relations naturally arises in many physical problems, for example Maxwell equations in periodic media, see 1–3,5], or when $\mathbf{L}(\mathbf{k})$ originates from a Hamiltonian. *We would like to stress that these symmetry conditions are not imposed to simplify studies* but rather to take into account fundamental symmetries of physical media. In fact, the opposite case when ((7) is assumed not to hold is much simpler. The symmetry creates resonant nonlinear interactions, which makes studies more intricate. Interestingly, many problems without symmetries can be put into the framework with symmetry by an extension of the relevant system (see Sect. 4).

Remark 15 (Band-crossing points). Band-crossing points are discussed in more detail in 1, Sect. 5.4], 2, Sects. 4.1, 4.2]. In particular, generically the set σ_{bc} of the band-crossing point is a manifold of the dimension $d-2$. Notice that there is an natural ambiguity in the definition of the normalized eigenvectors $\mathbf{g}_{n,\zeta}(\mathbf{k})$ of $\mathbf{L}(\mathbf{k})$ which is defined up to a complex number ξ with $|\xi|=1$. This ambiguity may not allow an eigenvector $\mathbf{g}_{n,\zeta}(\mathbf{k})$ which can be a locally smooth function in $\mathbf{k}$ to be a uniquely defined continuous function in $\mathbf{k}$ globally for all $\mathbf{k} \notin \sigma_{bc}$ because of a possibility of branching. But, importantly, the orthogonal projector $\Pi_{n,\zeta}(\mathbf{k})$ on $\mathbf{g}_{n,\zeta}(\mathbf{k})$ as defined by (11) is uniquely defined and, consequently, infinitely differentiable in $\mathbf{k}$ via the spectral theory, 35], for all $\mathbf{k} \notin \sigma_{bc}$. Since we consider $\hat{\mathbf{U}}(\mathbf{k})$ as an element of the space L^1 and σ_{bc} is of zero Lebesgue measure considering $\mathbf{k} \notin \sigma_{bc}$ is sufficient for us.

We introduce for vectors $\hat{\mathbf{u}} \in \mathbb{C}^{2J}$ their expansion with respect to the orthonormal basis$\left\{\mathbf{g}_{n,\zeta}(\mathbf{k})\right\}$:

$$\hat{\mathbf{u}}(\mathbf{k}) = \sum_{n=1}^{J} \sum_{\zeta=\pm} \hat{u}_{n,\zeta}(\mathbf{k})\, \mathbf{g}_{n,\zeta}(\mathbf{k}) = \sum_{n=1}^{J} \sum_{\zeta=\pm} \hat{\mathbf{u}}_{n,\zeta}(\mathbf{k}), \quad \hat{\mathbf{u}}_{n,\zeta}(\mathbf{k}) = \Pi_{n,\zeta}(\mathbf{k})\, \hat{\mathbf{u}}(\mathbf{k}), \tag{81}$$

and we refer to it as the *modal decomposition* of $\hat{\mathbf{u}}(\mathbf{k})$ and to $\hat{u}_{n,\zeta}(\mathbf{k})$ as the *modal coefficients* of $\hat{\mathbf{u}}(\mathbf{k})$. Evidently

$$\sum_{n=1}^{j} \sum_{\zeta=\pm} \Pi_{n,\zeta}(\mathbf{k}) = I_{2J}, \quad \text{where } I_{2J} \text{ is the } 2J \times 2J \text{ identity matrix.} \tag{82}$$

Notice that in view of the polynomial bound 80) we can define the action of the operator $\mathbf{L}(-\mathrm{i}\nabla_{\mathbf{r}})$ on any Schwartz function $\mathbf{Y}(\mathbf{r})$ by the formula

$$\mathbf{L}\,\widehat{(-\mathrm{i}\nabla_{\mathbf{r}})\,\mathbf{Y}}(\mathbf{k}) = \mathbf{L}(\mathbf{k})\,\hat{\mathbf{Y}}(\mathbf{k}), \quad \text{where the order of } \mathbf{L} \text{ does not exceed } p. \tag{83}$$

In a special case when all the entries of $\mathbf{L}(\mathbf{k})$ are polynomials (83) turns into the action of the differential operator with constant coefficients of order not exceeding p.

3.2. Nonlinear part. The nonlinear term $\hat{F}$ in (3) is assumed to be a general functional polynomial of the form

$$\hat{F}\left(\hat{\mathbf{U}}\right) = \sum_{m \in \mathfrak{M}_F} \hat{F}^{(m)}\left(\hat{\mathbf{U}}^m\right), \quad \text{where } \hat{F}^{(m)} \text{ is } m\text{-homogeneous polylinear operator,} \tag{84}$$

$$\mathfrak{M}_F = \left\{m_1, \ldots, m_p\right\} \subset \{2, 3, \ldots\} \text{ is a finite set, and } m_F = \max\{m : m \in \mathfrak{M}_F\}. \tag{85}$$

The integer m_F in (85) is called the *degree of the functional polynomial* $\hat{F}$. For instance, if $\mathfrak{M}_F = \{2\}$ or $\mathfrak{M}_F = \{3\}$ the polynomial $\hat{F}$ is respectively homogeneous quadratic or cubic. Every m-linear operator $\hat{F}^{(m)}$ in (84) is assumed to be of the form of a convolution

$$\hat{F}^{(m)}\left(\hat{\mathbf{U}}_1, \ldots, \hat{\mathbf{U}}_m\right)(\mathbf{k}, \tau) = \int_{\mathbb{D}_m} \chi^{(m)}\left(\mathbf{k}, \vec{k}\right) \hat{\mathbf{U}}_1\left(\mathbf{k}'\right) \ldots \hat{\mathbf{U}}_m\left(\mathbf{k}^{(m)}\left(\mathbf{k}, \vec{k}\right)\right) \tilde{\mathrm{d}}^{(m-1)d}\vec{k}, \tag{86}$$

$$\text{where } \mathbb{D}_m = \mathbb{R}^{(m-1)d}, \quad \tilde{\mathrm{d}}^{(m-1)d}\vec{k} = \frac{\mathrm{d}\mathbf{k}' \ldots \mathrm{d}\mathbf{k}^{(m-1)}}{(2\pi)^{(m-1)d}},$$
$$\mathbf{k}^{(m)}\left(\mathbf{k}, \vec{k}\right) = \mathbf{k} - \mathbf{k}' - \ldots - \mathbf{k}^{(m-1)}, \quad \vec{k} = \left(\mathbf{k}', \ldots, \mathbf{k}^{(m)}\right), \tag{87}$$

indicating that the nonlinear operator $F^{(m)}(\mathbf{U}_1, \ldots, \mathbf{U}_m)$ is translation invariant (it may be local or non-local). The quantities $\chi^{(m)}$ in (86) are called *susceptibilities*. For numerous examples of nonlinearities of the form similar to (84), (86) see 1–7] and references therein. In what follows the nonlinear term $\hat{F}$ in (3) will satisfy the following conditions.

Condition 16 (Nonlinearity). *The nonlinearity $\hat{F}\left(\hat{\mathbf{U}}\right)$ is assumed to be of the form* (84)–(86). *The susceptibility $\chi^{(m)}\left(\mathbf{k},\mathbf{k}',\ldots,\mathbf{k}^{(m)}\right)$ is infinitely differentiable for all $\mathbf{k}$ and $\mathbf{k}^{(j)}$ which are not band-crossing points, and is bounded, namely*

$$\left\|\chi^{(m)}\right\| = (2\pi)^{-(m-1)d} \sup_{\mathbf{k},\mathbf{k}',\ldots,\mathbf{k}^{(m)}\in\mathbb{R}^d\setminus\sigma_{bc}} \left|\chi^{(m)}\left(\mathbf{k},\mathbf{k}',\ldots,\mathbf{k}^{(m)}\right)\right| \le C_\chi, \quad m\in\mathfrak{M}_F, \tag{88}$$

where the norm $\left|\chi^{(m)}\left(\mathbf{k},\vec{k}\right)\right|$ of the m-linear tensor $\chi^{(m)}:\left(\mathbb{C}^{2J}\right)^m\to\left(\mathbb{C}^{2J}\right)^m$ for fixed $\mathbf{k},\vec{k}$ is defined by

$$\left|\chi^{(m)}\left(\mathbf{k},\vec{k}\right)\right| = \sup_{|\mathbf{x}_j|\le 1}\left|\chi^{(m)}\left(\mathbf{k},\vec{k}\right)(\mathbf{x}_1,\ldots,\mathbf{x}_m)\right|, \quad \text{where } |\mathbf{x}| \text{ is the Euclidean norm.} \tag{89}$$

When $\chi^{(m)}\left(\mathbf{k},\vec{k}\right)$ depend on small ϱ or, more generally, on ϱ^q, $q>0$, we similarly have $\chi^{(m)}\left(\mathbf{k},\vec{k},\varrho^q\right)$. Many results of this paper extend to this case, in particular if $\left\|\chi^{(m)}\left(\mathbf{k},\vec{k},\varrho^q\right)-\chi^{(m)}\left(\mathbf{k},\vec{k},0\right)\right\|\le C'_\chi\varrho^q$ for $\varrho\le 1$ then conditions of Corollary 38 are fulfilled.

Note that since the tensors $\chi^{(m)}\left(\mathbf{k},\vec{k}\right)$ are bounded, the dependence on $\left(\mathbf{k},\vec{k}\right)$ cannot be polynomial, therefore the original equation (1) does not include spatial derivatives but rather includes bounded "pseudodifferential" operators. Note that this type of susceptibilities with spatial dispersion is common in nonlinear optics, see 15,41,55].

3.3. Resonance invariant nk-spectrum. In this section, relying on given dispersion relations $\omega_n(\mathbf{k})\ge 0, n\in\{1,\ldots,J\}$, we consider resonance properties of nk-spectra S and the corresponding k-spectra K_S as defined in Definition 2, i.e.

$$S=\{(n_l,\mathbf{k}_{*l}),\ l=1,\ldots,N\}\subset\Sigma=\{1,\ldots,J\}\times\mathbb{R}^d,\ K_S=\left\{\mathbf{k}_{*l_i},\ i=1,\ldots,|K_S|\right\}. \tag{90}$$

We precede the formal description of the *resonance invariance* (see Definition 18) with the following guiding physical picture. Initially at $\tau=0$ the wave is a multi-wavepacket composed of modes from a small vicinity of the nk-spectrum S. As the wave evolves according to (3) the polynomial nonlinearity inevitably involves a larger set of modes $[S]_{\text{out}}\supseteq S$, but not all modes in $[S]_{\text{out}}$ are "equal" in developing significant amplitudes. The qualitative picture is that whenever a certain interaction phase function (see (134) below) is not zero, the fast time oscillations weaken effective nonlinear mode interaction and the energy transfer from the original modes in S to relevant modes from $[S]_{\text{out}}$, keeping their magnitudes vanishingly small as $\beta,\varrho\to 0$. There is a smaller set of modes $[S]_{\text{out}}^{\text{res}}$ which can interact with modes from S rather effectively and develop significant amplitudes. Now,

$$\text{if } [S]_{\text{out}}^{\text{res}}\subseteq S \text{ then } S \text{ is called resonance invariant.} \tag{91}$$

In simpler situations the resonance invariance conditions turn into the well-known in nonlinear optics phase and frequency matching conditions. For instance, if S contains $\left(n_0,\mathbf{k}_{*l_0}\right)$ and the dispersion relations allow for the second harmonic generation in another band n_1 so that $2\omega_{n_0}\left(\mathbf{k}_{*l_0}\right) = \omega_{n_1}\left(2\mathbf{k}_{*l_0}\right)$, then for S to be resonance invariant it must contain $\left(n_1, 2\mathbf{k}_{*l_0}\right)$ too.

Let us turn now to the rigorous constructions. First we introduce necessary notations. Let $m \geq 2$ be an integer, $\vec{l} = (l_1, .., l_m)$, $l_j \in \{1, \ldots, N\}$ be an integer vector from $\{1, \ldots, N\}^m$ and $\vec{\zeta} = \left(\zeta^{(1)}, , .., \zeta^{(m)}\right)$, $\zeta^{(j)} \in \{+1, -1\}$ be a binary vector from $\{+1, -1\}^m$. Note that a pair $\left(\vec{\zeta}, \vec{l}\right)$ naturally labels a sample string of the length m composed of elements $\left(\zeta^{(j)}, n_{l_j}, \mathbf{k}_{*l_j}\right)$ from the set $\{+1, -1\} \times S$. Let us introduce the sets

$$\Lambda = \{(\zeta, l) : l \in \{1, \ldots, N\}, \quad \zeta \in \{+1, -1\}\},$$
$$\Lambda^m = \left\{\vec{\lambda} = (\lambda_1, \ldots, \lambda_m), \quad \lambda_j \in \Lambda, \quad j = 1, \ldots, m\right\}. \tag{92}$$

There is a natural one-to-one correspondence between Λ^m and $\{-1, 1\}^m \times \{1, \ldots, N\}^m$ and we will write, exploiting this correspondence

$$\vec{\lambda} = \left((\zeta', l_1), \ldots, \left(\zeta^{(m)}, l_m\right)\right) = \left(\vec{\zeta}, \vec{l}\right), \quad \vec{\vartheta} \in \{-1, 1\}^m,\ \vec{l} \in \{1, \ldots, N\}^m \text{ for } \vec{\lambda} \in \Lambda^m. \tag{93}$$

Let us introduce the following linear combination:

$$\varkappa_m\left(\vec{\lambda}\right) = \varkappa_m\left(\vec{\zeta}, \vec{l}\right) = \sum_{j=1}^{m} \zeta^{(j)} \mathbf{k}_{*l_j} \text{ with } \zeta^{(j)} \in \{+1, -1\}, \tag{94}$$

and let $[S]_{K,\text{out}}$ be the set of all its values as $\mathbf{k}_{*l_j} \in K_S$, $\vec{\lambda} \in \Lambda^m$, namely

$$[S]_{K,\text{out}} = \bigcup\nolimits_{m \in \mathfrak{M}_F} \bigcup\nolimits_{\vec{\lambda} \in \Lambda^m} \left\{\varkappa_m\left(\vec{\lambda}\right)\right\}. \tag{95}$$

We call $[S]_{K,\text{out}}$ the *output k-spectrum* of K_S. Everywhere in this paper we consider *nk*-spectra S which satisfy the following condition:

$$[S]_{K,\text{out}} \bigcap \sigma_{bc} = \varnothing. \tag{96}$$

We also define the *output nk-spectrum* of S by

$$[S]_{\text{out}} = \left\{(n, \mathbf{k}) \in \{1, \ldots, J\} \times \mathbb{R}^d : n \in \{1, \ldots, J\}, \quad \mathbf{k} \in [S]_{K,\text{out}}\right\}. \tag{97}$$

We introduce the following functions:

$$\Omega_{1,m}\left(\vec{\lambda}\right)\left(\vec{k}_*\right) = \sum\nolimits_{j=1}^{m} \zeta^{(j)} \omega_{l_j}\left(\mathbf{k}_{*l_j}\right), \quad \vec{k}_* = \left(\mathbf{k}_{*1}, \ldots, \mathbf{k}_{*|K_S|}\right), \quad \text{where } \mathbf{k}_{*l_j} \in K_S, \tag{98}$$

$$\Omega\left(\zeta, n, \vec{\lambda}\right)\left(\mathbf{k}_{**}, \vec{k}_*\right) = -\zeta\omega_n\left(\mathbf{k}_{**}\right) + \Omega_{1,m}\left(\vec{\lambda}\right)\left(\vec{k}_*\right), \tag{99}$$

where $\zeta = \pm 1$, $m \in \mathfrak{M}_F$ as in (84). We introduce these functions to apply later to phase functions (134).

Now we introduce the *resonance equation*

$$\Omega\left(\zeta, n, \vec{\lambda}\right)\left(\zeta\varkappa_m\left(\vec{\lambda}\right), \vec{k}_*\right) = 0, \quad \vec{l} \in \{1, \ldots, N\}^m, \quad \vec{\zeta} \in \{-1, 1\}^m, \tag{100}$$

denoting by $P(S)$ the set of its solutions $\left(m, \zeta, n, \vec{\lambda}\right)$. Such a solution is called S-*internal* if

$$\left(n, \zeta \varkappa_m\left(\vec{\lambda}\right)\right) \in S, \text{ that is } n = n_{l_0}, \quad \zeta \varkappa_m\left(\vec{\lambda}\right) = \mathbf{k}_{*l_0}, \quad l_0 \in \{1, \ldots, N\},$$

and we denote the corresponding $l_0 = I\left(\vec{\lambda}\right)$. We also denote by $P_{\text{int}}(S) \subset P(S)$ the set of all S-internal solutions to (100).

Now we consider the simplest solutions to (100) which play an important role. Keeping in mind that the string $\vec{l}$ can contain several copies of a single value l, we can recast the sum in (98) as follows:

$$\begin{aligned}
&\Omega_{1,m}\left(\vec{\lambda}\right) = \Omega_{1,m}\left(\vec{\zeta}, \vec{l}\right) = \sum_{l=1}^{N} \delta_l \omega_l\left(\mathbf{k}_{*l}\right), \quad \text{where } \delta_l \\
&= \begin{cases} \sum_{j \in \vec{l}^{-1}(l)} \zeta^{(j)} & \text{if } \vec{l}^{-1}(l) \neq \varnothing \\ 0 & \text{if } \vec{l}^{-1}(l) = \varnothing \end{cases}, \\
&\vec{l}^{-1}(l) = \left\{j : l_j = l, \quad 1 \le j \le m\right\}, \quad \vec{l} = (l_1, \ldots, l_m), \quad 1 \le l \le N.
\end{aligned} \tag{101}$$

Let us call a solution $\left(m, \zeta, n, \vec{\lambda}\right) \in P(S)$ of (100) *universal* if it has the following properties: (i) only a single coefficient out of all δ_l in (101) is nonzero, namely for some I_0 we have $\delta_{I_0} = \pm 1$ and $\delta_l = 0$ for $l \neq I_0$; (ii) $n = n_{I_0}$ and $\zeta = \delta_{I_0}$. *A justification for calling such a solution universal comes from the fact that if it is a solution for one $\vec{k}_*$ it is a solution for any other $\vec{k}_* \in \mathbb{R}^d$*. We denote the set of universal solutions to (100) by $P_{\text{univ}}(S)$, and note that a universal solution is a S -internal solution with $I\left(\vec{\lambda}\right) = I_0$ implying

$$P_{\text{univ}}(S) \subseteq P_{\text{int}}(S). \tag{102}$$

Indeed, observe that for δ_l as in (101),

$$\varkappa_m\left(\vec{\lambda}\right) = \varkappa_m\left(\vec{\zeta}, \vec{l}\right) = \sum_{j=1}^{m} \zeta^{(j)} \mathbf{k}_{*l_j} = \sum_{l=1}^{N} \delta_l \mathbf{k}_{*l}, \tag{103}$$

implying $\varkappa_m\left(\vec{\lambda}\right) = \delta_{I_0} \mathbf{k}_{*I_0}$ and $\zeta \varkappa_m\left(\vec{\lambda}\right) = \delta_{I_0}^2 \mathbf{k}_{*I_0} = \mathbf{k}_{*I_0}$. Then Eq. (100) is obviously satisfied and $\left(n, \zeta \varkappa_m\left(\vec{\lambda}\right)\right) = \left(n_{I_0}, \mathbf{k}_{*I_0}\right) \in S$.

Example 17 (Universal solutions). Suppose there is just a single band, i.e. $J = 1$, a symmetric dispersion relation $\omega_1(-\mathbf{k}) = \omega_1(\mathbf{k})$, a cubic nonlinearity F with $\mathfrak{M}_F = \{3\}$. First let us take the simplest nk -spectrum $S_1 = \{(1, \mathbf{k}_*)\}$, that is $N = 1$. Then $\Omega_{1,3}\left(\vec{\lambda}\right)\left(\vec{k}_*\right) = \delta_1 \omega_1(\mathbf{k}_*)$ and $\varkappa_m\left(\vec{\lambda}\right) = \delta_1 \mathbf{k}_*$ where we use notation (101). The universal solution set has the form $P_{\text{univ}}(S_1) = \left\{\left(3, \zeta, 1, \vec{\lambda}\right) : \vec{\lambda} \in \Lambda_\zeta, \ \zeta = \pm\right\}$, where Λ_+ consists of vectors $(\lambda_1, \lambda_2, \lambda_3)$ of the form $((-, 1), (+, 1), (+, 1))$, $((+, 1), (-, 1), (+, 1))$ and $((+, 1), (+, 1), (-, 1))$. Obviously, $P_{\text{univ}}(S_1) = P_{\text{int}}(S_1)$. In the next example we take the nk-spectrum $S = \{(1, \mathbf{k}_*), (1, -\mathbf{k}_*)\}$, that is $N = 2$ and $\mathbf{k}_{*1} = \mathbf{k}_*, \mathbf{k}_{*2} = -\mathbf{k}_*$. This example is typical for two counterpropagating waves. Then $\Omega_{1,3}\left(\vec{\lambda}\right)\left(\vec{k}_*\right) = \sum_{j=1}^{3} \zeta^{(j)} \omega_{l_j}\left(\mathbf{k}_{*l_j}\right) = (\delta_1 + \delta_2)\, \omega_1(\mathbf{k}_*)$ and $\varkappa_m\left(\vec{\lambda}\right) = \sum_{j=1}^{m} \zeta^{(j)}$

$\mathbf{k}_{*l_j} = \delta_1 \mathbf{k}_{*1} + \delta_2 \mathbf{k}_{*2} = (\delta_1 - \delta_2)\, \mathbf{k}_*$ where we use notation (101). The universal solution set has the form $P_{\text{univ}}(S) = \left\{ \left(3, \zeta, 1, \vec{\lambda}\right) : \vec{\lambda} \in \Lambda_\zeta, \zeta = \pm \right\}$, where Λ_+ consists of vectors $(\lambda_1, \lambda_2, \lambda_3)$ of the form $((+,1),(-,1),(+,1))$, $((+,1),(-,1),(+,2))$, $((+,2),(-,2),$
$(+,1))$, $((+,2),(-,2),(+,2))$, and vectors obtained from the listed ones by permutations of coordinates $\lambda_1, \lambda_2, \lambda_3$. The solutions from $P_{\text{int}}(S)$ have to satisfy $|\delta_1 - \delta_2| = 1$ and $|\delta_1 + \delta_2| = 1$ which is possible only if $\delta_1 \delta_2 = 0$. Since $\zeta = \delta_1 + \delta_2$ we have $\zeta \varkappa_m \left(\vec{\lambda}\right) = \left(\delta_1^2 - \delta_2^2\right) \mathbf{k}_*$ and $\zeta \varkappa_m \left(\vec{\lambda}\right) = \mathbf{k}_{*1}$ if $|\delta_1| = 1$ or $\zeta \varkappa_m \left(\vec{\lambda}\right) = \mathbf{k}_{*2}$ if $|\delta_2| = 1$. Hence $P_{\text{int}}(S) = P_{\text{univ}}(S)$ in this case. Note that if we set $S_2 = \{(1, -\mathbf{k}_*)\}$ then $S = S_1 \cup S_2$ but $P_{\text{int}}(S)$ is larger than $P_{\text{int}}(S_1) \cup P_{\text{int}}(S_2)$. This can be interpreted as follows. When only modes from S_1 are excited, the modes from S_2 remain non-excited. But when both S_1 and S_2 are excited, there is a resonance effect of S_1 onto S_2, represented, for example, by $\vec{\lambda} = ((+,1),(-,1),(+,2))$, which involves the mode $\zeta \varkappa_m \left(\vec{\lambda}\right) = \mathbf{k}_{*2}$.

Now we are ready to define resonance invariant spectra. First, we introduce a subset $[S]_{\text{out}}^{\text{res}}$ of $[S]_{\text{out}}$ by the formula

$$[S]_{\text{out}}^{\text{res}} = \left\{ (n, \mathbf{k}_{**}) \in [S]_{\text{out}} : \mathbf{k}_{**} = \zeta \varkappa_m \left(\vec{\lambda}\right), \quad m \in \mathfrak{M}_F, \text{ where} \right. \\ \left. \left(m, \zeta, n, \vec{\lambda}\right) \text{ is a solution of (100)} \right\}, \tag{104}$$

calling it *resonant output spectrum* of S, and then we define

$$\text{resonance selection operation } \mathcal{R}(S) = S \cup [S]_{\text{out}}^{\text{res}}. \tag{105}$$

Definition 18 (Resonance invariant nk-spectrum). *The nk-spectrum S is called* **resonance invariant** *if $\mathcal{R}(S) = S$ or, equivalently, $[S]_{out}^{res} \subseteq S$. The nk-spectrum S is called* **universally resonance invariant** *if $\mathcal{R}(S) = S$ and $P_{univ}(S) = P_{int}(S)$.*

It is worth noticing that even when a nk-spectrum is not resonance invariant often it can be easily extended to a resonance invariant one. Namely, if $\mathcal{R}^j(S) \cap \sigma_{bc} = \varnothing$ for all j then the set

$$\mathcal{R}^\infty(S) = \bigcup_{j=1}^{\infty} \mathcal{R}^j(S) \subset \Sigma = \{1, \ldots, J\} \times \mathbb{R}^d$$

is resonance invariant. In addition to that, $\mathcal{R}^\infty(S)$ is always at most countable. Usually it is finite, i.e. $\mathcal{R}^\infty(S) = \mathcal{R}^p(S)$ for a finite p, see examples below and we also show below that $\mathcal{R}^\infty(S) = S$ for generic K_S.

Example 19 (*Resonance invariant nk-spectra for quadratic nonlinearity*). Suppose there is a single band, i.e. $J = 1$,with a symmetric dispersion relation, and a quadratic nonlinearity F, that is $\mathfrak{M}_F = \{2\}$. Let us assume that $\mathbf{k}_* \neq 0$, $\mathbf{k}_*, 2\mathbf{k}_*, \mathbf{0}$ are not band-crossing points and look at two examples. First, suppose that $2\omega_1(\mathbf{k}_*) \neq \omega_1(2\mathbf{k}_*)$ (no second harmonic generation) and $\omega_1(\mathbf{0}) \neq 0$. Let us set the nk-spectrum to be the set $S_1 = \{(1, \mathbf{k}_*)\}$, then S_1 is resonance invariant. Indeed, $K_{S_1} = \{\mathbf{k}_*\}$, $[S_1]_{K,\text{out}} = \{\mathbf{0}, 2\mathbf{k}_*, -2\mathbf{k}_*\}$, $[S_1]_{\text{out}} = \{(1,\mathbf{0}), (1, 2\mathbf{k}_*), (1, -2\mathbf{k}_*)\}$ and an elementary examination shows that $[S_1]_{\text{out}}^{\text{res}} = \varnothing \subset S_1$ implying $\mathcal{R}(S_1) = S_1$. For the second example let us assume $\omega_1(\mathbf{0}) \neq 0$ and $2\omega_1(\mathbf{k}_*) = \omega_1(2\mathbf{k}_*)$, that is the second harmonic generation is allowed. Here $[S_1]_{\text{out}}^{\text{res}} = \{(1, 2\mathbf{k}_*)\}$ and $\mathcal{R}(S_1) = \{(1, \mathbf{k}_*), (1, 2\mathbf{k}_*)\}$ implying $\mathcal{R}(S_1) \neq S_1$ and, hence, S_1 is not resonance invariant. Suppose now that $4\mathbf{k}_*, 3\mathbf{k}_* \notin \sigma_{bc}$ and $\omega_1(\mathbf{0}) \neq 0$, $\omega_1(4\mathbf{k}_*) \neq 2\omega_1(2\mathbf{k}_*)$, $\omega_1(3\mathbf{k}_*) \neq \omega_1(\mathbf{k}_*) + \omega_1(2\mathbf{k}_*)$ and let us set $S_2 = \{(1, \mathbf{k}_*), (1, 2\mathbf{k}_*)\}$. An elementary examination shows that S_2 is resonance invariant. Note that S_2 can be obtained by iterating the resonance selection operator, namely $S_2 = \mathcal{R}(\mathcal{R}(S_1))$. Note also that $P_{\text{univ}}(S_2) \neq P_{\text{int}}(S_2)$. Notice that $\omega_1(\mathbf{0}) = 0$ is a special case since $\mathbf{k} = \mathbf{0}$ is a band-crossing point, and it requires a special treatment.

Example 20 (*Resonance invariant nk-spectra for cubic nonlinearity*). Let us consider the one-band case with a symmetric dispersion relation and a cubic nonlinearity that is $\mathfrak{M}_F = \{3\}$. First we take $S_1 = \{(1, \mathbf{k}_*)\}$, we assume that $\mathbf{k}_*, 3\mathbf{k}_*$ are not band-crossing points, implying $[S_1]_{K,\text{out}} = \{\mathbf{k}_*, -\mathbf{k}_*, 3\mathbf{k}_*, -3\mathbf{k}_*\}$. We have $\Omega_{1,3}\left(\vec{\lambda}\right)\left(\vec{k}_*\right) = \sum_{j=1}^{3} \zeta^{(j)}\omega_1(\mathbf{k}_*) = \delta_1\omega_1(\mathbf{k}_*)$ and $\varkappa_m\left(\vec{\lambda}\right) = \delta_1\mathbf{k}_*$, where we use notation (101), δ_1 takes values $1, -1, 3, -3$. If $3\omega_1(\mathbf{k}_*) \neq \omega_1(3\mathbf{k}_*)$, then (100) has a solution only if $|\delta_1| = 1$ and $\delta_1 = \zeta$, hence $\zeta\varkappa_m\left(\vec{\lambda}\right) = \mathbf{k}_*$ and every solution is internal. Therefore, $[S_1]_{\text{out}}^{\text{res}} = \varnothing$ and $\mathcal{R}(S_1) = S_1$. Now consider the case associated with the third harmonic generation, namely $3\omega_1(\mathbf{k}_*) = \omega_1(3\mathbf{k}_*)$ and assume that $\omega_1(3\mathbf{k}_*) + 2\omega_1(\mathbf{k}_*) \neq \omega_1(5\mathbf{k}_*)$, $3\omega_1(3\mathbf{k}_*) \neq \omega_1(9\mathbf{k}_*)$, $2\omega_1(3\mathbf{k}_*) + \omega_1(\mathbf{k}_*) \neq \omega_1(7\mathbf{k}_*)$, $2\omega_1(3\mathbf{k}_*) - \omega_1(\mathbf{k}_*) \neq \omega_1(5\mathbf{k}_*)$. An elementary examination shows that the set $S_4 = \{(1, 3\mathbf{k}_*), (1, \mathbf{k}_*), (1, -\mathbf{k}_*)\ (1, -3\mathbf{k}_*)\}$ satisfies $\mathcal{R}(S_4) = S_4$. Consequently, a multiwavepacket having S_4 as its resonance invariant nk-spectrum involves the third harmonic generation and, according to Theorem 3, it is preserved under nonlinear evolution.

The above examples indicate that in simple cases the conditions on $\mathbf{k}_*$ which can make S non-invariant with respect to $\mathcal{R}$ have a form of several algebraic equations, therefore, for almost all $\mathbf{k}_*$ such spectra S are resonance invariant. The examples also show that if we fix S and dispersion relations then we can include S in the larger spectrum $S' = \mathcal{R}^p(S)$ using repeated application of the operation $\mathcal{R}$ to S, and often the resulting extended nk-spectrum S' is resonance invariant. We show in the following section that nk-spectrum S with generic K_S is universally resonance invariant.

Note that the concept of resonance invariant nk-spectrum gives a mathematical description of such fundamental concepts of nonlinear optics as phase matching, frequency matching, four wave interaction in cubic media and three wave interaction in quadratic media. If a multi-wavepacket has a resonance invariant spectrum, all these phenomena may take place in the internal dynamics of the multi-wavepacket, but do not lead to resonant interactions with continuum of all remaining modes.

3.4. Genericity of the nk-spectrum invariance condition. In simpler situations, when the number of bands J and wavepackets N are not too large, the resonance invariance of nk- spectrum can be easily verified as above in Examples 19, 20, but what one can say if J or N are large, or if the dispersion relations are not explicitly given? We show below that in properly defined non-degenerate cases a small variation of K_S makes S universally resonance invariant, i.e. the resonance invariance is a generic phenomenon.

Assume that the dispersion relations $\omega_n(\mathbf{k}) \geq 0$, $n \in \{1, \ldots, J\}$ are given. Observe then that $\Omega_m\left(\zeta, n, \vec{\lambda}\right) = \Omega_m\left(\zeta, n, \vec{\lambda}\right)\left(\mathbf{k}_{*1}, \ldots, \mathbf{k}_{*|K_S|}\right)$ defined by (99) is a continuous function of $\mathbf{k}_{*l} \notin \sigma_{bc}$ for every $m, \zeta, n, \vec{\lambda}$.

Definition 21 **(ω-degenerate dispersion relations).** *We call dispersion relations $\omega_n(\mathbf{k})$, $n = 1, \ldots, J$, ω-degenerate if there exists such a point $\mathbf{k}_* \in \mathbb{R}^d \setminus \sigma_{bc}$ that for all $\mathbf{k}$ in a neighborhood of $\mathbf{k}_*$ at least one of the following four conditions holds: (i) the relations are linearly dependent, namely $\sum_{n=0}^{J} C_n \omega_n(\mathbf{k}) = c_0$, where all C_n are integers, one of which is nonzero, and the c_0 is a constant; (ii) at least one of $\omega_n(\mathbf{k})$ is a linear function; (iii) at least one of $\omega_n(\mathbf{k})$ satisfies equation $C\omega_n(\mathbf{k}) = \omega_n(C\mathbf{k})$ with some n and integer $C \neq \pm 1$; (iv) at least one of $\omega_n(\mathbf{k})$ satisfies equation $\omega_n(\mathbf{k}) = \omega_{n'}(-\mathbf{k})$, where $n' \neq n$.*

Note that fulfillment of any of the four conditions in Definition 21 makes it impossible to turn some non-resonance invariant sets into resonance invariant ones by a variation of $\mathbf{k}_{*l}$. For instance, if $\mathfrak{M}_F = \{2\}$ as in Example 19 and $2\omega_1(\mathbf{k}) = \omega_1(2\mathbf{k})$ for all $\mathbf{k}$ in an open set G then the set $\{(1, \mathbf{k}_*)\}$ with $\mathbf{k}_* \in G$ cannot be made resonance invariant by a small variation of $\mathbf{k}_*$. Below we show that if dispersion relations are not ω -degenerate, then a small variation of $\mathbf{k}_{*l}$ turns non-resonance invariant sets into resonance invariant.

Theorem 22. *If $\Omega_m\left(\zeta, n_0, \vec{\lambda}\right)\left(\mathbf{k}'_{*1}, \ldots, \mathbf{k}'_{*|K_S|}\right) = 0$ on a cylinder G in $\left(\mathbb{R}^d \setminus \sigma_{bc}\right)^{|K_S|}$ which is a product of small balls $G_i \subset \left(\mathbb{R}^d \setminus \sigma_{bc}\right)$, then either $\left(m, \zeta, n_0, \vec{\lambda}\right) \in P_{univ}(S)$ or dispersive relations $\omega_n(\mathbf{k})$ are ω-degenerate as in Definition 21.*

Proof. Collecting similar terms in (100) we obtain the following equation for $\mathbf{k}_i$ from G_i:

$$\sum_{n=1}^{J} \sum_{i=1}^{|K_S|} \delta'_{in} \omega_n(\mathbf{k}_i) = \zeta \omega_{n_0}\left(\sum_{i=1}^{|K_S|} \delta'_i \mathbf{k}_i\right) \text{ where } \delta'_{in}, \delta'_i \text{ are integers.} \tag{106}$$

Comparing (106) with (101) we see that δ'_{in} may be non-zero only if $(n, \mathbf{k}_i) \in S$, that is $(n, \mathbf{k}_i) = (n_l, \mathbf{k}_l)$ with $l \in \{1, \ldots, N\}$, where $l = l(i, n)$ is uniquely determined and $\delta'_{in} = \delta_l$ with δ_l as in (101). If there are two nonzero coefficients δ_i in (106) we use an elementary Proposition 24 below, noticing that we are in case (ii) of Definition 21. If we do not have two nonzero δ'_i then either all $\delta'_i = 0$ or only one $\delta'_i = \delta'_{i_0} \neq 0$. If all $\delta'_i = 0$ then the right-hand side of (106) turns into $\omega_{n_0}(0)$ and, $G_i \subset \left(\mathbb{R}^d \setminus \sigma_{bc}\right)$, $\omega_{n_0}(0) \neq 0$. Hence, for every i the sum $\sum_{n=1}^{J} \delta'_{in} \omega_n(\mathbf{k}_i)$ is constant, one of δ'_{in} is non-zero and we are in case (i) of Definition 21. If only one $\delta'_i \neq 0$ with $i = i_0$ we have

$$\sum_{n=1}^{J} \sum_{i=1}^{|K_S|} \delta'_{in} \omega_n(\mathbf{k}_i) = \zeta \omega_{n_0}\left(\delta'_{i_0} \mathbf{k}_{i_0}\right) \text{ for all } \mathbf{k}_i \in G_i, \quad \mathbf{k}_{i_0} \in G_{i_0}, \tag{107}$$

implying linear dependence of the dispersion relations, namely

$$\sum_{n=1}^{J} \delta'_{in} \omega_n(\mathbf{k}_i) = C_i, \quad i \neq i_0, \text{ where } C_i \text{ are constant.}$$

The above equations would not imply linear dependence as in case (i) of Definition 21 only if

$$\delta'_{in} = 0, \quad i \neq i_0, \quad n = 1, \ldots, J, \tag{108}$$

and in this case the equality (107) takes the form

$$\sum_{n=1}^{J} \delta'_{i_0 n} \omega_n \left(\mathbf{k}_{i_0}\right) = \zeta \omega_{n_0} \left(\delta'_{i_0} \mathbf{k}_{i_0}\right) \text{ for all } \mathbf{k}_{i_0} \in G_{i_0}. \tag{109}$$

Note that in this case we deduce from (94) and (98) that $\sum_{n=1}^{J} \delta'_{i_0 n} = \delta'_{i_0}$. If $\left|\delta'_{i_0}\right| \neq 1$ we are in case (iii) of Definition 21, whereas if $\left|\delta'_{i_0}\right| = 1$ and $n \neq n_0$ we are in case (iv) of Definition 21. If $\left|\delta'_{i_0}\right| = 1$ and $n = n_0$ (109) turns into $\delta'_{i_0} \omega_{n_0} \left(\mathbf{k}_{i_0}\right) = \zeta \omega_{n_0} \left(\delta'_{i_0} \mathbf{k}_{i_0}\right)$. Since $\omega_{n_0} > 0$ it implies $\delta'_{i_0} = \zeta$ and $\omega_{n_0} \left(\mathbf{k}_{i_0}\right) = \omega_{n_0} \left(\zeta \delta'_{i_0} \mathbf{k}_{i_0}\right)$. Hence, in this case $\left(m, \zeta, n_0, \vec{\lambda}\right) \in P_{\text{univ}}(S)$, and since all possibilities are exhausted the proof is complete. □

Theorem 23 (Genericity of resonance invariance). *Assume that dispersive relations $\omega_n(\mathbf{k})$ are continuous and not ω-degenerate as in Definition 21. Let $\mathcal{K}_{rinv}$ be a set of points $\left(\mathbf{k}_{*1}, \ldots, \mathbf{k}_{*|K_S|}\right)$ such that there exists a universally resonance invariant nk-spectrum S for which its k-spectrum $K_S = \left\{\mathbf{k}_{*1}, \ldots, \mathbf{k}_{*|K_S|}\right\}$. Then $\mathcal{K}_{rinv}$ is open and everywhere dense set in $\left(\mathbb{R}^d \setminus \sigma_{bc}\right)^{|K_S|}$.*

Proof. The fact that $\mathcal{K}_{\text{rinv}}$ is open follows from Definition 18 and the continuity in $\mathbf{k}$ of the dispersion relations $\omega_n(\mathbf{k})$. Let G be a small open ball such that its closure $\bar{G} \subset \left(\mathbb{R}^d \setminus \sigma_{bc}\right)^{|K_S|}$. It suffices to prove that $\bar{G} \cap \mathcal{K}_{\text{rinv}}$ contains at least one point $\left(\mathbf{k}_{*1}, \ldots, \mathbf{k}_{*|K_S|}\right)$. For a given finite set $\mathfrak{M}_F$ let us consider all possible

$$\left(m, \zeta, n_0, \vec{\lambda}\right) \in \bigcup_{m \in \mathfrak{M}_F} \times \{-1, 1\} \times \{1, \ldots, J\} \times \Lambda^m$$

which are not universal solutions to (100), and for a given $\left(m, \zeta, n_0, \vec{\lambda}\right)$ let $G_0(m, \zeta, n_0, \vec{\lambda})$ be a set of solutions $\left(\mathbf{k}_1, \ldots, \mathbf{k}_{|K_S|}\right)$ to (100) in $\bar{G}$, and notice that it is a closed set. Let now $G_0(S) \subset \bar{G}$ be the union of the sets $G_0\left(m, \zeta, n_0, \vec{\lambda}\right)$ over all $\left(m, \zeta, n_0, \vec{\lambda}\right) \in P(S) \setminus P_{\text{univ}}(S)$ and let us show that $G_0(S) \neq G$. Indeed, suppose that $G_0(S) = G$ and hence G is a finite union of closed sets. According to Baire's theorem one of the sets $G_0\left(m, \zeta, n_0, \vec{\lambda}\right)$ with $\left(m, \zeta, n_0, \vec{\lambda}\right) \in P(S) \setminus P_{\text{univ}}(S)$ must have a nonempty interior. Then, according to Theorem 22, the dispersion relations $\omega_n(\mathbf{k})$ are ω-degenerate as in Definition 21 contradicting the conditions of the theorem. Hence, there is always a point $\left(\mathbf{k}_{*1}, \ldots, \mathbf{k}_{*|K_S|}\right) \in P(S) \setminus P_{\text{univ}}(S)$ that completes the proof. □

The proof of the next statement is elementary and we skip it.

Proposition 24. *Let $f_1(\mathbf{k})$, $f_2(\mathbf{k})$, $f_3(\mathbf{k})$ be real-valued and continuous functions respectively in neighborhoods of $\mathbf{k}_{*1}$, $\mathbf{k}_{*2}$, $\mathbf{k}_{*1}+\mathbf{k}_{*2}$ in $\mathbb{R}^d$. Assume that the following equation:*

$$f_1(\mathbf{k}_1) + f_2(\mathbf{k}_2) = f_3(\delta_1 \mathbf{k}_1 + \delta_2 \mathbf{k}_2) + C_0,$$

holds in these neighborhoods where C_0, δ_1, δ_2 are constants and $\delta_1 \delta_2 \neq 0$. Then all three functions $f_1(\mathbf{k})$, $f_2(\mathbf{k})$, $f_3(\mathbf{k})$ are linear in neighborhoods of $\mathbf{k}_{*1}$, $\mathbf{k}_{*2}$, $\mathbf{k}_{*1} + \mathbf{k}_{*2}$ respectively.

4. Reduction to a Standard Framework

Many well known nonlinear evolutionary equations and systems can be easily reduced to the framework of (1), (3) involving two small parameters ϱ and β and characterized by the following properties: (i) the linear part $\mathbf{L}$ has a large factor $\frac{1}{\varrho}$ before it; (ii) the nonlinearity $\mathbf{F}(\mathbf{U})$ is independent of ϱ, β or depends on ϱ regularly; (iii) the initial data depend on β so that they do not vanish as $\beta \to 0$; (iv) the solutions are considered on the time interval $0 \leq \tau \leq \tau_*$, where $\tau_* > 0$ does not depend on ϱ, β. Notice that solutions to (1), (3) under the above conditions exhibit nonlinear effects uniformly with respect to small ϱ, β on the time interval $0 \leq \tau \leq \tau_*$.

There are important classes of problems which can be readily reduced to the framework of (1), (3) by a simple rescaling.

Systems with a small factor before the nonlinearity. Consider a problem of the form

$$\partial_t \mathbf{v} = -\mathrm{i}\mathbf{L}\mathbf{v} + \alpha \mathbf{f}(\mathbf{v}), \quad \mathbf{v}|_{t=0} = \mathbf{h}, \quad 0 < \alpha \ll 1, \tag{110}$$

where initial data are bounded uniformly in α. Such problems are reduced to (1) by the time rescaling $\tau = t\alpha$. Note that now $\varrho = \alpha$ and the finite time interval $0 \leq \tau \leq \tau_*$ corresponds to the long time interval $0 \leq t \leq \tau_*/\alpha$.

Systems with small initial data on long time intervals. The equation here is

$$\partial_t \mathbf{v} = -\mathrm{i}\mathbf{L}\mathbf{v} + \mathbf{f}_0(\mathbf{v}), \quad \mathbf{v}|_{t=0} = \alpha_0 \mathbf{h}, \quad 0 < \alpha_0 \ll 1, \text{ where}$$
$$\mathbf{f}_0(\mathbf{v}) = \mathbf{f}_0^{(m)}(\mathbf{v}) + \mathbf{f}_0^{(m+1)}(\mathbf{v}) + \ldots, \tag{111}$$

and $\mathbf{f}^{(m)}(\mathbf{v})$ is a homogeneous polynomial of degree $m \geq 2$. After the rescaling $\mathbf{v} = \alpha_0 \mathbf{V}$, we obtain the following equation with a small nonlinearity:

$$\partial_t \mathbf{V} = -\mathrm{i}\mathbf{L}\mathbf{V} + \alpha_0^{m-1}\left[\mathbf{f}_0^{(m)}(\mathbf{V}) + \alpha_0 \mathbf{f}^{0(m+1)}(\mathbf{V}) + \ldots\right], \quad \mathbf{V}|_{t=0} = \mathbf{h}, \tag{112}$$

which is of the form of (110) with $\alpha = \alpha_0^{m-1}$. Note that nonlinearities $\mathbf{f}$ in (110) which are obtained from problems with small initial data and regular nonlinearities $\mathbf{f}_0(\mathbf{v})$ have a special form. Namely, they are almost homogeneous, $\mathbf{f}(\mathbf{V}) = \mathbf{f}_0^{(m)}(\mathbf{V}) + \alpha[\ldots]$ with leading term $\mathbf{f}_0^{(m)}(\mathbf{V})$. Introducing the slow time variable $\tau = t\alpha_0^{m-1}$ we get from the above an equation of the form (1), namely

$$\partial_\tau \mathbf{V} = -\frac{\mathrm{i}}{\alpha_0^{m-1}}\mathbf{L}\mathbf{V} + \left[\mathbf{f}^{(m)}(\mathbf{V}) + \alpha_0 \mathbf{f}^{(m+1)}(\mathbf{V}) + \ldots\right], \quad \mathbf{V}|_{\tau=0} = \mathbf{h}, \tag{113}$$

where the nonlinearity does not vanish as $\alpha_0 \to 0$. In this case $\varrho = \alpha_0^{m-1}$ and the finite time interval $0 \leq \tau \leq \tau_*$ corresponds to the long time interval $0 \leq t \leq \frac{\tau_*}{\alpha_0^{m-1}}$ with small $\alpha_0 \ll 1$. Note that Corollary 38 for ϱ-dependent nonlinearities can be applied to this case. This allows, in particular, to apply results of this paper to the Sine-Gordon equation where $\mathbf{f}_0(v) = \sin v$. Note that a different rescaling $\tau = t\alpha_0^m$ with $\varrho = \alpha_0^m$ would produce a large term $\varrho^{-1/m}\mathbf{f}^{(m)}(\mathbf{V})$. If the term $\mathbf{f}^{(m)}(\mathbf{V})$ is non-resonant for the initial data $\mathbf{h}$ such a term still produces a small contribution to the solution on interval $t \leq \tau_*/\alpha_0^m$ with small τ_*. The approach of this paper can be applied to this moderately singular case as well, but it would require more technical efforts and for the sake of simplicity we restrict ourselves to the regular case. The interaction of quadratic ($m = 2$) nonlinearity with the cubic term of the 1D model equation of form (111) was studied by Schneider 51].

High-frequency carrier waves. Sometimes high spatial frequency of carrier waves in the initial wavepackets after a rescaling creates a large parameter $\frac{1}{\varrho}$ at the linear part. For example, Nonlinear Schrodinger equation

$$\partial_\tau U = -\mathrm{i}\partial_x^2 U + \mathrm{i}\alpha |U|^2 U, \quad U|_{\tau=0} = h_1(\beta x) e^{\mathrm{i}Mk_{*1}x} + h_2(\beta x) e^{\mathrm{i}Mk_{*2}x} + c.c., \tag{114}$$

where $c.c.$ stands for complex conjugate of the prior term, and $M \gg 1$ is a large parameter, can be recast in the form (1). Indeed, changing variables $y = Mx$ in the above equation we obtain

$$\partial_\tau U = -\mathrm{i}\frac{1}{\varrho}\partial_y^2 U + \mathrm{i}\alpha |U|^2 U, \quad U|_{\tau=0} = h_1\left(\beta_1 y\right) e^{\mathrm{i}k_{*1}y} + h_2\left(\beta_1 y\right) e^{\mathrm{i}k_{*2}y} + c.c.,$$

where $\beta_1 = \frac{\beta}{M} \ll 1$, $\varrho = \frac{1}{M^2} \ll 1$. Note that though the nonlinearity$|U|^2 U$ in (114) is not complex homogeneous, it can be considered as a restriction of a system with a complex homogeneous nonlinearity as (67) is a restriction of (62).

First order hyperbolic equations and systems. Consider now the system (45), (46) for which the symmetry (7) does not hold. The system can be put into the standard framework by formally adding two more equations

$$\partial_\tau w_1 = \frac{c_1}{\varrho}\partial_x w_1 + F_1(w_1, w_2), \quad \partial_\tau w_2 = \frac{c_2}{\varrho}\partial_x w_2 + F_2(w_1, w_2), \tag{115}$$
$$w_1|_{\tau=0} = 0, \quad w_2|_{\tau=0} = 0,$$

which have only trivial solution $w_1 = w_2 = 0$ not affecting the solutions to the original system (45), (46). The extended system has the linear part with two-band dispersion relations $\omega_{1,\zeta}(k) = c_1\zeta|k|$, $\omega_{2,\zeta}(k) = c_2\zeta|k|$, $\zeta = \pm$, satisfying evidently (7).

5. Integrated Evolution Equation

Using the variation of constants formula we recast the modal evolution equation (3) into the following equivalent integral form:

$$\hat{\mathbf{U}}(\mathbf{k}, \tau) = \int_0^\tau \mathrm{e}^{\frac{-\mathrm{i}\left(\tau-\tau'\right)}{\varrho}\mathbf{L}(\mathbf{k})} \hat{F}\left(\hat{\mathbf{U}}\right)(\mathbf{k}, \tau)\, \mathrm{d}\tau' + \mathrm{e}^{\frac{-\mathrm{i}\zeta\tau}{\varrho}\mathbf{L}(\mathbf{k})} \hat{\mathbf{h}}(\mathbf{k}), \quad \tau \geq 0. \tag{116}$$

Then we factor $\hat{\mathbf{U}}(\mathbf{k}, \tau)$ into the slow variable $\hat{\mathbf{u}}(\mathbf{k}, \tau)$ and the fast oscillatory term as in (14), namely

$$\hat{\mathbf{U}}(\mathbf{k}, \tau) = \mathrm{e}^{-\frac{\mathrm{i}\tau}{\varrho}\mathbf{L}(\mathbf{k})} \hat{\mathbf{u}}(\mathbf{k}, \tau), \quad \hat{\mathbf{U}}_{n,\zeta}(\mathbf{k}, \tau) = \hat{\mathbf{u}}_{n,\zeta}(\mathbf{k}, \tau)\, \mathrm{e}^{-\frac{\mathrm{i}\tau}{\varrho}\zeta\omega_n(\mathbf{k})}, \tag{117}$$

where $\hat{\mathbf{u}}_{n,\zeta}(\mathbf{k}, \tau)$ are the modal coefficients of $\hat{\mathbf{u}}(\mathbf{k}, \tau)$ as in (81). Notice that $\hat{\mathbf{u}}_{n,\zeta}(\mathbf{k}, \tau)$ in (117) may depend on ϱ and (117) is just a change of variables and not an assumption. Consequently we obtain the following *integrated evolution equation* for $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{k}, \tau)$, $\tau \geq 0$,

$$\hat{\mathbf{u}}(\mathbf{k}, \tau) = \mathcal{F}\left(\hat{\mathbf{u}}\right)(\mathbf{k}, \tau) + \hat{\mathbf{h}}(\mathbf{k}), \quad \mathcal{F}\left(\hat{\mathbf{u}}\right) = \sum_{m \in \mathfrak{M}_F} \mathcal{F}^{(m)}\left(\hat{\mathbf{u}}^m(\mathbf{k}, \tau)\right), \tag{118}$$

$$\mathcal{F}^{(m)}\left(\hat{\mathbf{u}}^m\right)(\mathbf{k}, \tau) = \int_0^\tau \mathrm{e}^{\frac{\mathrm{i}\tau'}{\varrho}\mathbf{L}(\mathbf{k})} \hat{F}_m\left(\left(\mathrm{e}^{\frac{-\mathrm{i}\tau'}{\varrho}\mathbf{L}(\cdot)}\hat{\mathbf{u}}\right)^m\right)\left(\mathbf{k}, \tau'\right) \mathrm{d}\tau', \tag{119}$$

where $\hat{F}_m$ are defined by (84) and (86) in terms of the susceptibilities $\chi^{(m)}$, and $\mathcal{F}^{(m)}$ are bounded as in the following lemma.

Lemma 25 (Boundness of multilinear operators). *$\mathcal{F}^{(m)}$ defined by (86), (119) is a bounded operator from $E = C\left([0,\tau_*], L^1\right)$ into $C^1\left([0,\tau_*], L^1\right)$ satisfying*

$$\left\|\mathcal{F}^{(m)}\left(\hat{\mathbf{u}}_1 \ldots \hat{\mathbf{u}}_m\right)\right\|_E \le \tau_* \left\|\chi^{(m)}\right\| \prod_{j=1}^m \left\|\hat{\mathbf{u}}_j\right\|_E, \tag{120}$$

$$\left\|\partial_\tau \mathcal{F}^{(m)}\left(\hat{\mathbf{u}}_1 \ldots \hat{\mathbf{u}}_m\right)\right\|_E \le \left\|\chi^{(m)}\right\| \prod_j \left\|\hat{\mathbf{u}}_j\right\|_E. \tag{121}$$

Proof. Notice that since $\mathbf{L}(\mathbf{k})$ is Hermitian, $\left\|\exp\left\{-\mathrm{i}\mathbf{L}(\mathbf{k})\frac{\tau_1}{\varrho}\right\}\right\| = 1$. Using the Young inequality,

$$\left\|\hat{\mathbf{u}} * \hat{\mathbf{v}}\right\|_{L^1} \le \left\|\hat{\mathbf{u}}\right\|_{L^1} \left\|\hat{\mathbf{v}}\right\|_{L^1}, \tag{122}$$

together with (86), (119) we obtain

$$\left\|\mathcal{F}^{(m)}\left(\hat{\mathbf{u}}_1 \ldots \hat{\mathbf{u}}_m\right)(\cdot,\tau)\right\|_{L^1} \le \sup_{\mathbf{k},\vec{k}} \left|\chi^{(m)}\left(\mathbf{k},\vec{k}\right)\right|$$

$$\int_{\mathbb{R}^d}\int_0^\tau\int_{\mathbb{D}_m} \left|\hat{\mathbf{u}}_1\left(\mathbf{k}'\right)\right| \ldots \left|\hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right)\right)\right| \mathrm{d}\mathbf{k}' \ldots \mathrm{d}\mathbf{k}^{(m-1)}\mathrm{d}\tau_1 \mathrm{d}\mathbf{k} \le$$

$$\left\|\chi^{(m)}\right\| \int_0^\tau \left\|\hat{\mathbf{u}}_1(\tau_1)\right\|_{L^1} \ldots \left\|\hat{\mathbf{u}}_m(\tau_1)\right\|_{L^1} \mathrm{d}\tau_1 \le \tau_* \left\|\chi^{(m)}\right\| \left\|\hat{\mathbf{u}}_1\right\|_E \ldots \left\|\hat{\mathbf{u}}_m\right\|_E,$$

proving (120). Similarly we prove (121) by

$$\left\|\partial_\tau \mathcal{F}^{(m)}\left(\hat{\mathbf{u}}_1 \ldots \hat{\mathbf{u}}_m\right)(\cdot,\tau)\right\|_{L^1} \le \left\|\chi^{(m)}\right\|$$

$$\int_{\mathbb{R}^d}\int_{\mathbb{D}_m} \left|\hat{\mathbf{u}}_1\left(\mathbf{k}'\right)\right| \ldots \left|\hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right)\right)\right| \mathrm{d}\mathbf{k}' \ldots \mathrm{d}\mathbf{k}^{(m-1)}\mathrm{d}\mathbf{k} \le \left\|\chi^{(m)}\right\| \left\|\hat{\mathbf{u}}_1\right\|_E \cdots \left\|\hat{\mathbf{u}}_m\right\|_E.$$

□

Equation (118) can be recast as the following abstract equation in a Banach space:

$$\hat{\mathbf{u}} = \mathcal{F}\left(\hat{\mathbf{u}}\right) + \hat{\mathbf{h}}, \quad \hat{\mathbf{u}}, \hat{\mathbf{h}} \in E, \tag{123}$$

and it readily follows from Lemma 25 that $\mathcal{F}\left(\hat{\mathbf{u}}\right)$ has the following properties.

Lemma 26. *The operator $\mathcal{F}\left(\hat{\mathbf{u}}\right)$ defined by (118)–(119) satisfies the Lipschitz condition*

$$\left\|\mathcal{F}\left(\hat{\mathbf{u}}_1\right) - \mathcal{F}\left(\hat{\mathbf{u}}_2\right)\right\|_E \le \tau_* C_F \left\|\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2\right\|_E, \tag{124}$$

where $C_F \le C_\chi m_F^2 (4R)^{m_F - 1}$ if $\left\|\hat{\mathbf{u}}_1\right\|_E, \left\|\hat{\mathbf{u}}_2\right\|_E \le 2R$, with C_χ as in (88).

We also will use the following form of the contraction principle.

Lemma 27 (Contraction principle). *Consider equation*

$$\mathbf{x} = \mathcal{F}(\mathbf{x}) + \mathbf{h}, \quad \mathbf{x}, \mathbf{h} \in B, \tag{125}$$

where B is a Banach space, $\mathcal{F}$ is an operator in B. Suppose that for some constants $R_0 > 0$ and $0 < q < 1$ we have

$$\|\mathbf{h}\| \leq R_0, \quad \|\mathcal{F}(\mathbf{x})\| \leq R_0 \quad \text{if } \|\mathbf{x}\| \leq 2R_0, \tag{126}$$

$$\|\mathcal{F}(\mathbf{x}_1) - \mathcal{F}(\mathbf{x}_2)\| \leq q \|\mathbf{x}_1 - \mathbf{x}_2\| \quad \text{if } \|\mathbf{x}_1\|, \|\mathbf{x}_2\| \leq 2R_0. \tag{127}$$

Then there exists a unique solution $\mathbf{x}$ to Eq. (125) such that $\|\mathbf{x}\| \leq 2R_0$. Let $\|\mathbf{h}_1\|, \|\mathbf{h}_2\| \leq R_0$, then the two corresponding solutions $\mathbf{x}_1, \mathbf{x}_2$ satisfy

$$\|\mathbf{x}_1\|, \|\mathbf{x}_2\| \leq 2R_0, \quad \|\mathbf{x}_1 - \mathbf{x}_2\| \leq (1-q)^{-1} \|\mathbf{h}_1 - \mathbf{h}_2\|. \tag{128}$$

Let $\mathbf{x}_1, \mathbf{x}_2$ be the two solutions of correspondingly two equations of the form (125) with $\mathcal{F}_1$, $\mathbf{h}_1$ and $\mathcal{F}_2$, $\mathbf{h}_2$. Assume that that $\mathcal{F}_1(\mathbf{u})$ satisfies (126), (127) with a Lipschitz constant $q < 1$ and that $\|\mathcal{F}_1(\mathbf{x}) - \mathcal{F}_2(\mathbf{x})\| \leq \delta$ for $\|\mathbf{x}\| \leq 2R_0$. Then

$$\|\mathbf{x}_1 - \mathbf{x}_2\| \leq (1-q)^{-1} (\delta + \|\mathbf{h}_1 - \mathbf{h}_2\|). \tag{129}$$

Lemma 26 and the contraction principle as in Lemma 27 imply the following existence and uniqueness theorem.

Theorem 28. *Let $\|\mathbf{h}\|_E \leq R$, and let $\tau_* < 1/C_F$ where C_F, is a constant from Lemma 26. Then Eq. (118) has a solution $\hat{\mathbf{u}} \in E = C\left([0, \tau_*], L^1\right)$ which satisfies $\left\|\hat{\mathbf{u}}\right\|_E \leq 2R$, and such a solution is unique.*

The following existence and uniqueness theorem follows from Theorem 28.

Theorem 29. *Let (3) satisfy (88) and $\hat{\mathbf{h}} \in L^1\left(\mathbb{R}^d\right)$, $\left\|\hat{\mathbf{h}}\right\|_{L^1} \leq R$. Then there exists a unique solution to the modal evolution equation (3) in the functional space $C^1([0, \tau_*], L^1)$. The number τ_* depends on R and C_χ.*

Using the inequality (21) and applying the inverse Fourier transform we readily obtain the existence of an $F-$solution of (1) in $C^1\left([0, \tau_*], L^\infty\left(\mathbb{R}^d\right)\right)$ from the existence of the solution of Eq. (3) in $C^1\left([0, \tau_*], L^1\right)$. The existence of F-solutions in spaces of spatially smooth functions can be derived by replacing Lemma 25 with an estimate similar to the one in Lemma 50.

Let us recast now the system (118)–(119) into modal components using the projections $\Pi_{n,\zeta}(\mathbf{k})$ as in (11). The first step to introduce elementary modal susceptibilities $\chi^{(m)}_{n,\zeta,\vec{\xi}}$ having one-dimensional range in $\mathbb{C}^{2J}$ and vanishing if one of its arguments $\hat{\mathbf{u}}_j$ belongs to a $(2J-1)$-dimensional linear subspace in $\mathbb{C}^{2J}$ (j^{th} null-space of $\chi^{(m)}_{n,\zeta,\vec{\xi}}$). For example, in the linear case $m = 1$ when $\chi^{(1)}$ acts in $\mathbb{C}^{2J}$ and is presented in the standard orthonormal basis $\left\{\mathbf{e}_{n,\zeta}\right\}$ in $\mathbb{C}^{2J}$ by a $2J \times 2J$ matrix with elements $a^{(1)}_{\xi,\xi'} = a^{(1)}_{n,\zeta,n',\zeta'}$, where index $\xi = n, \zeta$ takes $2J$ values, the action of elementary susceptibility $\chi^{(1)}_{n,\zeta,n',\zeta'}$ on a vector $\mathbf{v} \in \mathbb{C}^{2J}$ is given by the formula $\chi^{(1)}_{n,\zeta,n',\zeta'}\mathbf{v} = a^{(1)}_{n,\zeta,n',\zeta'}\left(\mathbf{v} \cdot \mathbf{e}_{n',\zeta'}\right)\mathbf{e}_{n,\zeta}$, where $\left\{\mathbf{e}_{n,\zeta}\right\}$ is the standard orthonormal basis in $\mathbb{C}^{2J}$. Obviously $\chi^{(1)}_{n,\zeta,n',\zeta'}\mathbf{v} = \Pi_{n,\zeta}\chi^{(1)}\Pi_{n',\zeta'}\mathbf{v}$ and $\chi^{(1)}\mathbf{v} = \sum_{n,\zeta,n',\zeta'} \chi^{(1)}_{n,\zeta,n',\zeta'}\mathbf{v}$. The general definition follows.

Definition 30 (Elementary susceptibilities). *Let*

$$\vec{\xi} = \left(\vec{n}, \vec{\zeta}\right) \in \{1, \dots, J\}^m \times \{-1, 1\}^m = \Xi^m, (n, \zeta) \in \Xi \tag{130}$$

and $\chi^{(m)}\left(\mathbf{k}, \vec{k}\right)\left[\hat{\mathbf{u}}_1\left(\mathbf{k}'\right), \dots, \hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\right)\right]$ *be the* m*-linear symmetric tensor (susceptibility) as in* (86)*. We introduce* **elementary susceptibilities** $\chi^{(m)}_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right) : \left(\mathbb{C}^{2J}\right)^m \to \mathbb{C}^{2J}$*) as m-linear tensors defined for almost all* $\mathbf{k}, \vec{k}$ *by the following formula:*

$$\begin{aligned}
&\chi^{(m)}_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right)\left[\hat{\mathbf{u}}_1\left(\mathbf{k}'\right), \dots, \hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\right)\right] = \chi^{(m)}_{n,\zeta,\vec{n},\vec{\zeta}}\left(\mathbf{k}, \vec{k}\right)\left[\hat{\mathbf{u}}_1\left(\mathbf{k}'\right), \dots, \hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\right)\right] \\
&= \Pi_{n,\zeta}\left(\mathbf{k}\right)\chi^{(m)}\left(\mathbf{k}, \vec{k}\right)\left[\left(\Pi_{n_1,\zeta'}\left(\mathbf{k}'\right)\hat{\mathbf{u}}_1\left(\mathbf{k}'\right), \dots, \Pi_{n_m,\zeta^{(m)}}\left(\mathbf{k}^{(m)}\left(\mathbf{k}, \vec{k}\right)\right)\hat{\mathbf{u}}_m\right.\right. \\
&\quad \left.\left.\times\left(\mathbf{k}^{(m)}\left(\mathbf{k}, \vec{k}\right)\right)\right)\right].
\end{aligned} \tag{131}$$

Then using (82) and the *elementary susceptibilities* (131) we get

$$\begin{aligned}
&\chi^{(m)}\left(\mathbf{k}, \vec{k}\right)\left[\hat{\mathbf{u}}_1\left(\mathbf{k}'\right), \dots, \hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\right)\right] \\
&= \sum_{n,\zeta}\sum_{\vec{\xi}}\chi^{(m)}_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right)\left[\hat{\mathbf{u}}_1\left(\mathbf{k}'\right), \dots, \hat{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\right)\right].
\end{aligned} \tag{132}$$

Consequently the modal components $\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}$ of the operators $\mathcal{F}^{(m)}$ in (119) are m-linear *oscillatory integral operators* defined in terms of the elementary susceptibilities (132) as follows.

Definition 31 (Interaction phase). *Using notations from* (86) *we introduce for* $\vec{\xi} = \left(\vec{n}, \vec{\zeta}\right) \in \Xi^m$ *the operator*

$$\begin{aligned}
&\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left(\tilde{\mathbf{u}}_1 \dots \tilde{\mathbf{u}}_m\right)\left(\mathbf{k}, \tau\right) = \int_0^\tau \int_{\mathbb{D}_m} \exp\left\{\mathrm{i}\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right)\frac{\tau_1}{\varrho}\right\} \\
&\chi^{(m)}_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right)\left[\tilde{\mathbf{u}}_1\left(\mathbf{k}', \tau_1\right), \dots, \tilde{\mathbf{u}}_m\left(\mathbf{k}^{(m)}\left(\mathbf{k}, \vec{k}\right), \tau_1\right)\right]\tilde{\mathrm{d}}^{(m-1)d}\vec{k}\mathrm{d}\tau_1,
\end{aligned} \tag{133}$$

with the **interaction phase function** ϕ *defined by*

$$\begin{aligned}
&\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right) = \phi_{n,\zeta,\vec{n},\vec{\zeta}}\left(\mathbf{k}, \vec{k}\right) \\
&= \zeta\omega_n\left(\zeta\mathbf{k}\right) - \zeta'\omega_{n_1}\left(\zeta'\mathbf{k}'\right) - \dots - \zeta^{(m)}\omega_{n_m}\left(\zeta^{(m)}\mathbf{k}^{(m)}\right), \quad \mathbf{k}^{(m)} = \mathbf{k}^{(m)}\left(\mathbf{k}, \vec{k}\right).
\end{aligned} \tag{134}$$

Using $\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}$ in (133) we recast $\mathcal{F}^{(m)}\left(\mathbf{u}^m\right)$ in the system (118)–(119) as

$$\mathcal{F}^{(m)}\left[\hat{\mathbf{u}}_1 \dots, \hat{\mathbf{u}}_m\right]\left(\mathbf{k}, \tau\right) = \sum_{n,\zeta,\vec{\xi}}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left[\hat{\mathbf{u}}_1 \dots \hat{\mathbf{u}}_m\right]\left(\mathbf{k}, \tau\right), \tag{135}$$

yielding the following system for the modal components $\hat{\mathbf{u}}_{n,\zeta}\left(\mathbf{k}, \tau\right)$ as in (11),

$$\hat{\mathbf{u}}_{n,\zeta}\left(\mathbf{k}, \tau\right) = \sum_{m\in\mathfrak{M}_F}\sum_{\vec{\xi}\in\Xi^m}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left(\hat{\mathbf{u}}^m\right)\left(\mathbf{k}, \tau\right) + \hat{\mathbf{h}}_{n,\zeta}\left(\mathbf{k}\right), \quad (n, \zeta) \in \Xi. \tag{136}$$

6. Wavepacket Interaction System

The wavepacket preservation property of the nonlinear evolutionary system in any of its forms (1), (3), (118), (123), (136) is not easy to see directly. It turns out though that dynamics of wavepackets is well described by a system in a larger space E^{2N} based on the original equation (118) in the space E. We call it a *wavepacket interaction system*, which is useful in three ways: (i) the wavepacket preservation is quite easy to see and verify; (ii) it can be used to prove the wavepacket preservation for the original nonlinear problem; (iii) it can be used to study more subtle properties of the original problem, such as NLS approximation. We start with the system (118) where $\hat{\mathbf{h}}(\mathbf{k})$ is a multiwavepacket with a given nk-spectrum $S = \{(\mathbf{k}_{*l}, n_l),\ l = 1, \ldots, N\}$ as in (31) and k-spectrum $K_S = \{\mathbf{k}_{*i},\ i = 1, \ldots, |K_S|\}$ as in (32).

When constructing the wavepacket interaction system it is convenient to have relevant functions to be explicitly localized about the k-spectrum K_S of the initial data. We implement that by making up the following cutoff functions based on (25), (26),

$$\begin{aligned}\Psi_{i,\vartheta}(\mathbf{k}) &= \Psi(\mathbf{k}, \vartheta\mathbf{k}_{*i}) \\ &= \Psi\left(\beta^{-(1-\epsilon)}(\mathbf{k} - \vartheta\mathbf{k}_{*i})\right), \quad \mathbf{k}_{*i} \in K_S, \quad i = 1, \ldots, |K_S|, \quad \vartheta = \pm \end{aligned} \tag{137}$$

with ϵ as in Definition 1 and $\beta > 0$ small enough to satisfy

$$\beta^{1/2} \leq \pi_0, \text{ where } \pi_0 = \pi_0(S) < \frac{1}{2}\min_{\mathbf{k}_{*i} \in K_S} \operatorname{dist}\{\mathbf{k}_{*i}, \sigma_{bc}\}. \tag{138}$$

In what follows we use notations from (92) and

$$\vec{l} = (l_1, \ldots, l_m) \in \{1, \ldots, N\}^m, \quad \vec{\vartheta} = \left(\vartheta', \ldots, \vartheta^{(m)}\right) \in \{-1, 1\}^m, \quad \vec{\lambda} = \left(\vec{l}, \vec{\vartheta}\right) \in \Lambda^m, \tag{139}$$

$$\vec{n} = (n_1, \ldots, n_m) \in \{1, \ldots, J\}^m, \quad \vec{\zeta} \in \{-1, 1\}^m, \tag{140}$$

$$\vec{\xi} = \left(\vec{n}, \vec{\zeta}\right) \in \Xi^m \quad, \vec{k} = \left(\mathbf{k}', \ldots, \mathbf{k}^{(m)}\right) \in \mathbb{R}^m, \text{ where } \Xi^m \text{ as in (130)}.$$

Based on the above we introduce now the *wavepacket interaction system*,

$$\begin{aligned}\hat{\mathbf{w}}_{l,\vartheta}(\cdot) &= \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}(\cdot)\mathcal{F}\left(\sum_{(l',\vartheta')\in\Lambda}\hat{\mathbf{w}}_{l',\vartheta'}\right) \\ &\quad + \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}(\cdot)\hat{\mathbf{h}}, (l, \vartheta) \in \Lambda, \\ \vec{\mathbf{w}} &= \left(\hat{\mathbf{w}}_{1,+}, \hat{\mathbf{w}}_{1,-}, \ldots, \hat{\mathbf{w}}_{N,+}, \hat{\mathbf{w}}_{N,-}\right) \in E^{2N}, \quad \hat{\mathbf{w}}_{l,\vartheta} \in E, (l, \vartheta) \in \Lambda,\end{aligned} \tag{141}$$

with $\Psi(\cdot, \vartheta\mathbf{k}_{*i})$, $\Pi_{n,\vartheta}$ as in (137), (11), $\mathcal{F}$ defined by (118), and the norm in E^{2N} defined based on (17) by the formula

$$\|\vec{\mathbf{w}}\|_{E^{2N}} = \sum_{l,\vartheta}\left\|\hat{\mathbf{w}}_{l,\vartheta}\right\|_E, \quad E = C\left([0, \tau_*], L^1\right).$$

The index (l, ϑ) which takes $2N$ values labels equations and variables, the right-hand side of (141) is well-defined for all $\vec{\mathbf{w}} \in E^{2N}$ and the equality (141) is understood as equality of elements of E^{2N}. We also use the following concise form of the wave interaction system (141):

$$\vec{\mathbf{w}} = \mathcal{F}_{\Psi}(\vec{\mathbf{w}}) + \vec{\mathbf{h}}_{\Psi}, \text{ where} \tag{142}$$
$$\vec{\mathbf{h}}_{\Psi} = \left(\Psi_{i_1,+}\Pi_{n_1,+}\hat{\mathbf{h}}, \Psi_{i_1,-}\Pi_{n_1,-}\hat{\mathbf{h}}, \ldots, \Psi_{i_N,+}\Pi_{n_N,+}\hat{\mathbf{h}}, \Psi_{i_N,-}\Pi_{n_N,-}\hat{\mathbf{h}}\right) \in E^{2N}.$$

The following lemma is analogous to Lemmas 25, 26.

Lemma 32. *Polynomial operator $\mathcal{F}_{\Psi}(\vec{\mathbf{w}})$ is bounded in E^{2N}, $\mathcal{F}_{\Psi}(\mathbf{0}) = \mathbf{0}$, and it satisfies the Lipschitz condition*

$$\|\mathcal{F}_{\Psi}(\vec{\mathbf{w}}_1) - \mathcal{F}_{\Psi}(\vec{\mathbf{w}}_2)\|_{E^{2N}} \le C\tau_* \|\vec{\mathbf{w}}_1 - \vec{\mathbf{w}}_2\|_{E^{2N}}, \tag{143}$$

where C depends only on C_χ as in (88), on the degree of $\mathcal{F}$ and on $\|\vec{\mathbf{w}}_1\|_{E^{2N}} + \|\vec{\mathbf{w}}_2\|_{E^{2N}}$, and it does not depend on β and ϱ.

Proof. We consider every operator $\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}(\vec{\mathbf{w}})$ defined by (133) and prove its boundedness and the Lipschitz property as in Lemma 25 using the inequality $\left|\exp\left\{\mathrm{i}\phi_{n,\zeta,\vec{\xi}}\frac{\tau_1}{\varrho}\right\}\right| \le 1$ and estimates (25), (88). Note that the integration in τ_1 yields the factor τ_* and consequent summation with respect to $n, \zeta, \vec{\xi}$ yields (143). □

Lemma 32 and the contraction principle as in Lemma 27 yield the following statement.

Theorem 33. *Let $\left\|\vec{\mathbf{h}}_{\Psi}\right\|_{E^{2N}} \le R$. Then there exists $R_1 > 0$ and $\tau_* > 0$ such that Eq. (141) has a solution $\vec{\mathbf{w}} \in E^{2N}$ which satisfies $\|\vec{\mathbf{w}}\|_{E^{2N}} \le R_1$ and such a solution is unique.*

Lemma 34. *Every function $\hat{\mathbf{w}}_{l,\zeta}(\mathbf{k}, \tau)$ corresponding to the solution of (142) from E^{2N} is a wavepacket with nk-pair $(\mathbf{k}_{*l}, n_l)$ with the degree of regularity which can be any $s > 0$.*

Proof. Note that according to (137) and (142) the function

$$\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k}, \tau) = \Psi\left(\mathbf{k}, \vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\mathcal{F}(\mathbf{k}, \tau), \quad \|\mathcal{F}(\tau)\|_{L^1} \le C, \quad 0 \le \tau \le \tau_*$$

involves the factor $\Psi_{l,\vartheta}(\mathbf{k}) = \Psi\left(\beta^{-(1-\epsilon)}(\mathbf{k} - \vartheta\mathbf{k}_{*l})\right)$ where ϵ is as in Definition 1. Hence,

$$\Pi_{n,\vartheta'}\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k}, \tau) = 0 \quad \text{if } n \ne n_l \text{ or } \vartheta' \ne \vartheta, \tag{144}$$
$$\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k}, \tau) = \Psi\left(\mathbf{k}, \vartheta\mathbf{k}_{*i_l}\right)\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k}, \tau), \hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k}, \tau) = 0 \quad \text{if } |\mathbf{k} - \vartheta\mathbf{k}_{*l}| \ge \beta^{1-\epsilon}, \tag{145}$$

and, consequently, Definition 1 for $\hat{\mathbf{w}}_{l,\vartheta}$ is satisfied with $\hat{D}_h = 0$ for any $s > 0$ and $C' = 0$ in (30). □

Now we would like to show that if $\hat{\mathbf{h}}$ is a multiwavepacket, then the function

$$\hat{\mathbf{w}}(\mathbf{k},\tau)=\sum_{(l,\vartheta)\in\Lambda}\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k},\tau)=\sum_{\lambda\in\Lambda}\hat{\mathbf{w}}_{\lambda}(\mathbf{k},\tau) \tag{146}$$

is an approximate solution of Eq. (123) (see notation (92)). To do that we introduce

$$\Psi_{\infty}(\mathbf{k})=1-\sum_{\vartheta=\pm}\sum_{i=1}^{|K_S|}\Psi(\mathbf{k},\vartheta\mathbf{k}_{*i})=1-\sum_{\vartheta=\pm}\sum_{\mathbf{k}_{*i}\in K_S}\Psi\left(\frac{\mathbf{k}-\vartheta\mathbf{k}_{*i}}{\beta^{1-\epsilon}}\right). \tag{147}$$

Expanding the m-linear operator $\mathcal{F}^{(m)}\left(\left(\sum_{l,\vartheta}\hat{\mathbf{w}}_{l,\vartheta}\right)^m\right)$ and using notations (92), (93) we get

$$\mathcal{F}^{(m)}\left(\left(\sum_{l,\vartheta}\hat{\mathbf{w}}_{l,\vartheta}\right)^m\right)=\sum_{\vec{\lambda}\in\Lambda^m}\mathcal{F}^{(m)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right),\quad\text{where} \tag{148}$$

$$\vec{\mathbf{w}}_{\vec{\lambda}}=\hat{\mathbf{w}}_{\lambda_1}\ldots\hat{\mathbf{w}}_{\lambda_m},\quad\vec{\lambda}=(\lambda_1,\ldots,\lambda_m)\in\Lambda^m. \tag{149}$$

The next statement shows that (146) defines an approximate solution to the integrated evolution equation (118).

Theorem 35. *Let $\hat{\mathbf{h}}$ be a multi-wavepacket with resonance invariant nk-spectrum S with regularity degree s, $\vec{\mathbf{w}}$ be a solution of (142) and $\hat{\mathbf{w}}(\mathbf{k},\tau)$ be defined by (146). Let*

$$\hat{\mathbf{D}}\left(\hat{\mathbf{w}}\right)=\hat{\mathbf{w}}-\mathcal{F}\left(\hat{\mathbf{w}}\right)-\hat{\mathbf{h}}. \tag{150}$$

Then there exists $\beta_0>0$ such that we have the estimate

$$\left\|\hat{\mathbf{D}}\left(\hat{\mathbf{w}}\right)\right\|_E\le C\varrho+C\beta^s,\ \textit{if}\ 0<\varrho\le 1,\quad\beta\le\beta_0. \tag{151}$$

Proof. Let

$$\mathcal{F}^{-}\left(\hat{\mathbf{w}}\right)=\left(1-\sum_{l,\vartheta}\Psi_{i_l,\vartheta}\Pi_{n_l,\vartheta}\right)\mathcal{F}\left(\hat{\mathbf{w}}\right),\quad\hat{\mathbf{h}}^{-}=\hat{\mathbf{h}}-\sum_{l,\vartheta}\Psi_{i_l,\vartheta}\Pi_{n_l,\vartheta}\hat{\mathbf{h}}. \tag{152}$$

Summation of (141) with respect to l,ϑ yields

$$\hat{\mathbf{w}}=\sum_{l,\vartheta}\Psi_{i_l,\vartheta}\Pi_{n_l,\vartheta}\mathcal{F}\left(\hat{\mathbf{w}}\right)+\sum_{l,\vartheta}\Psi_{i_l,\vartheta}\Pi_{n_l,\vartheta}\hat{\mathbf{h}}.$$

Hence, from (141) and (150) we obtain

$$\hat{\mathbf{D}}\left(\hat{\mathbf{w}}\right)=\hat{\mathbf{h}}^{-}-\mathcal{F}^{-}\left(\hat{\mathbf{w}}\right). \tag{153}$$

Using (28) and (30) we consequently obtain

$$\left\|\Pi_{n_l,\vartheta}\hat{\mathbf{h}}_i\right\|_{L^1}\le C\beta^s\quad\text{if } n_l\ne n_i;\quad\left\|\Psi_{i_l,\vartheta}\hat{\mathbf{h}}_i\right\|_{L^1}\le C\beta^s\quad\text{if } \mathbf{k}_{*i_l}\ne\mathbf{k}_{*i},$$
$$\left\|\hat{\mathbf{h}}^{-}\right\|_E\le C_1\beta^s. \tag{154}$$

Now, to show (151) it is sufficient to prove that

$$\left\| \mathcal{F}^{-}\left(\hat{\mathbf{w}}\right)\right\|_{E} \leq C_{2}\varrho. \tag{155}$$

Obviously,

$$\mathcal{F}^{-}\left(\hat{\mathbf{w}}\right) = \left(1 - \sum_{l,\vartheta} \Psi_{i_l,\vartheta} \Pi_{n_l,\vartheta}\right) \sum_{m} \mathcal{F}^{(m)}\left(\hat{\mathbf{w}}^{m}\right). \tag{156}$$

Note that

$$\sum_{l,\vartheta} \Psi_{i_l,\vartheta} \Pi_{n_l,\vartheta} = \sum_{\vartheta=\pm} \sum_{(n,k_*)\in S} \Psi\left(\cdot, \vartheta \mathbf{k}_*\right) \Pi_{n,\vartheta}. \tag{157}$$

Using (82) and (147) we consequently obtain

$$\sum_{\vartheta=\pm} \sum_{(n,k_*)\in \Sigma} \Psi\left(\cdot, \vartheta \mathbf{k}_*\right) \Pi_{n,\vartheta} + \Psi_{\infty} = 1, \tag{158}$$

$$\left(1 - \sum_{l,\vartheta} \Psi_{i_l,\vartheta} \Pi_{n_l,\vartheta}\right) = \Psi_{\infty} + \sum_{\vartheta=\pm} \sum_{(n,k_*)\in \Sigma\setminus S} \Psi\left(\cdot, \vartheta \mathbf{k}_*\right) \Pi_{n,\vartheta}, \tag{159}$$

with Σ defined in (90). Let us expand now $\mathcal{F}^{(m)}\left(\hat{\mathbf{w}}^{m}\right)$ using (148). According to (156) and (159) to prove (155) it is sufficient to prove that for every string $\vec{\lambda} \in \Lambda^{m}$ the following inequalities hold:

$$\left\| \Psi_{\infty} \Pi_{n,\vartheta} \mathcal{F}^{(m)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)\right\| \leq C_{3}\varrho \quad \text{for } (n,\vartheta) \in \Lambda, \text{ and} \tag{160}$$

$$\left\| \Psi\left(\cdot, \vartheta \mathbf{k}_*\right) \Pi_{n,\vartheta} \mathcal{F}^{(m)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)\right\| \leq C_{3}\varrho, \quad \text{if } (n, \mathbf{k}_*) \in \Sigma \setminus S. \tag{161}$$

We will use (144) and (145) to obtain the above estimates. According to (135)

$$\mathcal{F}^{(m)}\left[\vec{\mathbf{w}}_{\vec{\lambda}}\right](\mathbf{k}, \tau) = \sum_{n,\zeta} \sum_{\vec{\xi}} \mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left[\hat{\mathbf{w}}_{\lambda_1} \ldots \hat{\mathbf{w}}_{\lambda_m}\right](\mathbf{k}, \tau). \tag{162}$$

Note that according to (144) if $\lambda_i = \left(l, \vartheta'\right)$

$$\hat{\mathbf{w}}_{\lambda_i} = \Pi_{n,\vartheta} \hat{\mathbf{w}}_{\lambda_i}, \quad \text{if } n = n_l \text{ and } \vartheta' = \vartheta. \tag{163}$$

Let us introduce the notation

$$\vec{n}\left(\vec{l}\right) = \left(n_{l_1}, \ldots, n_{l_m}\right), \quad \vec{\xi}\left(\vec{\lambda}\right) = \left(\vec{n}\left(\vec{l}\right), \vec{\vartheta}\right), \quad \text{for } \vec{\lambda} = \left(\vec{l}, \vec{\vartheta}\right) \in \Lambda^{m}. \tag{164}$$

Since

$$\Pi_{n',\vartheta} \Pi_{n,\vartheta'} = 0, \quad \text{if } n \neq n' \text{ or } \vartheta' \neq \vartheta, \tag{165}$$

then (163) implies

$$\begin{gathered} \mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left[\hat{\mathbf{w}}_{\lambda_1} \ldots \hat{\mathbf{w}}_{\lambda_m}\right] = 0 \quad \text{if } \vec{\xi} = \left(\vec{n}, \vec{\zeta}\right) \neq \vec{\xi}\left(\vec{\lambda}\right), \quad \text{and, hence,} \\ \mathcal{F}^{(m)}\left[\vec{\mathbf{w}}_{\vec{\lambda}}\right](\mathbf{k}, \tau) = \sum_{n,\zeta} \mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}\left(\vec{\lambda}\right)}\left[\hat{\mathbf{w}}_{\lambda_1} \ldots \hat{\mathbf{w}}_{\lambda_m}\right](\mathbf{k}, \tau), \end{gathered} \tag{166}$$

where we use notation (93), (164). Note also that

$$\Pi_{n',\vartheta}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}} = 0 \quad \text{if } n' \neq n \text{ or } \vartheta \neq \zeta, \tag{167}$$

and, hence, we have nonzero $\Pi_{n',\vartheta}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)$ only if

$$\vec{\xi} = \vec{\xi}\left(\vec{\lambda}\right), \quad n' = n, \vartheta = \zeta. \tag{168}$$

By (133)

$$\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)(\mathbf{k},\tau) = \int_0^\tau \int_{\mathbb{D}_m} \exp\left\{\mathrm{i}\phi_{n,\zeta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\mathbf{k},\vec{k}\right)\frac{\tau_1}{\varrho}\right\} \tag{169}$$

$$\chi^{(m)}_{n,\zeta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\mathbf{k},\vec{k}\right)\left[\hat{\mathbf{w}}_{\lambda_1}\left(\mathbf{k}',\tau_1\right),\ldots,\hat{\mathbf{w}}_{\lambda_m}\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right),\tau_1\right)\right]\tilde{\mathrm{d}}^{(m-1)d}\vec{k}\mathrm{d}\tau_1,$$

Now we use (145) and notice that according to the convolution identity in (86),

$$\left|\hat{\mathbf{w}}_{\lambda_1}\left(\mathbf{k}',\tau_1\right)\right|\cdot\ldots\cdot\left|\hat{\mathbf{w}}_{\lambda_m}\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right),\tau_1\right)\right| = 0 \text{ if } \left|\mathbf{k} - \sum_i \vartheta_i \mathbf{k}_{*l_i}\right| \geq m\beta^{1-\epsilon}. \tag{170}$$

Hence the integral (169) is nonzero only if $\left(\mathbf{k},\vec{k}\right)$ belongs to the set

$$B_\beta = \left\{\left(\mathbf{k},\vec{k}\right) : \left|\mathbf{k}^{(i)} - \vartheta_i \mathbf{k}_{*l_i}\right| \leq \beta^{1-\epsilon}, \quad i = 1,\ldots,m, \quad \left|\mathbf{k} - \sum_i \vartheta_i \mathbf{k}_{*l_i}\right| \leq m\beta^{1-\epsilon}\right\}. \tag{171}$$

We will prove now that if $(n, \mathbf{k}_{*i}) \notin S$, then for small β one of the following alternatives holds:

$$\text{either } \Psi\left(\cdot,\vartheta\mathbf{k}_{*i}\right)\Pi_{n',\vartheta}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right) = 0, \tag{172}$$

$$\text{or (168) holds and } \left|\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)\right| \geq c > 0 \text{ for } \left(\mathbf{k},\vec{k}\right) \in B_\beta. \tag{173}$$

Note that since $\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)$ is smooth, then using notation (94) we get

$$\left|\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right) - \phi_{n',\zeta,\vec{\xi}}\left(\mathbf{k}_{**},\vec{k}_*\right)\right| \leq C\beta^{1-\epsilon} \text{ for } \left(\mathbf{k},\vec{k}\right) \in B_\beta, \tag{174}$$

$$\vec{\vartheta} = (\vartheta_1,\ldots,\vartheta_m), \quad \mathbf{k}_{**} = \zeta\sum_i \vartheta_i \mathbf{k}_{*l_i} = \zeta\varkappa_m\left(\vec{\vartheta},\vec{l}\right).$$

Hence the alternative (173) holds if

$$\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k}_{**},\vec{k}_*\right) \neq 0, \tag{175}$$

and, consequently, it suffices to prove that either (172) or (175) holds. Combining (171) with $\Psi\left(\mathbf{k},\vartheta\mathbf{k}_{*i}\right) = 0$ for $\left|\mathbf{k} - \vartheta\mathbf{k}_{*i}\right| \geq \beta^{1-\epsilon}$ we find that $\Psi_{i,\vartheta}\mathcal{F}^{(m)}\left[\vec{\mathbf{w}}_{\vec{\lambda}}\right]$ can be nonzero for small β only in a small neighborhood of a point $\zeta\varkappa_m\left(\vec{\vartheta},\vec{l}\right) \in [S]_{K,\mathrm{out}}$, and that is possible only if

$$\mathbf{k}_{**} = \zeta\varkappa_m\left(\vec{\vartheta},\vec{l}\right) = \vartheta\mathbf{k}_{*i}, \quad \mathbf{k}_{*i} \in K_S. \tag{176}$$

Let us show that the equality

$$\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k}_{**},\vec{k}_{*}\right)=0 \tag{177}$$

is impossible for $\mathbf{k}_{**}$ as in (176) and $n'=n$ as in (167), keeping in mind that $(n,\mathbf{k}_{*i})\notin S$. It follows from (99) and (134) that Eq. (177) has the form of the resonance equation (100). Since nk-spectrum S is resonance invariant, in view of Definition 18 the resonance equation (177) may have a solution only if $\mathbf{k}_{**}=\mathbf{k}_{*i}$, $i=i_l$, $n=n_l$, with $\left(n_l,\mathbf{k}_{*i_l}\right)\in S$. Since $(n,\mathbf{k}_{*i})\notin S$ that implies (177) does not have a solution and, hence, (175) holds when $(n,\mathbf{k}_{*i})\notin S$. Notice that Theorem 33 and (121) yield bounds

$$\left\|\hat{\mathbf{w}}_{\lambda_i}\right\|_E\leq R_1,\quad \left\|\partial_\tau\hat{\mathbf{w}}_{\lambda_i}\right\|_E\leq C.$$

These bounds combined with Lemma 36, proven below, imply that if (175) holds then (161) holds. Now let us turn to (160). According to (147) and (170) the term $\Psi_\infty\Pi_{n',\vartheta}\mathcal{F}^{(m)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)$ can be non-zero only if $\zeta\varkappa_m\left(\vec{\lambda}\right)=\mathbf{k}_{**}\notin K_S$. Since nk-spectrum S is resonance invariant we conclude as above that inequality (175) holds in this case as well. The fact that the set of all $\varkappa_m\left(\vec{\lambda}\right)$ is finite, combined with inequality (175), imply (173) for sufficiently small β. Using Lemma 36 as above we derive (160). Hence, all terms in the expansion (156) are either zero or satisfy (160) or (161) implying consequently (155) and (151). $\square$

Here is the lemma used in the above proof.

Lemma 36. *Assume that*

$$\left|\Psi_{i,\vartheta'}\Pi_{n',\zeta}\chi^{(m)}_{n,\zeta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)\left[\hat{\mathbf{w}}_{\lambda_1}\left(\mathbf{k}',\tau_1\right),\ldots,\hat{\mathbf{w}}_{\lambda_m}\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right),\tau_1\right)\right]\right|=0\ \textit{for}\ \left(\mathbf{k},\vec{k}\right)\in B_\beta,$$
$$\textit{and}\ \left|\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)\right|\geq\omega_*>0\ \textit{for}\ \left(\mathbf{k},\vec{k}\right)\notin B_\beta,\ \textit{with}\ B_\beta\ \textit{as in (171)}. \tag{178}$$

Then

$$\left\|\Psi\left(\cdot,\vartheta'\mathbf{k}_{*i}\right)\Pi_{n',\zeta}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)\right\|_E\leq \tag{179}$$
$$\frac{4\varrho}{\omega_*}\left\|\chi^{(m)}\right\|\prod_j\left\|\hat{\mathbf{w}}_{\lambda_j}\right\|_E+\frac{2\varrho\tau_*}{\omega_*}\left\|\chi^{(m)}\right\|\sum_i\left\|\partial_\tau\hat{\mathbf{w}}_{\lambda_i}\right\|_E\prod_{j\neq i}\left\|\hat{\mathbf{w}}_{\lambda_j}\right\|_E.$$

Proof. Notice that the oscillatory factor in (133) equals

$$\exp\left\{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)\frac{\tau_1}{\varrho}\right\}=\frac{\varrho}{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)}\partial_{\tau_1}\exp\left\{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)\frac{\tau_1}{\varrho}\right\}.$$

Denoting $\phi_{n,\zeta,\vec{\xi}} = \phi$, $\Psi_{i,\vartheta'}\Pi_{n',\zeta}\chi^{(m)}_{n,\zeta,\vec{\xi}} = \chi^{(m)}_{\vec{\eta}}$ and integrating (133) by parts with respect to τ_1 we obtain

$$\begin{aligned}
&\Psi\left(\mathbf{k},\vartheta'\mathbf{k}_{*i}\right)\Pi_{n',\zeta}\mathcal{F}^{(m)}_{n,\zeta,\vec{\xi}}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)(\mathbf{k},\tau)\\
&= \int_B \Psi\left(\mathbf{k},\vartheta'\mathbf{k}_{*i}\right)\frac{\varrho \mathrm{e}^{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)\frac{\tau}{\varrho}}}{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)}\chi^{(m)}_{\vec{\eta}}\left(\mathbf{k},\vec{k}\right)\hat{\mathbf{w}}_{\lambda_1}\left(\mathbf{k}',\tau\right)\ldots\hat{\mathbf{w}}_{\lambda_m}\\
&\quad\times\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right),\tau\right)\tilde{\mathrm{d}}^{(m-1)d}\vec{k}\\
&\quad-\int_B \Psi\left(\mathbf{k},\vartheta'\mathbf{k}_{*i}\right)\frac{\varrho}{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)}\chi^{(m)}_{\vec{\eta}}\left(\mathbf{k},\vec{k}\right)\hat{\mathbf{w}}_{\lambda_1}\left(\mathbf{k}',0\right)\ldots\hat{\mathbf{w}}_{\lambda_m}\\
&\quad\times\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right),0\right)\tilde{\mathrm{d}}^{(m-1)d}\vec{k}\\
&\quad-\int_0^\tau\int_B \Psi\left(\mathbf{k},\vartheta'\mathbf{k}_{*i}\right)\frac{\varrho \mathrm{e}^{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)\frac{\tau_1}{\varrho}}}{\mathrm{i}\phi\left(\mathbf{k},\vec{k}\right)}\chi^{(m)}_{\vec{\eta}}\left(\mathbf{k},\vec{k}\right)\partial_{\tau_1}\\
&\quad\times\left[\hat{\mathbf{w}}_{\lambda_1}\left(\mathbf{k}'\right)\ldots\hat{\mathbf{w}}_{\lambda_m}\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right)\right)\right]\tilde{\mathrm{d}}^{(m-1)d}\vec{k}d\tau_1,
\end{aligned}\tag{180}$$

where B is the set of $\mathbf{k}^{(i)}$ for which (171) holds. The relations (88) and (25) imply $\left|\chi^{(m)}_{\vec{\eta}}\left(\mathbf{k},\vec{k}\right)\right| \leq \left\|\chi^{(m)}\right\|$. Using then (178), the Leibnitz formula and (122) we obtain (179). □

The main result of this subsection is the next theorem which, when combined with Lemma 34, implies the wavepacket preservation, namely that the solution $\hat{\mathbf{u}}_{n,\vartheta}(\mathbf{k},\tau)$ of (136) is a multi-wavepacket for all $\tau \in [0,\tau_*]$.

Theorem 37. *Assume that conditions of Theorem 35 are fulfilled. Let $\hat{\mathbf{u}}_{n,\vartheta}(\mathbf{k},\tau)$ for $n = n_l$ and $\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k},\tau)$ be the solutions to respective systems (136), (141), $\hat{\mathbf{w}}$ be defined by (146). Then there exists $\beta_0 > 0$ such that*

$$\left\|\hat{\mathbf{u}}_{n_l,\vartheta} - \Pi_{n_l,\vartheta}\hat{\mathbf{w}}\right\|_E \leq C\varrho + C'\beta^s \text{ for } 0 < \beta \leq \beta_0. \tag{181}$$

Proof. Note that $\hat{\mathbf{u}}_{n,\vartheta} = \Pi_{n,\vartheta}\hat{\mathbf{u}}$ where $\hat{\mathbf{u}}$ is a solution of (118) and, according to Theorem 28, $\left\|\hat{\mathbf{u}}\right\|_E \leq 2R$. Comparing Eqs. (118) and (150), which are $\hat{\mathbf{u}} = \mathcal{F}\left(\hat{\mathbf{u}}\right) + \hat{\mathbf{h}}$ and $\hat{\mathbf{w}} = \mathcal{F}\left(\hat{\mathbf{w}}\right) + \hat{\mathbf{h}} + \hat{\mathbf{D}}\left(\hat{\mathbf{w}}\right)$, we find that Lemma 27 can be applied. Then we notice that by Lemma 26 $\mathcal{F}$ has the Lipschitz constant $C_F\tau_*$ for such $\hat{\mathbf{u}}$. Taking $C_F\tau_* < 1$ as in Theorem 28 we obtain (181) from (128). □

Notice that Theorem 5 is a direct corollary of Theorem 37 and Lemma 34. The following corollary shows that inequality (181) and, therefore, Theorems 5 and 3 on preservation of wavepackets hold in the case when the coefficients of operator $\hat{\mathbf{F}}\left(\hat{\mathbf{U}}\right)$ in (3), (86) regularly depends on small ϱ, $\hat{\mathbf{F}}\left(\hat{\mathbf{U}}\right) = \hat{\mathbf{F}}\left(\hat{\mathbf{U}},\varrho\right)$.

Corollary 38 (Parameter dependent nonlinearity). *Assume that conditions of Theorem 35 are fulfilled. Consider a perturbed Eq. (118) $\hat{\mathbf{u}}(\mathbf{k},\tau) = \mathcal{F}\left(\hat{\mathbf{u}}\right)(\mathbf{k},\tau) +$*

$\mathcal{F}_1\left(\hat{\mathbf{u}},\varrho\right)(\mathbf{k},\tau)+\hat{\mathbf{h}}(\mathbf{k})$, *where operator* $\mathcal{F}_1\left(\hat{\mathbf{u}},\varrho\right)$ *satisfies the inequality* $\left\|\mathcal{F}_1\left(\hat{\mathbf{u}},\varrho\right)\right\|_E\leq C\varrho^q$ *for* $\left\|\hat{\mathbf{u}}\right\|_E\leq 2R$ *with some* q, $0<q\leq 1$. *Let* $\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k},\tau)$ *be the solution of (141). Then* $\left\|\Pi_{n,\vartheta}\hat{\mathbf{u}}-\hat{\mathbf{w}}_{l,\vartheta}\right\|_E\leq C\varrho^q+C'\beta^s$.

Proof. The statement follows from (181) and Lemma 27. □

The following theorem shows that any multi-wavepacket solution to (118) yields a solution to the wavepacket interaction system (141).

Theorem 39. *Let* $\hat{\mathbf{u}}(\mathbf{k},\tau)$ *be a solution of* (118) *and assume that* $\hat{\mathbf{u}}(\mathbf{k},\tau)$ *and* $\hat{\mathbf{h}}(\mathbf{k})$ *are multiwavepackets with nk-spectrum* $S=\{(n_l,\mathbf{k}_{*l}),l=1,\ldots,N\}$ *and the regularity degree s. Let also* $\Psi_{i_l,\vartheta}=\Psi_{i_l,\vartheta}$ *be defined by* (137). *Then functions* $\hat{\mathbf{w}}'_{l,\vartheta}(\mathbf{k},\tau)=\Psi_{i_l,\vartheta}\Pi_{n_l,\vartheta}\hat{\mathbf{u}}(\mathbf{k},\tau)$ *are a solution to the system* (141) *with* $\hat{\mathbf{h}}(\mathbf{k})$ *replaced by* $\hat{\mathbf{h}}'(\mathbf{k},\tau)$ *satisfying*

$$\left\|\hat{\mathbf{h}}(\mathbf{k})-\hat{\mathbf{h}}'(\mathbf{k},\tau)\right\|_{L^1}\leq C\beta^s,\ 0\leq\tau\leq\tau_*. \tag{182}$$

Proof. Multiplying (118) by $\Psi_{i_l,\vartheta}\Pi_{n_l,\vartheta}$ we get

$$\begin{aligned}\hat{\mathbf{w}}'_{l,\vartheta}&=\Psi\left(\cdot,\vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\mathcal{F}\left(\hat{\mathbf{u}}\right)(\mathbf{k},\tau)+\Psi\left(\cdot,\vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\hat{\mathbf{h}}(\mathbf{k}),\quad\hat{\mathbf{w}}'_{l,\vartheta}\\&=\Psi\left(\cdot,\vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\hat{\mathbf{u}}.\end{aligned}\tag{183}$$

Since $\hat{\mathbf{u}}(\mathbf{k},\tau)$ is a multiwavepacket with regularity s we have

$$\left\|\hat{\mathbf{u}}(\cdot,\tau)-\hat{\mathbf{w}}'(\cdot,\tau)\right\|_{L^1}\leq C_\epsilon\beta^s\ \text{ where }\ \hat{\mathbf{w}}'(\cdot,\tau)=\sum_{l,\vartheta}\Psi\left(\cdot,\vartheta\mathbf{k}_{*i_l}\right)\hat{\mathbf{u}}(\cdot,\tau). \tag{184}$$

Let us recast (183) in the form

$$\hat{\mathbf{w}}'_{l,\vartheta}=\Psi\left(\cdot,\vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\mathcal{F}\left(\hat{\mathbf{w}}'\right)(\mathbf{k},\tau)+\Psi\left(\cdot,\vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\left[\hat{\mathbf{h}}(\mathbf{k})+\hat{\mathbf{h}}''(\mathbf{k},\tau)\right], \tag{185}$$

$$\hat{\mathbf{h}}''(\mathbf{k},\tau)=\left[\mathcal{F}\left(\hat{\mathbf{u}}\right)-\mathcal{F}\left(\hat{\mathbf{w}}'\right)\right](\mathbf{k},\tau).$$

Denoting $\hat{\mathbf{h}}(\mathbf{k})+\hat{\mathbf{h}}''(\mathbf{k},\tau)=\hat{\mathbf{h}}'(\mathbf{k},\tau)$ we observe that (185) has the form of (141) with $\hat{\mathbf{h}}(\mathbf{k})$ replaced by $\hat{\mathbf{h}}'(\mathbf{k},\tau)$. Inequality (182) follows then from (184) and (124). □

7. Reduction of Wavepacket Interaction System to a Minimal Interaction System

Our goal in this section is to substitute the wavepacket interaction system (141) with a simpler (minimal) *interaction system* which describes the evolution of wavepackets with the same accuracy. We fix the nk-spectrum $S=\{(n_l,\mathbf{k}_{*l}),\ l=1,\ldots,N\}$ of the initial multiwavepacket and assume everywhere below that it is resonance invariant. The minimal interaction system is built based on operators $\mathbf{L}$ and $\hat{\mathbf{F}}\left(\hat{\mathbf{U}}\right)$ and on S. We want the minimal interaction system to satisfy the following requirements. Firstly, the approximation of solutions of (141) by solutions of the minimal interaction system of the order (μ,ν) has to be of the order ϱ in suitable region of parameters (ϱ,β) (which is larger for larger μ,ν). Secondly, the minimal interaction system of the order (μ,ν)

should be defined by S and by the values of $\mathbf{L}(\mathbf{k})$ and its derivatives of the order up to μ and by the values $\chi^{(m)}\left(\mathbf{k}, \vec{k}\right)$ and its derivatives of order up to ν at $\mathbf{k}_{*l} \in S_K$.

The construction of the minimal interaction system consists of the following consecutive steps: (i) introduction of a time averaged wavepacket interaction system obtained by discarding non-resonant terms in the nonlinearity; (ii) reduction of the system for vector components $\hat{\mathbf{v}}_{l,\vartheta}$ to an equivalent one for scalar amplitudes $\hat{v}_{l,\vartheta}$; (iii) change of variables $\mathbf{k} = \vartheta\mathbf{k}_{*l} + \beta\boldsymbol{\eta}$ in the equation for $\hat{v}_{l,\vartheta}$ resulting in a regular dependence of coefficients on small $\beta\boldsymbol{\eta}$; (iv) substitution of the general dependence on $\beta\boldsymbol{\eta}$ in the linear part with a certain polynomial one of the order μ, and the general dependence on $\beta\boldsymbol{\eta}$ of coefficients of the nonlinearity with a certain trigonometric polynomial of the order ν; (v) substitution of the cutoff functions $\Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)$ from (141), which were preserved up to this step, with 1.

As a result we obtain a minimal interaction system with weakly universal nonlinearity, which in the simplest case, where S is just a single element $(\mathbf{k}_*, n)$, is equivalent to the classical NLS equation, and in the case when S consists of only two elements $(\mathbf{k}_*, n)$, $(-\mathbf{k}_*, n)$, is equivalent to the classical coupled modes system.

7.1. Time averaged wavepacket interaction system. Here we modify the wavepacket interaction system (141), substituting its nonlinearity with a certain universal or conditionally universal one obtained by the time averaging, and prove that this substitution produces a small error of order ϱ. As the first step we recast (141) in a slightly different form by using expansions (148), (162) together with (166) and (167) and writing the nonlinearity in Eq. (141) in the form

$$\Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\mathcal{F}(\cdot,\tau) = \sum_{m\in\mathfrak{M}_F}\sum_{\vec{\lambda}\in\Lambda^m}\Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\mathcal{F}^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right),\ \vec{\lambda} = \left(\vec{l}, \vec{\zeta}\right),$$

$$\mathcal{F}^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right)(\mathbf{k},\tau) = \mathcal{F}^{(m)}_{n,\zeta,\vec{n},\vec{\zeta}}\left[\hat{\mathbf{w}}_{\lambda_1}\ldots\hat{\mathbf{w}}_{\lambda_m}\right](\mathbf{k},\tau), \vec{n} = \vec{n}\left(\vec{l}\right),\ \ (n,\zeta) = (n_l,\vartheta),$$

with $\mathcal{F}^{(m)}_{n,\zeta,\vec{n},\vec{\zeta}}$ as in (133) and $\vec{n}\left(\vec{l}\right)$ as in (164). Consequently, the wavepacket interaction system (141) can be written in an equivalent form

$$\hat{\mathbf{w}}_{l,\vartheta} = \sum_{m\in\mathfrak{M}_F}\sum_{\vec{\lambda}\in\Lambda^m}\Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\mathcal{F}^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right) + \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\hat{\mathbf{h}},\quad l=1,\ldots N,\quad \vartheta=\pm. \tag{186}$$

The construction of the above mentioned time averaged equation reduces to discarding certain terms in the original system (186). First we introduce the following sets of indices related to the resonance equation (100) and Ω_m defined by (99):

$$\Lambda^m_{n_l,\vartheta} = \left\{\vec{\lambda} = \left(\vec{l}, \vec{\zeta}\right) \in \Lambda^m : \Omega_m\left(\vartheta, n_l, \vec{\lambda}\right) = 0\right\}, \tag{187}$$

and then the *time-averaged nonlinearity by*

$$\mathcal{F}_{\mathrm{av},n_l,\vartheta}\left(\vec{\mathbf{w}}\right) = \sum_{m\in\mathfrak{M}_F}\mathcal{F}^{(m)}_{n_l,\vartheta},\ \mathcal{F}^{(m)}_{n_l,\vartheta} = \sum_{\vec{\lambda}\in\Lambda^m_{n_l,\vartheta}}\mathcal{F}^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{\mathbf{w}}_{\vec{\lambda}}\right). \tag{188}$$

Note that the nonlinearity $\mathcal{F}^{(m)}_{\mathrm{av},n_l,\vartheta}(\vec{\mathbf{w}})$ can be obtained from $\mathcal{F}^{(m)}_{n_l,\vartheta}$ by the averaging formula (70) where A_T is defined by formula (69) with frequencies $\phi_j = \omega_{n_j}\left(\mathbf{k}_{*i_j}\right)$. Consequently, the desired equation with time-averaged nonlinearity is

$$\hat{\mathbf{v}}_{l,\vartheta} = \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\mathcal{F}_{\mathrm{av},n_l,\vartheta}(\vec{\mathbf{v}}) + \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\Pi_{n_l,\vartheta}\hat{\mathbf{h}}, \quad l = 1, \dots N, \vartheta = \pm, \tag{189}$$

which similarly to (142) we recast concisely as

$$\vec{\mathbf{v}} = \mathcal{F}_{\mathrm{av},\Psi}(\vec{\mathbf{v}}) + \vec{\mathbf{h}}_{\Psi}. \tag{190}$$

The following lemma is analogous to Lemmas 32, 26.

Lemma 40. *Operator $\mathcal{F}_{\mathrm{av},\Psi}(\vec{\mathbf{v}})$ is bounded for bounded $\vec{\mathbf{v}} \in E^{2N}$, $\mathcal{F}_{\mathrm{av},\Psi}(\mathbf{0}) = \mathbf{0}$. Polynomial operator $\mathcal{F}_{\mathrm{av},\Psi}(\vec{\mathbf{v}})$ satisfies the Lipschitz condition*

$$\left\|\mathcal{F}_{\mathrm{av},\Psi}(\vec{\mathbf{v}}_1) - \mathcal{F}_{\mathrm{av},\Psi}(\vec{\mathbf{v}}_2)\right\|_{E^{2N}} \leq C\tau_* \left\|\vec{\mathbf{v}}_1 - \vec{\mathbf{v}}_2\right\|_{E^{2N}}, \tag{191}$$

where C depends only on C_χ as in (88), on the power of $\mathcal{F}$ and on $\|\vec{\mathbf{v}}_1\|_{E^{2N}} + \|\vec{\mathbf{v}}_2\|_{E^{2N}}$, and, in particular, it does not depend on β.

From Lemma 40 and the contraction principle we obtain the following theorem similarly to Theorem 33.

Theorem 41. *Let $\left\|\vec{\mathbf{h}}_\Psi\right\|_{E^{2N}} \leq R$. Then there exists $R_1 > 0$ and $\tau_* > 0$ such that Eq. (190) has a solution $\vec{\mathbf{v}} \in E^{2N}$ satisfying $\|\vec{\mathbf{v}}\|_{E^{2N}} \leq R_1$, and such a solution is unique.*

Theorem 42. *Let $\hat{\mathbf{v}}_{l,\vartheta}(\mathbf{k}, \tau)$ be the solution of (189) and $\hat{\mathbf{w}}_{l,\vartheta}(\mathbf{k}, \tau)$ be the solution of (141). Then the $\hat{\mathbf{v}}_{l,\vartheta}(\mathbf{k}, \tau)$ is a wavepacket satisfying (144), (145) with $\hat{\mathbf{w}}$ replaced by $\hat{\mathbf{v}}$. In addition to that, there exists $\beta_0 > 0$ such that*

$$\left\|\hat{\mathbf{v}}_{l,\vartheta} - \hat{\mathbf{w}}_{l,\vartheta}\right\|_E \leq C\varrho, \quad l = 1, \dots, N; \quad \vartheta = \pm, \text{ for } 0 < \varrho \leq 1, \quad 0 < \beta \leq \beta_0. \tag{192}$$

Proof. Formula (144), (145) for $\hat{\mathbf{v}}_{l,\vartheta}(\mathbf{k}, \tau)$ follow from (189). We note that $\vec{\mathbf{w}}$ is an approximate solution of (189), namely we have an estimate for $\hat{\mathbf{D}}_{\mathrm{av}}\left(\hat{\mathbf{w}}\right) = \hat{\mathbf{w}} - \mathcal{F}_{\mathrm{av},\Psi} - \hat{\mathbf{h}}_\Psi$ which is similar to (150), (151):

$$\left\|\hat{\mathbf{D}}_{\mathrm{av}}\left(\hat{\mathbf{w}}\right)\right\| = \left\|\hat{\mathbf{w}} - \mathcal{F}_{\mathrm{av},\Psi} - \hat{\mathbf{h}}\right\|_E \leq C\varrho, \quad \text{if } 0 < \varrho \leq 1, \quad \beta \leq \beta_0. \tag{193}$$

The proof of (193) is similar to the proof of (155) with minor simplifications thanks to the absence of terms with Ψ_∞. Using (193) we apply Lemma 27 and obtain (192). □

7.2. Averaged system for scalar amplitudes. Now we recast (189) in the form of an equivalent system of scalar equations for *amplitudes* $\hat{v}_{l,\vartheta} = \hat{v}_{\lambda}$ of solutions $\hat{\mathbf{v}}_{\lambda_l}$ defined based on (11), namely

$$\hat{\mathbf{v}}_{\lambda_l}(\mathbf{k}) = \Psi\left(\mathbf{k}, \zeta^{(l)}\mathbf{k}_{*i_l}\right)\Pi_{n_l,\zeta^{(l)}}(\mathbf{k})\,\hat{\mathbf{v}}_{\lambda_l}(\mathbf{k}) = \hat{v}_{l,\zeta^{(l)}}(\mathbf{k})\,\mathbf{g}_{n_l,\zeta^{(l)}}(\mathbf{k}). \tag{194}$$

Note that according to (145) support of $\hat{v}_{l,\zeta^{(l)}}$ is localized near $\zeta\mathbf{k}_{*i_l}$, and we can assume that $\mathbf{g}_{n_l,\zeta^{(l)}}(\mathbf{k})$ depend smoothly on $\mathbf{k}$ near this point. Multiplying (189) by $\mathbf{g}_{n_l,\zeta_l}(\mathbf{k})$ (with the standard scalar product in $\mathbb{C}^{2j}$) and using (194) we obtain the following system of scalar amplitude equations:

$$\hat{v}_{l,\vartheta} = \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right) f_{\mathrm{av},n_l,\vartheta}(\vec{v}) + \Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)\hat{h}_{n_l,\vartheta}, \quad l = 1,\ldots,N,\ \vartheta = \pm, \text{ where} \tag{195}$$

$$\hat{h}_{n_l,\vartheta} = \mathbf{g}_{n_l,\vartheta}\cdot\Pi_{n_l,\vartheta}\hat{\mathbf{h}}, \quad f_{\mathrm{av},n_l,\vartheta}(\vec{v}) = \sum_{m\in\mathfrak{M}_F}\sum_{\vec{\lambda}\in\Lambda^m_{n_l,\vartheta}} f^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{v}_{\vec{\lambda}}\right). \tag{196}$$

According to (169) the m-linear operators in the above equation are given by

$$f^{(m)}_{n,\vartheta,\vec{\xi}}\left(\vec{v}_{\vec{\lambda}}\right)(\mathbf{k},\tau) = \int_0^{\tau}\int_{\mathbb{D}_m} \mathrm{e}^{\mathrm{i}\phi_{n,\vartheta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)\frac{\tau_1}{\varrho}} Q^{(m)}_{n,\vartheta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)\prod_{i=1}^{m}\hat{v}_{\lambda_i}\tilde{\mathrm{d}}^{(m-1)d}\vec{k}\,\mathrm{d}\tau_1, \tag{197}$$

$$Q^{(m)}_{n,\vartheta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right) = \mathbf{g}_{n,\vartheta}(\mathbf{k})\cdot\chi^{(m)}_{n,\vartheta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)\left[\mathbf{g}_{\lambda_1}(\mathbf{k}'),\ldots,\mathbf{g}_{\lambda_m}\left(\mathbf{k}^{(m)}\left(\mathbf{k},\vec{k}\right)\right)\right]. \tag{198}$$

The concise form for the system (195) of scalar equations for amplitudes is

$$\vec{v} = f_{\Psi}(\vec{v}) + \hat{h}_{\Psi}, \quad \vec{v}\in E^{2N}_{\mathrm{sc}}, \tag{199}$$

where the components $\hat{v}_{l,\vartheta}$ of $\vec{v}$ belong to the space E_{sc} of scalar functions with the norm defined by (17), (18) applied to scalar functions. Note that $Q^{(m)}_{n,\vartheta,\vec{\xi}}\left(\mathbf{k},\vec{k}\right)$ can be extended in an arbitrary way as bounded functions for arguments $\mathbf{k}, \vec{k}$, where (171) is not satisfied, for example the extension can be zero, the extension does not affect solutions of (195) because this equation involves factors $\Psi\left(\cdot, \vartheta\mathbf{k}_{*i_l}\right)$ and (145) holds.

Lemma 43. *Operator f_{Ψ} is bounded for bounded $\vec{v}\in E^{2N}_{\mathrm{sc}}$ and $f_{\Psi}(\mathbf{0}) = \mathbf{0}$. The polynomial operator $f_{\Psi}(\vec{v})$ satisfies the Lipschitz condition*

$$\|f_{\Psi}(\vec{v}_1) - f_{\Psi}(\vec{v}_2)\|_{E^{2N}_{\mathrm{sc}}} \le C\tau_*\|\vec{v}_1 - \vec{v}_2\|_{E^{2N}_{\mathrm{sc}}},$$

where C depends only on C_{χ} as in (88), on the order of $\mathcal{F}$ as a polynomial and on $\|\vec{v}_1\|_{E^{2N}} + \|\vec{v}_2\|_{E^{2N}}$, and it does not depend on β.

From Lemma 40 and the contraction principle we obtain the following theorem similarly to Theorem 33.

Theorem 44. *Let $\left\|\hat{h}_{\Psi}\right\|_{E^{2N}_{\mathrm{sc}}} \le R$. Then there exists $R_1 > 0$ and $\tau_* > 0$ such that* (199) *has a solution $\vec{v}\in E^{2N}_{\mathrm{sc}}$ satisfying $\|\vec{v}\|_{E^{2N}_{\mathrm{sc}}} \le R_1$, and such a solution is unique.*

7.3. Rescaled amplitude equations. According to (145) amplitudes $\hat{v}_{l,\vartheta}(\zeta\mathbf{k}_{*l}+\boldsymbol{\eta})$ are localized about the point $\boldsymbol{\eta}=\mathbf{0}$, and to study its behavior in a vicinity of $\boldsymbol{\eta}=\mathbf{0}$ we introduce a group of dilation operators

$$\left(B_\beta\hat{v}\right)(\boldsymbol{\eta})=\beta^d\hat{v}(\beta\boldsymbol{\eta}),\quad \beta>0, \tag{200}$$

which preserve the L^1-norm and commute with the convolution, i.e.

$$\left\|B_\beta\hat{v}\right\|_{L^1}=\left\|\hat{v}\right\|_{L^1},\quad B_\beta\hat{v}*B_\beta\hat{w}=B_\beta\left(\hat{v}*\hat{w}\right). \tag{201}$$

We introduce then a rescaled and shifted version of initial data $\hat{h}_{n_l,\vartheta}$ in (196) by the formula

$$\hat{H}_{n_l,\vartheta}(\mathbf{k})=B_\beta\hat{h}_{n_l,\vartheta}(\mathbf{k}+\vartheta\mathbf{k}_{*l}),\ \hat{h}_{n_l,\vartheta}(\mathbf{k})=\beta^{-d}\hat{H}_{n_l,\vartheta}\left(\beta^{-1}(\mathbf{k}-\vartheta\mathbf{k}_{*l})\right), \tag{202}$$

where B_β is defined by (200), $|\mathbf{k}-\vartheta\mathbf{k}_{*l}|\leq\beta^{1-\epsilon}$, and new variables

$$\boldsymbol{\eta}_l=\beta^{-1}(\mathbf{k}-\vartheta\mathbf{k}_{*l}),\quad l=1,\ldots,N,\quad \vec{\eta}=\left(\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_N\right). \tag{203}$$

In this and the following sections we assume that $\hat{H}_{n_l,\vartheta}(\beta,\boldsymbol{\eta})$ are defined for all $\boldsymbol{\eta}\in\mathbb{R}^d$, including $|\boldsymbol{\eta}|\geq\beta^{-\epsilon}$. Though (195) involves $\hat{h}_{n_l,\vartheta}$ with a cutoff factor, namely $\Psi\left(\mathbf{k},\vartheta\mathbf{k}_{*i_l}\right)\hat{h}_{n_l,\vartheta}(\mathbf{k})=\Psi\left(\mathbf{k},\vartheta\mathbf{k}_{*i_l},\beta^{1-\epsilon}\right)\hat{h}_{n_l,\vartheta}(\mathbf{k})$ as in (26), we will later use $\hat{H}_{n_l,\vartheta}(\beta,\boldsymbol{\eta})$ defined for all $\boldsymbol{\eta}$, and assume that

$$\left\|\left(1-\Psi\left(\beta^\epsilon\boldsymbol{\eta}\right)\right)\hat{H}_{n_l,\vartheta}(\beta,\boldsymbol{\eta})\right\|_{L^1}\leq C\beta^s, \tag{204}$$

where (i) $\Psi(\beta^\epsilon\boldsymbol{\eta})=\Psi\left(\boldsymbol{\eta},0,\beta^{-\epsilon}\right)$ is as in (25), (26); (ii) ϵ and s are the same as in Definition 1; (iii) condition (204) is consistent with (29) and (30).

For a solution $\hat{v}_{l,\vartheta}(\mathbf{k},\tau)$ of (195) using (145) we introduce the following functions.

$$\hat{z}_{l,\vartheta}(\boldsymbol{\eta},\tau)=\beta^d\hat{v}_{l,\vartheta}(\vartheta\mathbf{k}_{*l}+\beta\boldsymbol{\eta},\tau),\ \hat{z}_{l,\vartheta}(\boldsymbol{\eta},\tau)=\Psi\left(\beta^\epsilon\boldsymbol{\eta}\right)\hat{z}_{l,\vartheta}(\boldsymbol{\eta},\tau),\ \boldsymbol{\eta}\in\mathbb{R}^d, \tag{205}$$

which satisfy a rescaled version of (195) provided below. Note that since $\left(\vec{n},\vec{\zeta}\right)=\vec{\lambda}\in\Lambda^m_{n_l,\vartheta}$ and the nk-spectrum S is resonance invariant we have $\varkappa_m\left(\vec{\lambda}\right)=\sum_i\zeta^{(i)}\mathbf{k}_{*l_i}=\zeta\mathbf{k}_{*l}=\vartheta\mathbf{k}_{*l}$. Since $\mathbf{k},\vec{k}$ satisfy the convolution identity (87) the variables $\boldsymbol{\eta},\vec{\eta}$ defined by (203) satisfy a similar identity as well, namely

$$\boldsymbol{\eta}=\sum_{i=1}^m\boldsymbol{\eta}^{(i)},\quad \boldsymbol{\eta}^{(m)}(\mathbf{k},\vec{\eta})=\boldsymbol{\eta}-\sum_{i=1}^{m-1}\boldsymbol{\eta}^{(i)}. \tag{206}$$

Change of variables (203) in the integral operator $f_{\mathrm{av},n_l,\vartheta}$ defined by (197) yields the following *amplitude system* for $z_{l,\vartheta}$ which is equivalent to (195):

$$\hat{z}_{l,\vartheta}(\boldsymbol{\eta})=\Psi\left(\beta^\epsilon\boldsymbol{\eta}\right)f_{\mathrm{av},n_l,\vartheta,\beta}(\vec{z})(\boldsymbol{\eta})+\Psi\left(\beta^\epsilon\boldsymbol{\eta}\right)\hat{H}_{n_l,\vartheta}(\boldsymbol{\eta}),\quad l=1,\ldots N,\quad \vartheta=\pm. \tag{207}$$

According to (137), (196) and (197),

$$\Psi\left(\mathbf{k},\vartheta\mathbf{k}_{*i_l},\beta^{1-\epsilon}\right)=\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right),\quad f_{\mathrm{av},n_l,\vartheta,\beta}\left(\vec{z}\right)=\sum_{m\in\mathfrak{M}_F}f^{(m)}_{\mathrm{av},n_l,\vartheta,\beta}\left(\vec{z}\right),\tag{208}$$

$$f^{(m)}_{\mathrm{av},n_l,\vartheta,\beta}\left(\vec{z}\right)=\sum_{\vec{\lambda}\in\Lambda^m_{n_l,\vartheta}}f^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right),\beta}\left(\vec{z}_{\vec{\lambda}}\right),$$

$$f^{(m)}_{n,\vartheta,\vec{\xi}\left(\vec{\lambda}\right),\beta}\left(\vec{z}_{\vec{\lambda}}\right)(\boldsymbol{\eta},\tau)=\int_0^{\tau}\int_{\boldsymbol{\eta}'+\cdots+\boldsymbol{\eta}^{(m)}=\boldsymbol{\eta}}\exp\left\{\mathrm{i}\phi_{n,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vartheta\mathbf{k}_{*l}+\beta\boldsymbol{\eta},\vec{k}_*+\beta\vec{\eta}\right)\frac{\tau_1}{\varrho}\right\}\tag{209}$$

$$Q^{(m)}_{n,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vartheta\mathbf{k}_{*l}+\beta\boldsymbol{\eta},\vec{k}_*+\beta\vec{\eta}\right)\prod\nolimits_{i=1}^{m}\hat{z}_{\lambda_i}\left(\boldsymbol{\eta}^{(i)}\right)\tilde{\mathrm{d}}^{(m-1)d}\vec{\eta}\,\mathrm{d}\tau_1.$$

Note that the condition (171) on the domain of integration takes in the new variables the form

$$\left|\boldsymbol{\eta}^{(i)}\right|\leq\beta^{-\epsilon},\quad i=1,\ldots,m\text{ and }|\boldsymbol{\eta}|\leq m\beta^{-\epsilon}.\tag{210}$$

Finally, we rewrite the *amplitude system* (207) in the concise form

$$\vec{z}=\Psi\left(\beta^{\epsilon}\cdot\right)f_{\mathrm{av},\beta}\left(\vec{z}\right)+\Psi\left(\beta^{\epsilon}\cdot\right)\hat{H}_{\beta},\ \vec{z}\in E^{2N}_{\mathrm{sc}}.\tag{211}$$

Let us show now that (211) is of the form of (118) with $2J$-component vector $\hat{\mathbf{u}}$ substituted with $2N$-component vector $\vec{z}$, the matrix $\mathbf{L}(\mathbf{k})$ substituted with a diagonal matrix $\vec{L}$ with entries $\vartheta\omega_{n_l}(\vartheta\mathbf{k}_{*l}+\beta\boldsymbol{\eta})$. For that we introduce the *S-averaged tensor* $Q^{(m)}_{\mathrm{av}}$ defined on $\vec{z}\in\mathbb{C}^{2Nm}$ by the formula

$$Q^{(m)}_{\mathrm{av},n,\vartheta}\left(\beta\boldsymbol{\eta},\beta\vec{\eta},\vec{z}\right)=\sum_{\vec{\lambda}\in\Lambda^m_{n,\vartheta}}Q^{(m)}_{n,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vartheta\mathbf{k}_{*l}+\beta\boldsymbol{\eta},\vec{k}_*+\beta\vec{\eta}\right)\prod\nolimits_{i=1}^{m}\hat{z}_{\lambda_i}\tag{212}$$

which depends on S through $\Lambda^m_{n,\vartheta}$ and acts from $\mathbb{C}^{2Nm}$ into $\mathbb{C}^{2N}$. Note that $\hat{z}_{\lambda_i}$ and $Q^{(m)}_{n,\vartheta,\vec{\xi}}$ are scalar factors, $\hat{z}_{\lambda_i}$ is a scalar projection in $\mathbb{C}^{2N}$ onto a line along the λ_i^{th} eigenvector of $\vec{L}$. Hence, the right-hand side of (212) is a sum of elementary susceptibilities obtained from $Q^{(m)}_{\mathrm{av}}$ as in (132) and (207) has the form of (136). Note that non-zero terms in (212) contain products $\hat{z}_{\lambda_i}$ which satisfy (100). Therefore, if $\beta=0$ and S is resonance invariant, $Q^{(m)}_{\mathrm{av}}$ has the form of weakly universal nonlinearity; if S is universally resonance invariant then $Q^{(m)}_{\mathrm{av}}$ has the form of a universal nonlinearity as in (65).

7.4. Amplitude system with polynomial dispersion relations. Now we introduce an *amplitude system with polynomial dispersion* which is similar to (207) and provides (i) sufficiently accurate approximation to (207); (ii) standard polynomial dependence of coefficients on $\boldsymbol{\eta}$, $\vec{\eta}$ in the sense clarified below. The amplitude system has the form

$$\hat{u}_{l,\vartheta}=\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right)f^{(\mu,\nu)}_{n_l,\vartheta}\left(\vec{u}\right)+\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right)\hat{H}_{n_l,\vartheta},\quad l=1,\ldots N,\ \vartheta=\pm,\tag{213}$$

$$f^{(\mu,\nu)}_{n_l,\vartheta}\left(\vec{u}\right)=\sum_{m\in\mathfrak{M}_F}\sum_{\vec{\lambda}\in\Lambda^m_{n_l,\vartheta}}f^{(m,\mu,\nu)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}\left(\vec{u}_{\vec{\lambda}}\right),\tag{214}$$

where $\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right)$ are cutoff-factors defined in (208), (137) and approximations $f^{(m,\mu,\nu)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}$ for $f^{(m)}_{n_l,\vartheta,\vec{\xi}\left(\vec{\lambda}\right)}$ are defined below. The indices $\mu = 1, 2$, $\nu = 0, 1$ determine the order of approximation: (i) μ determines the order of approximation of the dispersion relation by a polynomial of the degree μ; (ii) ν determines the order of approximation of the susceptibility coefficients (198) by a trigonometric polynomial of the degree ν. As before, we recast (213) in a concise form,

$$\vec{u} = \Psi_{\beta} f^{(\mu,\nu)}\left(\vec{u}\right) + \Psi_{\beta}\hat{H}, \tag{215}$$

where $\Psi_{\beta}\left(\boldsymbol{\eta}\right) = \Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right)$. Finally, we eliminate in (213) the cutoff factor $\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right)$ by setting $\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right) = \Psi\left(\mathbf{0}\right) = 1$, and introduce the amplitude system with *weakly universal nonlinearity and polynomial dispersion* without cutoff

$$\hat{u}_{l,\vartheta}\left(\boldsymbol{\eta}\right) = f^{(\mu,\nu)}_{n_l,\vartheta}\left(\vec{u}\right)\left(\boldsymbol{\eta}\right) + \hat{H}_{n_l,\vartheta}\left(\boldsymbol{\eta}\right), \quad l = 1, \dots N, \quad \vartheta = \pm, \tag{216}$$

which can be written in the form of (215) with $\Psi_{\beta} = 1$.

Let us turn now to the construction of the approximations. For every nk-pair $\left(\mathbf{k}_{*l}, n_l\right)$ we introduce the Taylor polynomials of order μ of the dispersion relation $\omega_{n_l}\left(\mathbf{k}_{*l} + \beta\boldsymbol{\eta}\right)$:

$$\gamma_1\left(\mathbf{k}_{*l}, n_l, \beta\boldsymbol{\eta}\right) = \omega_{n_l}\left(\mathbf{k}_{*l}\right) + \beta\omega'_{n_l}\left(\mathbf{k}_{*l}\right)\boldsymbol{\eta},$$
$$\gamma_2\left(\mathbf{k}_{*l}, n_l, \beta\boldsymbol{\eta}\right) = \gamma_1\left(\mathbf{k}_{*l}, n_l, \beta\boldsymbol{\eta}\right) + \frac{\beta^2}{2}\left(\boldsymbol{\eta}, \omega''_{n_l}\left(\mathbf{k}_{*l}\right)\boldsymbol{\eta}\right),$$

and similarly γ_3 for $\mu = 3$. Obviously we have the inequality (see (171))

$$\left|\omega_{n_l}\left(\mathbf{k}_{*l} + \beta\boldsymbol{\eta}\right) - \gamma_{\mu}\left(\mathbf{k}_{*l}, n_l, \beta\boldsymbol{\eta}\right)\right| \leq C\beta^{(\mu+1)(1-\epsilon_1)}, \quad \left(\mathbf{k}, \vec{k}\right) \in B_{\beta}. \tag{217}$$

The phase function $\phi_{n,\zeta,\vec{\xi}}\left(\mathbf{k}, \vec{k}\right)$, $\vec{\xi} = \left(\vec{n}, \vec{\zeta}\right)$, defined by (134), is approximated then by a polynomial phase function

$$\begin{aligned}&\phi^{(\mu)}_{n_l,\zeta,\vec{\xi}}\left(\zeta\mathbf{k}_{*l}, \vec{k}_*, \beta\boldsymbol{\eta}, \beta\vec{\eta}\right)\\ &= \zeta\gamma_{\mu}\left(\mathbf{k}_{*l}, n_l, \beta\boldsymbol{\eta}\right) - \zeta'\gamma_{\mu}\left(\mathbf{k}_{*l_1}, n', \beta\boldsymbol{\eta}'\right) - \ldots - \zeta^{(m)}\gamma_{\mu}\left(\mathbf{k}_{*l_m}, n^{(m)}, \beta\boldsymbol{\eta}^{(m)}\right).\end{aligned} \tag{218}$$

Note that since $\vec{\xi} = \vec{\xi}\left(\vec{\lambda}\right)$ with $\vec{\lambda} \in \Lambda^m_{n_l,\vartheta}$ defined by (187), Eq. (100) is fulfilled. Hence, $\phi^{(\mu)}_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \mathbf{0}, \mathbf{0}\right) = 0$ and the function $\phi^1_{n_l,\vartheta,\vec{\xi}}$ depends linearly on $\boldsymbol{\eta}, \vec{\eta}$ and $\phi^2_{n_l,\vartheta,\vec{\xi}}$ is quadratic, namely

$$\phi^1_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \beta\boldsymbol{\eta}, \beta\vec{\eta}\right) = \beta\phi^1_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta}\right), \tag{219}$$

$$\phi^2_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \beta\boldsymbol{\eta}, \beta\vec{\eta}\right) = \beta\phi^1_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta}\right) + \beta^2\phi^{2,2}_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta}\right). \tag{220}$$

In the case $\mu = 2$ the polynomial phase function involves two parameters ϱ_1, ϱ_2:

$$\begin{aligned} &\phi^2_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \beta\boldsymbol{\eta}, \beta\vec{\eta}\right)\frac{\tau_1}{\varrho} \\ &\quad = \mathrm{i}\phi^1_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta}\right)\frac{\tau_1}{\varrho_1} + \mathrm{i}\phi^{2,2}_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta}\right)\frac{\tau_1}{\varrho_2}, \end{aligned} \tag{221}$$

$$\varrho_1 = \frac{\varrho}{\beta}, \quad \varrho_2 = \frac{\varrho}{\beta^2}; \quad 0 < \varrho_1 < \infty, \quad 0 < \varrho_2 \le \infty, \tag{222}$$

where ϱ_1 and ϱ_2 may be large or small depending on the relation between ϱ and β. Sometimes it is convenient to consider ϱ_1 and ϱ_2 as independent parameters. If $\mu = 1$ we formally set $\varrho_2 = \infty$, $\frac{\tau_1}{\varrho_2} = 0$. If (171) holds we have the estimate

$$\left| \mathrm{e}^{\left\{\mathrm{i}\phi^{\mu}_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l},\vec{k}_*,\beta\boldsymbol{\eta},\beta\vec{\eta}\right)\frac{\tau_1}{\varrho}\right\}} - \mathrm{e}^{\left\{\mathrm{i}\phi_{n_l,\vartheta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}+\beta\boldsymbol{\eta},\vec{k}_*+\beta\vec{\eta}\right)\frac{\tau_1}{\varrho}\right\}} \right| \le C\tau_* \frac{\beta^{(\mu+1)(1-\epsilon)}}{\varrho}, \ \mu = 1, 2. \tag{223}$$

To ensure that the approximation error is small for given μ we assume that ϱ and β satisfy

$$\varrho \to 0, \ \beta \to 0, \quad \frac{\beta^{(\mu+1)(1-\epsilon)}}{\varrho} \to 0. \tag{224}$$

Now we approximate the dependence of $Q^{(m)}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l} + \beta\boldsymbol{\eta}, \vec{k}_* + \beta\vec{\eta}\right)$ on $\boldsymbol{\eta}, \vec{\eta}$ given by (198) by trigonometric polynomials. Zero order approximation with $\nu = 0$ is given by

$$Q^{(m,0)}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l} + \beta\boldsymbol{\eta}, \vec{k}_* + \beta\vec{\eta}\right) = Q^{(m)}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*\right). \tag{225}$$

To define the first order approximation we modify the standard Taylor expansion using trigonometric polynomials instead of algebraic ones. Taking the first derivative with respect to β at $\beta = 0$,

$$Q^{(m)\prime}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \boldsymbol{\eta}, \vec{k}_*, \vec{\eta}\right) = \left.\frac{d}{d\beta}\right|_{\beta=0} Q^{(m)}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l} + \beta\boldsymbol{\eta}, \vec{k}_* + \beta\vec{\eta}\right),$$

which obviously is a linear function with respect to $\boldsymbol{\eta}, \vec{\eta}$, we express then $\boldsymbol{\eta}$ in terms of $\vec{\eta}$ using (206):

$$Q^{(m)\prime}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \boldsymbol{\eta}, \vec{k}_*, \vec{\eta}\right) = \sum_{j=1}^{m} q^{(m),j}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*\right)\cdot\boldsymbol{\eta}^{(j)}, \quad \boldsymbol{\eta}^{(j)} = \left(\eta^{(j)}_1, \dots, \eta^{(j)}_d\right).$$

Then the first order approximation is

$$Q^{(m,1)}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l} + \beta\boldsymbol{\eta}, \vec{k}_* + \beta\vec{\eta}\right) = Q^{(m)}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*\right) + \sum_{j=1}^{m} q^{(m),j}_{n,\zeta,\vec{\xi}}\left(\vartheta\mathbf{k}_{*l}, \vec{k}_*\right)\cdot\sin\beta\boldsymbol{\eta}^{(j)},$$

where $\sin\boldsymbol{\eta}^{(j)} = \left(\sin\eta^{(j)}_1, \dots, \sin\eta^{(j)}_d\right)$. An advantage of this approximation is that the multiplication by $\sin\eta^{(j)}_1$ is a bounded operator which equals the Fourier transform of a finite-difference operator whereas the multiplication by $\eta^{(j)}_1$ corresponds to the partial derivative and is unbounded. Since the original nonlinearity does not involve unbounded operators, the use of bounded operators is natural and convenient. In fact, it is well known that the presence of the derivatives in the nonlinearity of NLS-type equations causes well known technical difficulties, see 14]. In our approach the approximating equation provides the same accuracy and its nonlinearity involves only bounded finite-difference operators bypassing those difficulties altogether.

According to Condition 16 the susceptibility is smooth and if (210) holds we have the following inequality:

$$\left| Q^{(m)}_{n,\zeta,\vec{\xi}} \left(\vartheta \mathbf{k}_{*l} + \beta \boldsymbol{\eta}, \vec{k}_* + \beta \vec{\eta} \right) - Q^{(m,\nu)}_{n,\zeta,\vec{\xi}} \left(\vartheta \mathbf{k}_{*l}, \vec{k}_*, \beta \boldsymbol{\eta}, \beta \vec{\eta} \right) \right| \leq C \beta^{(\nu+1)(1-\epsilon_1)}. \quad (226)$$

We introduce components $f^{(m,\mu,\nu)}_{n_l,\vartheta,\vec{\lambda}}$ of the weakly universal nonlinearity $f^{(\mu,\nu)}$ by the formula

$$f^{(m,\mu,\nu)}_{n_l,\vartheta,\vec{\lambda}} \left(\vec{z}_{\vec{\lambda}} \right) (\boldsymbol{\eta}, \tau) = \int_0^\tau \int_{\boldsymbol{\eta}' + \dots + \boldsymbol{\eta}^{(m)} = \boldsymbol{\eta}} \mathrm{e}^{\mathrm{i}\phi^1_{n_l,\vartheta,\vec{\xi}} \left(\vartheta \mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta} \right) \frac{\tau_1}{\varrho_1} + \mathrm{i}\phi^{2,2}_{n_l,\vartheta,\vec{\xi}} \left(\vartheta \mathbf{k}_{*l}, \vec{k}_*, \boldsymbol{\eta}, \vec{\eta} \right) \frac{\tau_1}{\varrho_2}} \quad (227)$$
$$Q^{(m,\nu)}_{n_l,\vartheta,\vec{\xi}} \left(\vartheta \mathbf{k}_{*l}, \vec{k}_* \right) \prod_{i=1}^m \hat{z}_{\lambda_i} \left(\boldsymbol{\eta}^{(i)} \right) \tilde{\mathrm{d}}^{(m-1)d} \vec{k} \, \mathrm{d}\tau_1.$$

As before, we establish standard properties of the operator $f^{(\mu,\nu)}$ defined by the above formula.

Lemma 45. *Operator* $\Psi_\beta f^{(\mu,\nu)}$ *is bounded for bounded* $\vec{u} \in E^{2N}_{\mathrm{sc}}$, $f_\Psi(\mathbf{0}) = \mathbf{0}$. *The polynomial operator* $\Psi_\beta f^{(\mu,\nu)}$ *satisfies the Lipschitz condition*

$$\left\| \Psi_\beta f^{(\mu,\nu)}(\vec{u}_1) - \Psi_\beta f^{(\mu,\nu)}(\vec{u}_2) \right\|_{E^{2N}_{\mathrm{sc}}} \leq C \tau_* \left\| \vec{u}_1 - \vec{u}_2 \right\|_{E^{2N}_{\mathrm{sc}}}, \quad (228)$$

where C *depends only on* C_χ *as in (88), on the power of* $\mathcal{F}$ *and on* $\|\vec{u}_1\|_{E^{2N}_{\mathrm{sc}}} + \|\vec{u}_2\|_{E^{2N}_{\mathrm{sc}}}$. *In particular, it does not depend on* $\beta \geq 0$ *and on* $0 < \varrho_1 < \infty$, $0 < \varrho_2 \leq \infty$.

From Lemma 40 and the contraction principle we obtain the following theorem completely similar to Theorem 33.

Theorem 46. *Let* $\left\| \hat{h}_\Psi \right\|_{E^{2N}_{\mathrm{sc}}} \leq R$. *Then there exists* $R_1 > 0$ *and* $\tau_* > 0$ *such that Eq.* (190) *has a solution* $\vec{z} \in E^{2N}_{\mathrm{sc}}$ *satisfying* $\|\vec{z}\|_{E^{2N}_{\mathrm{sc}}} \leq R_1$. *Such a solution is unique and* $\hat{z}_{l,\vartheta}(\mathbf{k},\tau) = 0$ *if* $|\mathbf{k}| \geq \beta^{-\epsilon}$.

Theorem 47. *Let* $\hat{u}_{l,\vartheta}(\mathbf{k},\tau)$ *be a solution to* (213) *and* $\hat{z}_{l,\vartheta}(\mathbf{k},\tau)$ *be the solution of* (211). *Then the following inequality holds*:

$$\left\| \hat{u}_{l,\vartheta} - \hat{z}_{l,\vartheta} \right\|_{E_{\mathrm{sc}}} \leq C \beta^{(\mu+1)(1-\epsilon)} + C \varrho^{-1} \beta^{(\mu+1)(1-\epsilon)}, \quad l = 1, \dots, N; \quad \vartheta = \pm, \quad (229)$$

for all $0 < \varrho \leq 1$ *and* $0 < \beta \leq \beta_0$, *where* ϵ *is the same as in Definition 1,* β_0 *is sufficiently small.*

Proof. To obtain (229) we note that $u_{l,\vartheta}$ is an approximate solution of (211), namely

$$\vec{u} - \Psi_\beta f^{(\mu,\nu)}(\vec{u}) - \hat{h}_\Psi = \hat{D} \text{ where } \hat{D} \text{ is small.}$$

To estimate $\left\|\hat{D}\right\|$ observe that integrals involving $\vec{u}$ have the integration domain as in (171). Hence, using (226) and (223) we obtain

$$\left\|\hat{D}\right\|_{E^{2N}_{\mathrm{sc}}} \leq C\beta^{(\mu+1)(1-\epsilon)} + C\varrho^{-1}\beta^{(\mu+1)(1-\epsilon)},$$

and applying Lemma 27 we get (229). □

7.5. Decay of solutions and elimination of cutoff factors. In this subsection we show how to remove the cutoff function in (213) and to obtain the averaged interaction system with a weakly universal nonlinearity. If $\mu = 1$, $\nu = 0$ and the nk-spectrum S is resonance-invariant, the amplitude system coincides with the system (62) with a weakly universal nonlinearity. For $\mu > 1$ or $\nu > 0$ the amplitude system involves additional terms. In particular, if $\mu = 2$, $\nu = 0$ and $S = \{(\mathbf{k}_*, n)\}$ is just a single element then the linear part has the second order and the nonlinearity is universal, and the amplitude system turns into the classical NLS system:

$$\partial_\tau u_\zeta = \zeta \frac{1}{\varrho}\gamma_2(\mathbf{k}_*, n, -i\zeta\beta\nabla_r\boldsymbol{\eta}) + b_\zeta u_\zeta^2 u_{-\zeta}, \quad u_\zeta(0) = \hat{H}_\zeta, \quad \zeta = \pm.$$

This system is equivalent to (51) when $\hat{H}_- = \hat{H}_+^*$, $b_- = b_+^*$, $u_- = u_+^*$. When $\nu > 0$ the nonlinearity involves additional terms with finite difference operators.

The possibility to remove cutoff functions is based on the fast decay of $\hat{u}(\mathbf{k})$ as $|\mathbf{k}| \to \infty$, which is equivalent to high smoothness of $u(\mathbf{r})$. The factor Ψ_β can be replaced by 1 with a small error when data $\hat{H}(\mathbf{k})$ decay sufficiently fast. To describe the decay we introduce weighted Banach spaces of scalar functions $\hat{H}(\mathbf{k})$ described as follows.

Definition 48 (Weight function). *For $a \geq 0$ we call a positive function $\psi(r)$, $r \geq 0$, a weight function from class $W(a)$ if it satisfies the following conditions: (i) $\psi(0) > 0$, $\psi(r_1) \geq \psi(r_2)$ for $r_1 \geq r_2 \geq 0$; (ii) $\psi(r_1 + r_2) \leq \psi(r_1) + \psi(r_2) + C$, where C does not depend on r_1, r_2 (ψ is sublinear); (iii) $\psi(r) - a\ln r \geq C' > 0$ for all $r > 0$ ($\psi(r)$ is superlogarithmic).*

We introduce $L^1(\psi)$ as a space of scalar functions $\hat{H}(\mathbf{k})$, $\mathbf{k} \in \mathbb{R}^d$ with the norm

$$\left\|\hat{H}\right\|_{L^1(\psi)} = \int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}|)} \left|\hat{H}(\mathbf{k})\right| \mathrm{d}\mathbf{k}. \tag{230}$$

For vector-functions we use the same formula with Euclidean norm $|\cdot|$. In the simplest case of $\psi(r) = a\ln(1 + r)$ we have $\psi \in W(a)$ and obtain $L^1(\psi) = L^{1,a}$ with the norm (19). If the weight function belongs to $W(a)$ for all a the space $L^1(\psi)$ consists of the Fourier transforms of infinitely smooth functions. The following lemma shows that $L^1(\psi)$ is closed with respect to the convolution.

Lemma 49. *Let* $\hat{H}_1, \hat{H}_2 \in L^1(\psi)$ *and*

$$\hat{H}_3(\mathbf{k}) = \int_{\mathbb{R}^d} \hat{H}_1\left(\mathbf{k}-\mathbf{k}'\right) \hat{H}_2\left(\mathbf{k}-\mathbf{k}'\right) \mathrm{d}\mathbf{k}'.$$

$$\text{Then } \left\| \hat{H}_3(\mathbf{k}) \right\|_{L^1(\psi)} \leq C \left\| \hat{H}_1(\mathbf{k}) \right\|_{L^1(\psi)} \left\| \hat{H}_1(\mathbf{k}) \right\|_{L^1(\psi)}. \tag{231}$$

Proof. Using Definition 48 (ii) we obtain

$$\begin{aligned} \mathrm{e}^{\psi(|\mathbf{k}|)} \left| \hat{H}_3(\mathbf{k}) \right| &\leq \int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}|)} \left| \hat{H}_1\left(\mathbf{k}-\mathbf{k}'\right) \right| \left| \hat{H}_2\left(\mathbf{k}'\right) \right| \mathrm{d}\mathbf{k}' \\ &\leq \mathrm{e}^C \int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}'|)} \mathrm{e}^{\psi(|\mathbf{k}-\mathbf{k}'|)} \left| \hat{H}_1\left(\mathbf{k}-\mathbf{k}'\right) \right| \left| \hat{H}_2\left(\mathbf{k}'\right) \right| \mathrm{d}\mathbf{k}'. \end{aligned}$$

Applying Young's inequality (122) we obtain

$$\int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}|)} \left| \hat{H}_3(\mathbf{k}) \right| \mathrm{d}\mathbf{k} \leq \mathrm{e}^C \int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}|)} \left| \hat{H}_1(\mathbf{k}) \right| \mathrm{d}\mathbf{k}' \int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}|)} \left| \hat{H}_2(\mathbf{k}) \right| \mathrm{d}\mathbf{k}',$$

implying (231). □

Let us introduce the norm in the space $E_{\mathrm{sc}}(\psi)$ by the formula (17)

$$\left\| \hat{H}(\cdot,\cdot) \right\|_{E(\psi)} = \left\| \hat{H}(\cdot,\cdot) \right\|_{C\left([0,\tau_*],L^1(\psi)\right)} = \sup_{0\leq\tau\leq\tau_*} \int_{\mathbb{R}^d} \mathrm{e}^{\psi(|\mathbf{k}|)} \left| \hat{H}(\mathbf{k},\tau) \right| \mathrm{d}\mathbf{k}. \tag{232}$$

Using (231) instead of (18) we obtain as in Lemma 25 the following statement.

Lemma 50. *Operator* $\Psi_\beta f^{(s,\nu)}$ *in* (215) *is bounded for bounded* $\vec{u} \in E_{\mathrm{sc}}^{2N}(\psi)$*,* $f(\mathbf{0}) = \mathbf{0}$*, and satisfies the Lipschitz condition*

$$\left\| \Psi_\beta f^{(s,\nu)}(\vec{u}_1) - \Psi_\beta f^{(s,\nu)}(\vec{u}_2) \right\|_{E_{\mathrm{sc}}^{2N}(\psi)} \leq C\tau_* \left\| \vec{u}_1 - \vec{u}_2 \right\|_{E_{\mathrm{sc}}^{2N}(\psi)}, \tag{233}$$

where C *depends only on* C_χ *as in (88), on the power of polynomial* $f^{(s,\nu)}$ *and on* $\|\vec{u}_1\|_{E_{\mathrm{sc}}^{2N}(\psi)} + \|\vec{u}_1\|_{E_{\mathrm{sc}}^{2N}(\psi)}$ *and does not depend on* $\beta \geq 0$ *and on* $0 < \varrho_1 < \infty$*,* $0 < \varrho_2 \leq \infty$*.*

From Lemma 40 and the contraction principle we obtain the following theorem completely similar to Theorem 33.

Theorem 51. *Let* $\left\| \hat{H} \right\|_{E_{\mathrm{sc}}^{2N}(\psi)} \leq R$*. Then there exists* $R_1 > 0$ *and* $\tau_* > 0$ *such that Eq.* (215) *has a solution* $\vec{u} \in E_{\mathrm{sc}}^{2N}(\psi)$ *which satisfies* $\|\vec{u}\|_{E_{\mathrm{sc}}^{2N}(\psi)} \leq R_1$*, and such a solution is unique.*

The following lemma shows that Ψ can be replaced by one with a small error.

Lemma 52. *Let* $\left\| \hat{H} \right\|_{L^1(\psi)} \leq C$*,* $\psi \in W(a)$*,* Ψ *as in (25). If* $s > 0$*,* $\epsilon > 0$ *and* $\frac{s}{\epsilon} < a$*, then* (204) *holds.*

Proof. We have

$$\int \left(1-\Psi\left(\beta^{\epsilon}\boldsymbol{\eta}\right)\right)\left|\hat{H}\left(\boldsymbol{\eta}\right)\right| \mathrm{d}\boldsymbol{\eta} \leq \int_{|\boldsymbol{\eta}|\geq\beta^{-\epsilon}} \left|\hat{H}\left(\boldsymbol{\eta}\right)\right| \mathrm{d}\boldsymbol{\eta} = \int_{|\boldsymbol{\eta}|\geq\beta^{-\epsilon}} \mathrm{e}^{-\psi(|\boldsymbol{\eta}|)} \left|\mathrm{e}^{\psi(|\boldsymbol{\eta}|)}\hat{H}\left(\boldsymbol{\eta}\right)\right| \mathrm{d}\boldsymbol{\eta} \tag{234}$$

$$\leq \int_{|\boldsymbol{\eta}|\geq\beta^{-\epsilon}} \mathrm{e}^{-\psi\left(\beta^{-\epsilon}\right)} \left|\mathrm{e}^{\psi(|\mathbf{k}|)}\hat{H}\left(\boldsymbol{\eta}\right)\right| \mathrm{d}\boldsymbol{\eta} \leq \beta^{s}\mathrm{e}^{\ln\left(\beta^{-\epsilon}\right)s/\epsilon-\psi\left(\beta^{-\epsilon}\right)} \left\|\hat{H}\right\|_{L^{1}(\psi)}.$$

According to Definition 48 (iii),

$$\ln\left(\beta^{-\epsilon}\right)s/\epsilon-\psi\left(\beta^{-\epsilon}\right)\leq a\ln\left(\beta^{-\epsilon}\right)-\psi\left(\beta^{-\epsilon}\right)\leq C,$$

and we obtain (204) from (234). □

Theorem 53. *Let $\left\|\hat{H}\right\|_{E^{2N}_{\mathrm{sc}}(\psi)}\leq R$, where the weight function ψ belongs to $W(a)$ and let $\frac{s}{\epsilon}<a$. Let $\vec{u}$ and $\vec{u}_0$ be solutions to respectively the minimal equation with cutoff factor and without cutoff factor respectively. Then there exists C_s and β_0 such that*

$$\|\vec{u}-\vec{u}_0\|_{E^{2N}_{\mathrm{sc}}(\psi)}\leq C_s\beta^s,\quad 0<\beta\leq\beta_0. \tag{235}$$

Proof. We show that $\vec{u}$ is an approximate solution to $\vec{u}_0=f^{(\mu,\nu)}\left(\vec{u}_0\right)+\hat{H}$. Namely,

$$\vec{u}=\Psi_\beta f^{(\mu,\nu)}\left(\vec{u}\right)+\Psi_\beta\hat{H}=f^{(\mu,\nu)}\left(\vec{u}\right)+\hat{H}+\hat{D},\quad \hat{D}=\left(\Psi_\beta-1\right)f^{(\mu,\nu)}\left(\vec{u}\right)+\left(\Psi_\beta-1\right)\hat{H}.$$

According to Lemma 49 if $\vec{u}\in E^{2N}_{\mathrm{sc}}\left(\psi\right)$ then $f^{(\mu,\nu)}\left(\vec{u}\right)\in E^{2N}_{\mathrm{sc}}\left(\psi\right)$. Applying Lemma 52 we obtain

$$\left\|\hat{D}\right\|_{E^{2N}_{\mathrm{sc}}(\psi)}\leq C\beta^s,\quad 0<\beta\leq\beta_0. \tag{236}$$

Lemma 27 combined with (236) yields (235). □

Now we give the theorem on approximation by solutions of a minimal system without cutoff.

Theorem 54. *Let $\hat{H}_{l,\zeta}\left(\mathbf{k}\right)$, $l=1,\ldots,N$ be functions bounded in $L^1\left(\psi\right)$, where ψ belongs to $W\left(a\right)$, let $\frac{s}{\epsilon}<a$. Let $\hat{h}_{l,\zeta}\left(\mathbf{k}\right)$ be defined by (202) and $\Psi\hat{\mathbf{h}}_{l,\zeta}\left(\mathbf{k}\right)=\Psi\hat{h}_{l,\zeta}\left(\mathbf{k}\right)\mathbf{g}_{n_l,\zeta}\left(\mathbf{k}\right)$. Let $\hat{\mathbf{u}}\left(\mathbf{k},\tau\right)$ be a solution of Eq. (118) with multiwavepacket initial data of the form (33). Let $u_{l,\vartheta}\left(\mathbf{k},\tau\right)$ be a solution to the system with a weakly universal nonlinearity (216) with initial data $u_{l,\vartheta}\left(\mathbf{k},0\right)=\hat{H}_{l,\vartheta}\left(\mathbf{k}\right)$ and*

$$\hat{\mathbf{u}}_{\min}\left(\mathbf{k},\tau\right)=\sum_{\vartheta}\sum_{l=1}^{N}\beta^{-d}u_{l,\vartheta}\left(\beta^{-1}\left(\mathbf{k}-\zeta\mathbf{k}_{*i_l}\right),\tau\right)\mathbf{g}_{n_l,\vartheta}\left(\mathbf{k}\right).$$

Then

$$\left\|\hat{\mathbf{u}}-\hat{\mathbf{u}}_{\min}\right\|_E\leq C_{\epsilon,s}\beta^s+C\beta^{(\nu+1)(1-\epsilon)}+C\varrho^{-1}\beta^{(\mu+1)(1-\epsilon)}+C\varrho. \tag{237}$$

Proof. We take $\hat{\mathbf{u}}=\sum_\vartheta\sum_{l=1}^{N}u_{l,\vartheta}$ and estimate $\left\|\hat{\mathbf{u}}\left(\mathbf{k},\tau\right)-\hat{\mathbf{u}}_{\min}\left(\mathbf{k},\tau\right)\right\|_E$ applying subsequently Theorems 37, 42, formulas (194) and (205), Theorem 47 and finally Theorem 53 to obtain inequality (237). □

Note that Theorem 7 is a direct corollary of Theorem 54.

Remark 55. Note that (216) is the Fourier integral version of the following system of equations based on weakly universal nonlinearity and is slightly more general than (62),

$$\partial_\tau u_{l,\vartheta} = \frac{1}{\varrho_1}\omega'_{n_l}(\mathbf{k}_{*l}) \cdot \nabla_x u_{l,\vartheta} + \frac{i}{2\varrho_2}\nabla_r \cdot \omega''_{n_l}(\mathbf{k}_{*l})\nabla_r u_{l,\vartheta} + f^{(\mu,\nu)}_{n_l,\vartheta}(\vec{u}, \delta\vec{u}), \quad (238)$$
$$u_{l\vartheta}|_{\tau=0} = \hat{H}_{l\vartheta}, \text{ where } \delta_i u_l(\mathbf{r}) = u_j(\mathbf{r}+e_i) - u_j(\mathbf{r}-e_i),$$

where ϱ_1, ϱ_1 are as in (222) and e_i is *ith* standard ort in $\mathbb{R}^d$. In the case when (52) holds $1/\varrho_2$ is bounded or small and the dependence on the coefficient $1/\varrho_2$ is regular for small ϱ and β and $u_{\vartheta,j}(\mathbf{k},\tau)$ may be looked at as a shape function. When $\varrho_1 = \varrho$ and $1/\varrho_2$ is substituted by zero we obtain an equation exactly of the form (62).

When $\nu = 0$, $\mu = 1$ and the nk-spectrum S is universally resonance invariant as in Definition 18, the nonlinearities $f^{(1,0)}_{n_l,\vartheta,0}$ are universal of the form (65). When the nk-spectrum S is resonance invariant but not universally resonance invariant, the nonlinearities are weakly universal, but may be not universal, that allows, in particular, for the second and the third harmonic generation.

Acknowledgement. Effort of A. Babin and A. Figotin is sponsored by the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under grant number FA9550-04-1-0359. We also would like to express our deep gratitude to the reviewer for the thorough analysis of our work and valuable suggestions which helped to improve the presentation of our results.

References

1. Babin, A., Figotin, A.: Nonlinear Photonic Crystals: I. Quadratic nonlinearity. Waves in Random Media **11**, R31–R102 (2001)
2. Babin, A., Figotin, A.: Nonlinear Photonic Crystals: II. Interaction classification for quadratic nonlinearities. Waves in Random Media **12**, R25–R52 (2002)
3. Babin, A., Figotin, A.: Nonlinear Photonic Crystals: III. Cubic Nonlinearity. Waves in Random Media **13**, R41–R69 (2003)
4. Babin, A., Figotin, A.: Nonlinear Maxwell Equations in Inhomogenious Media. Commun. Math. Phys. **241**, 519–581 (2003)
5. Babin, A., Figotin, A.: Polylinear spectral decomposition for nonlinear Maxwell equations. In: Agranovich, M.S., Shubin, M.A. (eds.) *Partial Differential Equations, Advances in Mathematical Sciences*, American Mathematical Society Translations-Series 2, Vol. **206**, Providence, RI: Amer. Math. Soc., 2002, pp. 1–28
6. Babin, A., Figotin, A.: Nonlinear Photonic Crystals: IV Nonlinear Schrodinger Equation Regime. Waves in Random and Complex Media, **15**(2), 145–228 (2005)
7. Babin, A., Figotin, A.: Linear Superposition In Nonlinear Wave Dynamics. Rev. Math. Phys. **18**(9), 971–1053 (2006)
8. Babin, A., Mahalov, A., Nicolaenko, B.: Global regularity of 3D rotating Navier-Stokes equations for resonant domains. Indiana Univ. Math. J. **48**(3), 1133–1176 (1999)
9. Babin, A., Mahalov, A., Nicolaenko, B.: Fast Singular Oscillating Limits and Global Regularity for the 3D Primitive Equations of Geophysics. M2AN **34**(2), 201–222 (2000)
10. Ben Youssef, W., Lannes, D.: The long wave limit for a general class of 2D quasilinear hyperbolic problems. Comm. Par. Differ. Eqs. **27**(5–6), 979–1020 (2002)
11. Bogoliubov, N.N., Mitropolsky, Y.A.: *Asymptotic Methods In The Theory Of Non-Linear Oscillations*. Delhi: Hindustan Pub. Corp., 1961
12. Boyd, R.: *Nonlinear Optics*. London:Academic Press, 1992
13. Bona, J.L., Colin, T., Lannes, D.: Long wave approximations for water waves. Arch. Rat. Mech. Anal. **178**(3), 373–410 (2005)
14. Bourgain, J.: *Global solutions of nonlinear Schrödinger equations. American Mathematical Society Colloquium Publications* **46**. Providence, RI: Amer. Math. Soc., 1999

15. Butcher, P., Cotter, D.: *The Elements of Nonlinear Optics*. Cambridge: Cambridge Univ. Press, 1993
16. Cazenave, T.: *Semilinear Schrödinger equations*. Courant Lecture Notes in Mathematics 10, New York:New York University, Courant Institute of Mathematical Sciences, Providence, RI: Amer. Math. Soc. 2003
17. Colin, T.: Rigorous derivation of the nonlinear Schrödinger equation and Davey-Stewartson systems from quadratic hyperbolic systems. Asymptot. Anal. **31**(1), 69–91 (2002)
18. Colin, T., Lannes, D.: Justification of and long-wave correction to Davey-Stewartson systems from quadratic hyperbolic systems. Discrete Contin. Dyn. Syst. **11**(1), 83–100 (2004)
19. Craig, W., Groves, M.D.: Normal forms for wave motion in fluid interfaces. Wave Motion **31**(1), 21–41 (2000)
20. Craig, W., Sulem, C., Sulem, P.-L.: Nonlinear modulation of gravity waves: a rigorous approach. Nonlinearity **5**(2), 497–522 (1992)
21. Dobrokhotov, S.Yu., Maslov, V.P., Omelyanov, G.A.: Multiwave interaction in weakly nonlinear media with dispersion. In: *Mathematical mechanisms of turbulence*, i, Kiev: Akad. Nauk Ukrain. SSR, Inst. Mat., 1986, pp. 25–45
22. Dineen, S.: *Complex Analysis on Infinite Dimensional Spaces*. Berlin-Heidelberg-New york: Springer, 1999
23. Giannoulis, J., Mielke, A.: The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities. Nonlinearity **17**(2), 551–565 (2004)
24. Goodman, R.H., Weinstein, M.I., Holmes, P.J.: Nonlinear propagation of light in one-dimensional periodic structures. J. Nonlinear Sci. **11**(2), 123–168 (2001)
25. Groves, M.D., Schneider, G.: Modulating pulse solutions for quasilinear wave equations. J. Differ. Eq. **219**(1), 221–258 (2005)
26. Hayashi, N., Naumkin, P.: Asymptotics of small solutions to nonlinear Schrödinger equations with cubic nonlinearities. Int. J. Pure Appl. Math. **3**(3), 255–273 (2002)
27. Hille, E., Phillips, R.S.: *Functional Analysis and Semigroups*. Providence RI:AMS, 1991
28. Infeld, E., Rowlands, G.: *Nonlinear Waves, Solitons, and Chaos*. 2nd ed., Cambridge: Cambridge University Press, 2000
29. Joly, J.-L., Metivier, G., Rauch, J.: Diffractive nonlinear geometric optics with rectification. Indiana Univ. Math. J. **47**(4), 1167–1241 (1998)
30. Kalyakin, L.A.: Long-wave asymptotics. Integrable equations as the asymptotic limit of nonlinear systems. Usp. Mat. Nauk **44**(1)(265), 5–34, 247 (1989); translation in Russ. Math. Surv. **44**(1), 3–42 (1989)
31. Kalyakin, L.A.: Asymptotic decay of a one-dimensional wave packet in a nonlinear dispersive medium. Math. USSR Sb. **60**(2), 457–483 (1988)
32. Krieger, J., Schlag, W.: Stable manifolds for all monic supercritical focusing nonlinear Schrödinger equations in one dimension. J. Amer. Math. Soc. (Electronic) **19**(4), 815–920 (2006)
33. Kuksin, S.B.: Fifteen years of KAM for PDE. Geometry, topology, and mathematical physics, Amer. Math. Soc. Transl. Ser. 2, **212**, Providence, RI: Amer. Math. Soc., 2004, pp. 237–258
34. Kirrmann, P., Schneider, G., Mielke, A.: The validity of modulation equations for extended systems with cubic nonlinearities. Proc. Roy. Soc. Edinburgh Sect. A **122**(1–2), 85–91 (1992)
35. Kato, T.: *Perturbation Theory for Linear Operators*. Berlin-Heidelberg-New York: Springer, 1980
36. Lax, P.D.: Integrals of nonlinear equations of evolution and solitary waves. Comm. Pure Appl. Math. **21**, 467–490 (1968)
37. Mitropolskii, Yu.A., Nguyen, V.D.: *Applied asymptotic methods in nonlinear oscillations. Solid Mechanics and its Applications* **55**. Dordrecht: Kluwer Academic Publishers Group, 1997
38. Maslov. V.P.: Non-standard characteristics in asymptotic problems. Usp. Mat. Nauk **38**:6, 3–36 (1983), translation in Russ. Math. Surv. **38**:6, 1–42 (1983)
39. Maslov, V.P.: Mathematical aspects of integral optics. Russ. J. Math. Phys. **8**(1), 83–105 (2001)
40. Mielke, A., Schneider, G., Ziegra, A.: Comparison of inertial manifolds and application to modulated systems. Math. Nachr. **214**, 53–69 (2000)
41. Moloney, J., Newell, A.: *Nonlinear Optics*. Advanced Book Program, Boulder, CO: Westview Press, 2004
42. Mills, D.: *Nonlinear Optics*. Berlin-Heidelberg-New York: Springer-Verlag, 1991
43. Nayfeh, A.H.: *Perturbation Methods*. New York: Wiley, 1973
44. Ostrovsky, L., Potapov, A.: *Modulated Waves*. Baltimore MD: The John Hopkins Univ. Press, 1999
45. Pankov, A.: *Travelling Waves And Periodic Oscillations In Fermi-Pasta-Ulam Lattices*. London: Imperial College Press, 2005
46. Phillips, O.M.: Wave Interactions. In: Leibovich, S., Seebass, A.R. (eds.) *Nonlinear Waves*. Ithaca and London: Cornell Univ. Press, 1974
47. Pierce, R.D., Wayne, C.E.: On the validity of mean-field amplitude equations for counterpropagating wavetrains. Nonlinearity **8**(5), 769–779 (1995)
48. Sauter, E.G.: *Nonlinear Optics*. New york: Wiley-Interscience, 1996

49. Schlag, W.: Spectral theory and nonlinear partial differential equations: a survey. Discrete Contin. Dyn. Syst. **15**(3), 703–723 (2006)
50. Schneider, G.: Justification of modulation equations for hyperbolic systems via normal forms. NoDEA Nonlinear Differential Equations Appl. **5**(1), 69–82 (1998)
51. Schneider, G.: Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. J. Differ. Eq. **216**(2), 354–386 (2005)
52. Schneider, G., Uecker, H.: Nonlinear coupled mode dynamics in hyperbolic and parabolic periodically structured spatially extended systems. Asymptot. Anal. **28**(2), 163–180 (2001)
53. Schneider, G., Uecker, H.: Existence and stability of modulating pulse solutions in Maxwell's equations describing nonlinear optics. Z. Angew. Math. Phys. **54**(4), 677–712 (2003)
54. Schneider, G., Wayne, C.E.: Estimates for the three-wave interaction of surface water waves. European J. Appl. Math. **14**(5), 547–570 (2003)
55. Sipe, J.E., Bhat, N., Chak, P., Pereira, S.: Effective field theory for the nonlinear optical properties of photonic crystals. Phys. Rev. E **69**, 016604 (2004)
56. Slusher, R.E., Eggleton, B.J.: *Nonlinear Photonic Crystals*. Berlin-Heidelberg-New York: Springer-Verlag, 2003
57. Sulem, C., Sulem, P.-L.: *The Nonlinear Schrodinger Equation*. Berlin-Heidelberg-New York: Springer, 1999
58. Volkov, S.N., Sipe, J.E.: Nonlinear optical interactions of wave packets in photonic crystals: Hamiltonian dynamics of effective fields. Phys. Rev. E **70**, 066621 (2004)
59. Soffer, A., Weinstein, M.I.: Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations. Invent. Math. **136**(1), 9–74 (1999)
60. Weissert, T.P.: *The Genesis of Simulation in Dynamics: pursuing the Fermi-Pasta-Ulam problem*. New York: Springer-Verlag, 1997
61. Whitham, G.: *Linear and Nonlinear Waves*. New York: John Wiley & Sons, 1974

Communicated by P. Constantin