HAL open science

Short Pulses Approximations in Dispersive Media

Mathieu Colin, David Lannes

To cite this version:

Mathieu Colin, David Lannes. Short Pulses Approximations in Dispersive Media. 2007. hal-00201084

HAL Id: hal-00201084

https://hal.science/hal-00201084

Preprint submitted on 23 Dec 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SHORT PULSES APPROXIMATIONS IN DISPERSIVE MEDIA

MATHIEU COLIN AND DAVID LANNES

Abstract. We derive various approximations for the solutions of nonlinear hyperbolic systems with fastly oscillating initial data. We first provide error estimates for the so-called slowly varying envelope, full dispersion, and Schrödinger approximations in a Wiener algebra; this functional framework allows us to give precise conditions on the validity of these models; we give in particular a rigorous proof of the "practical rule" which serves as a criterion for the use of the slowly varying envelope approximation (SVEA). We also discuss the extension of these models to short pulses and more generally to large spectrum waves, such as chirped pulses. We then derive and justify rigorously a modified Schrödinger equation with improved frequency dispersion. Numerical computations are then presented, which confirm the theoretical predictions.

1. Introduction

1.1. **General setting.** The nonlinear Schrödinger equation has been derived as an asymptotic model for many physical problems involving the propagation of slowly modulated oscillating plane waves. Typical examples are water-waves [16, 24] or ferromagnetism [20] and of course nonlinear optics (see e.g. [10, 21]) for which it plays a central role. It is aimed at approximating the solution u of a nonlinear hyperbolic system with fast oscillating initial condition, say,

$$\begin{cases} \partial_t u + A(\partial)u + \dfrac{1}{\varepsilon}Eu = \varepsilon F(u), \\ u_{|_{t=0}} = U^0(x)e^{i\frac{\mathbf{k}\cdot x}{\varepsilon}} + \text{c.c.}; \end{cases} \tag{1}$$

here $\varepsilon \ll 1$ is a small parameter corresponding to the wavelength of the oscillations in dimensionless variables (see Assumption 1 below for more precisions). The so-called Schrödinger approximation can be decomposed into two steps:

-a- *Slowly Varying Envelope Approximation (SVEA)*: one writes the solution u as the product of a fast oscillating wave train and an envelope U:

$$u(t,x) \sim U(t,x)e^{i\frac{\mathbf{k}\cdot x - \omega(\mathbf{k})t}{\varepsilon}} + \text{c.c.}, \tag{2}$$

where $(\omega(\mathbf{k}), \mathbf{k})$ solves the *dispersion relation* (or, equivalently, belongs to the *characteristic variety*, see (13) below).

-b- The envelope $U(t,x)$ is approximated by the solution of the nonlinear Schrödinger equation

$$\partial_t U + (\mathbf{c_g}\cdot\nabla)U - \varepsilon\frac{i}{2}\mathcal{R}(\partial,\partial)U = \varepsilon\widetilde{F}(U), \qquad U_{|_{t=0}} = U^0, \tag{3}$$

where the group velocity $\mathbf{c_g}$, the second order differential operator $\mathcal{R}(\partial,\partial)$ and the nonlinearity $\widetilde{F}$ can be explicitly given in terms of the data.

This work was supported by the ACI Jeunes Chercheurs et Jeunes Chercheuses "Dispersion et nonlinéarités".

The Schrödinger approximation (2)-(3) has been rigorously justified [14, 15, 17] for times of order $O(1/\varepsilon)$ in the usual situation where the typical scale for the space variations of U^0 is of order 1. This means that the number of oscillations (or optical cycles) in the laser pulse is of order $O(1/\varepsilon)$. Recently, however, lasers with ultrashort pulses have been developed, for which the number of optical cycles is much smaller; for such pulses, the Schrödinger equation proves completely inaccurate and various authors proposed other ways of describing the asymptotics of (1) when $\varepsilon \to 0$. Some of these results are briefly recalled below; their common point is that they all abandon the *SVEA*, because the widely accepted "practical rule"

$$\text{The SVEA (2) is valid if } |\nabla U^0|_\infty \ll \frac{1}{\varepsilon} \tag{4}$$

is enforced when the pulses get very small.
Alterman and Rauch [1, 2, 3] modeled short pulses by replacing the fast oscillating term in the initial condition by a fast decaying one; more precisely, they modified the initial condition for (1) as follows:

$$u_{|t=0} = U^0(x)e^{i\frac{\mathbf{k}\cdot x}{\varepsilon}} + \text{c.c.} \quad \rightsquigarrow \quad u_{|t=0} = U^0(x, \frac{\mathbf{k}\cdot x}{\varepsilon}), \tag{5}$$

with $U^0(x,z) \to 0$ as $z \to \infty$, and the SVEA (2) is consequently replaced by

$$u(t,x) \sim U\big(t, x, \frac{\mathbf{k}\cdot x - \omega(\mathbf{k})t}{\varepsilon}\big),$$

with $U(t,x,z) \to 0$ as $z \to \infty$. The Schrödinger equation (3) is then replaced by

$$\partial_t\partial_z U + (\mathbf{c_g}\cdot\nabla)\partial_z U + \varepsilon\frac{1}{2}\mathcal{R}(\partial,\partial)U = \varepsilon\partial_z\widetilde{F}(U), \qquad U_{|t=0} = U^0; \tag{6}$$

this approximation (rigorously justified) uses the fact that the group velocity $\mathbf{c_g}$ does not depend on $|\mathbf{k}|$ and is therefore only valid in *nondispersive media* ($E = 0$ in (1)). Alterman and Rauch's approach has been generalized in [13, 23] taking into account the particularities of the optical susceptibility of some cubic nonlinear media such as silica, and finally obtaining a *quasilinear* variant of (6), which is rigorously justified in the linear case.
In order to model the propagation of ultrashort pulses in *dispersive* media, Barrailh and Lannes [6] chose another approach based on the functional tools developed in [18], which consist in replacing the initial condition for (1) as follows

$$u_{|t=0} = U^0(x)e^{i\frac{\mathbf{k}\cdot x}{\varepsilon}} + \text{c.c.} \quad \rightsquigarrow \quad u_{|t=0} = U^0(x, 0, \frac{\mathbf{k}\cdot x}{\varepsilon}),$$

where the Fourier transform of the initial profile $U^0(x,T,Z)$ with respect to T and Z is an *H^s-valued measure of bounded variation*. This general framework allows one to consider initial data of the form (1) –for which the bounded variation measure is obviously $U^0(x)\delta_{(\omega(\mathbf{k}),\mathbf{k})}$– and of the form (5). A generalization of (6) is then derived and rigorously justified; this equation however has the drawback of being quite complicated because the transport operator $\partial_t + \mathbf{c_g}\cdot\nabla$ must be replaced by a nonlocal operator modeling the fact that the group velocity $\mathbf{c_g}$ depends on the frequency in dispersive media.

An important characteristic of ultrashort pulses that we did not mention so far is that their frequency spectrum is broad (while for usual wave packets as the initial condition of (1), and taking the Fourier transform with respect to the fast scale x/ε, it is essentially contained in a $O(\varepsilon)$ neighborhood of $\mathbf{k}$; see for instance [4]). Since

the dispersion relation of the Schrödinger equation (3) is a second order Taylor expansion of the exact dispersion relation of (1) at $\mathbf{k}$, the error is quite important if frequencies far from $\mathbf{k}$ must be taken into account; in addition to the violation of the practical rule (4), this is another reason why the Schrödinger approximation breaks down for short pulses. This phenomenon is not specific to short pulses since it occurs for all pulses with large frequency spectrum (typical examples are *chirped pulses*).

An alternative way to replace the NLS approximation for such pulses is therefore to focus on the dispersive properties of the asymptotic model. Instead of abandoning the SVEA (2) as in [1, 2, 3, 23, 13, 6], various authors [11, 12] chose to make this approximation but kept the full dispersive properties of the original equations (1), thus avoiding the disastrous (for large spectrum pulses in dispersive media) second order Taylor expansion of the dispersion relation. Consequently, the approximations thus obtained might have a slightly smaller range of validity, but are undoubtly simpler and moreover provide a more precise approximation ($O(\varepsilon)$ versus $o(1)$). For instance, the equation derived in [12], and which we call *full dispersion model* here, reads

$$\partial_t U + \frac{i}{\varepsilon}(\omega(\mathbf{k}+\varepsilon D) - \underline{\omega})U = \varepsilon \widetilde{F}(U), \qquad U(x) = U^0(x),$$

where $\omega(\cdot)$ parameterizes the graph of the relevant sheet of the characteristic variety (quite obviously (3) can be deduced from this equation by Taylor expanding $\omega(\cdot)$ at $\mathbf{k}$). This model is much simpler than the one derived in [6] and furnishes very satisfactory results; however it is still a nonlocal equation and its resolution requires spectral methods. One of the goals of this paper is to derive a new approximation, with, in practical, the same dispersive qualities as the full dispersion model, but keeping the same level of complexity as the usual Schrödinger equation (3).

Before describing with more details the results of this paper, let us introduce here two kinds of initial data for (1), which we will often refer to throughout the article:

- Short-pulses: the initial profile $U^0(x)$ in (1) is taken of the form

$$U^0(x) = f(\frac{x-x_0}{\beta}), \tag{7}$$

  with $0 < \beta \leq 1$ and f a smooth function; the case $\beta = 1$ corresponds to classical laser pulses, and short pulses to $\beta \ll 1$ (see Fig. 1); the number of optical cycles for such a short pulse is thus $O(\varepsilon/\beta)$;
- Chirped pulses: the initial profile $U^0(x)$ in (1) is taken of the form

$$U^0(x) = f(x-x_0)\cos(\frac{1}{\beta}\cos(\frac{x-x_0}{\beta})), \tag{8}$$

  with $0 < \beta \leq 1$ and f a smooth function; the case $\beta = 1$ corresponds to classical laser pulses, and chirped pulses to $\beta \ll 1$ (see Fig. 2).

More precisely, we propose here to:

(1) Provide a framework simpler than [6] but general enough to handle any kind of large spectrum pulses (such as ultrashort and chirped pulses). The idea here is to keep the SVEA and to work with envelopes which are in a Wiener algebra. The reason of this choice is that when $\beta \to 0$, the Wiener norm

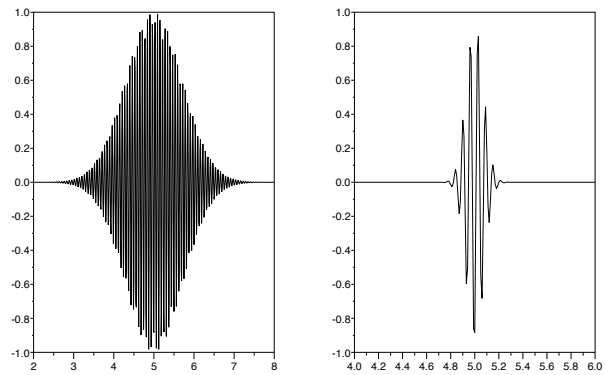

FIGURE 1. Initial condition $U^0(x)e^{i\frac{\mathbf{k}\cdot x}{\varepsilon}}$ +c.c. with U^0 as in (7) and $f = e^{-x^2}$, $x_0 = 5$, $\varepsilon = 0.01$, and for $\beta = 1$ and $\beta = 0.1$

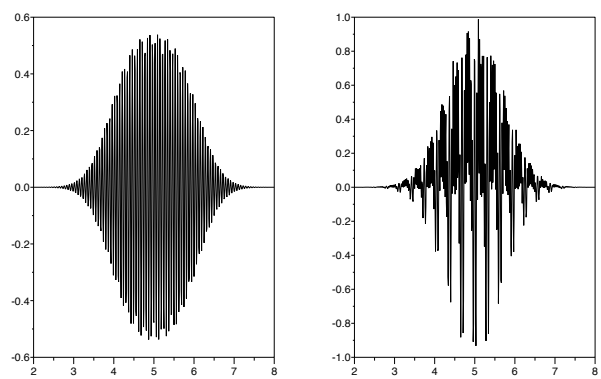

FIGURE 2. Initial condition $U^0(x)e^{i\frac{\mathbf{k}\cdot x}{\varepsilon}}$ +c.c. with U^0 as in (8) and $f = e^{-x^2}$, $x_0 = 5$, $\varepsilon = 0.01$, and for $\beta = 1$ and $\beta = 0.1$

(which controls the L^∞-norm) of the initial envelopes (7) remain bounded while any Sobolev norm controlling the L^∞-norm grows to infinity;

(2) Rigorously prove the "practical rule" (4). We show that the SVEA (2) makes sense if $|\nabla U^0|_W \ll \frac{1}{\varepsilon}$, where $|\cdot|_W$ is the Wiener norm (see (16) below); in the case of short-pulses (7), this condition is equivalent to $\varepsilon \ll \beta$;

(3) Establish precise error estimates for the full dispersion and Schrödinger models. This allows us to give precise estimates on the range of validity of these models;

(4) Derive and rigorously justify a new family of Schrödinger equations with improved frequency dispersion and whose formulation is purely differential

(without nonlocal operator). The idea is to approximate the nonlocal operator of the full dispersion model by a suitable rational function, following an idea which proved very useful in water-waves theory (derivation of the BBM equation from the KdV equation [7], derivation of Boussinesq models with improved frequency dispersion [8, 9]). These equations read

$$\begin{aligned}&\big(1-i\varepsilon\mathbf{b}\cdot\nabla-\varepsilon^2\nabla\cdot B\nabla\big)\partial_t U+(\mathbf{c_g}\cdot\nabla)U\\&\quad-\varepsilon\frac{i}{2}\mathcal{R}(\partial,\partial)U+i\varepsilon\nabla\cdot(\nabla\omega_1(\underline{\mathbf{k}})\mathbf{b}^T)\nabla U+\varepsilon^2\mathbf{C}(\nabla)U=\widetilde{F}(U),\end{aligned}\tag{9}$$

where $\mathbf{b}\in\mathbb{C}^d$, $B\in\mathcal{M}_{d\times d}(\mathbb{R})$ and $\mathbf{C}:\mathbb{C}^d\times\mathbb{C}^d\times\mathbb{C}^d\to\mathbb{C}$ is a trilinear mapping (taking $\mathbf{b}=0$, $B=0$ and $\mathbf{C}=0$ gives therefore (3));

(5) Present some numerical comparisons between the different asymptotic models, for ultrashort and chirped pulses.

1.2. **Organization of the paper.** Section 2 is devoted to the derivation and justification of the different asymptotic models. Basic assumptions, definitions, and tools are first given in §2.1 and we then proceed to prove the envelope approximation (2) in §2.2; in particular, precise error estimates are established which allow one to give a rigorous proof of the practical rule (4). In §2.3, we then derive and give precise error estimates in a Wiener algebra setting of the full dispersion model (§2.3.1) and the Schrödinger model (§2.3.2). Finally the new Schrödinger equations (9) are derived in §2.3.3.

In Section 3, we present numerical computations in order to make a comparison of the different asymptotic models.

1.3. **Notations.** - We denote generically by Cst a constant whose value may change from one line to another;
- We use the abbreviation c.c. for "complex conjugate", so that for all $a\in\mathbb{C}$, $a+\text{c.c.}=a+\overline{a}$;
- If $\mathbf{a},\mathbf{b}\in\mathbb{C}^d$ then $\mathbf{a}\cdot\mathbf{b}$ denotes the scalar (not hermitian) product: $\mathbf{a}\cdot\mathbf{b}=\sum_{j=1}^d\mathbf{a}_j\mathbf{b}_j$.
- For all $x\in\mathbb{R}^d$, we write $x=(x_1,\dots,x_d)$, and ∂_j stands for $\partial_j=\partial_{x_j}$; we also write $\nabla=(\partial_1,\dots,\partial_d)^T$ and, for all multi-index $\alpha\in\mathbb{N}^d$, $\partial^\alpha=\partial_1^{\alpha_1}\dots\partial_d^{\alpha_d}$;
- For all $j=1,\dots,d$, we write $D_j=\frac{1}{i}\partial_j$ and $D=(D_1,\dots,D_d)=\frac{1}{i}\nabla^T$;
- The Fourier transform of a distribution $u\in\mathfrak{S}(\mathbb{R}^d)$ is denoted either by $\mathcal{F}u$ or $\widehat{u}$;
- We use the classical notation for Fourier multipliers:

$$f(D)u=\mathcal{F}^{-1}\big(\xi\mapsto f(\xi)\widehat{u}(\xi)\big);$$

- If $f\in L^\infty(\mathbb{R}^d)$, we simply write $|f|_\infty=|f|_{L^\infty}$.

2. Asymptotic results

This section is devoted to the study of the asymptotic behavior when ε goes to zero of the solution to the initial value problem (ivp)

$$\begin{cases}\partial_t u+A(\partial)u+\dfrac{E}{\varepsilon}u=\varepsilon T(u,\overline{u},u),\\ u_{|_{t=0}}=U^0(x)e^{i\frac{\mathbf{k}\cdot x}{\varepsilon}}+\text{c.c.},\end{cases}\tag{10}$$

where $A(\partial)$, E and $T(\cdot,\cdot,\cdot)$ are defined in the next section.

2.1. **Basic assumptions and tools.** We make the following assumption on the operators $A(\partial)$, E and $T(\cdot,\cdot,\cdot)$ which appear in (10):

Assumption 1. i. *The system (10) is symmetric hyperbolic in the sense that for some $n \geq 1$:*

- *One has $A(\partial) = \sum_{j=1}^d A_j \partial_j$, and the A_j are $n \times n$ symmetric, real-valued, matrices;*
- *The $n \times n$ matrix E is real and skew-symmetric;*

ii. *The mapping*

$$T = \begin{array}{ccc} \mathbb{C}^{3n} & \to & \mathbb{C}^n \\ (u^1, u^2, u^3) & \mapsto & T(u^1, u^2, u^3), \end{array}$$

is linear with respect to u^1, u^2 and u^3.

Example 1. *A standard model for the propagation of a beam in a Kerr medium is the Maxwell-Lorentz system which can be written in dimensionless form as*

$$\left\{ \begin{array}{l} \partial_t B + \text{curl } E = 0, \\ \partial_t E - \text{curl } B + \frac{1}{\varepsilon} Q = 0, \\ \partial_t Q - \frac{1}{\varepsilon}(E - P) = |P|^2 P, \\ \partial_t P - \frac{1}{\varepsilon} Q = 0, \end{array} \right. \tag{11}$$

where (E, B) is the electromagnetic field, P the polarization and $Q = \varepsilon \partial_t P$. This system is of the form (10) and satisfies Assumption 1 with $n = 4d$ and

$$A(\partial) = \begin{pmatrix} 0 & \nabla\times & 0 & 0 \\ -\nabla\times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & I_{d\times d} & 0 \\ 0 & -I_{d\times d} & 0 & I_{d\times d} \\ 0 & 0 & -I_{d\times d} & 0 \end{pmatrix}$$

(the entries in the above matrices are $d\times d$ matrices); denoting $u^j = (B^j, E^j, Q^j, P^j)$ $(j = 1, 2, 3)$, the nonlinearity is given by

$$T(u^1, u^2, u^3) = \begin{pmatrix} 0 \\ 0 \\ (P^1 \cdot P^2) P^3 \\ 0 \end{pmatrix}.$$

Example 2. *A simple toy model is the following Klein-Gordon system*

$$\partial_t u + \begin{pmatrix} 0 & \nabla^T \\ \nabla & 0 \end{pmatrix} u + \frac{1}{\varepsilon} \begin{pmatrix} 0 & -\mathbf{v} \\ \mathbf{v}^T & 0 \end{pmatrix} u = |u|^2 \begin{pmatrix} 0 & -\mathbf{v} \\ \mathbf{v}^T & 0 \end{pmatrix} u, \tag{12}$$

with $u : \mathbb{R}_t^+ \times \mathbb{R}_x^d \mapsto \mathbb{C}^{1+d}$ and $\mathbf{v} \in \mathbb{R}^d \backslash \{0\}$.
Quite obviously, (12) is of the form (10) and satisfies Assumption 1 with $n = 1+d$,

$$A(\partial) = \begin{pmatrix} 0 & \nabla^T \\ \nabla & 0 \end{pmatrix}, \qquad E = \begin{pmatrix} 0 & -\mathbf{v}^T \\ \mathbf{v} & 0 \end{pmatrix}$$

and $T(u^1, u^2, u^3) = (u^1 \cdot u^2) \begin{pmatrix} 0 & -\mathbf{v}^T \\ \mathbf{v} & 0 \end{pmatrix} u^3$.

Under Assumption 1, the matrix $A(\mathbf{k}) + \frac{E}{i} = \sum_{j=1}^d A_j \mathbf{k}_j + \frac{E}{i}$ is hermitian for all $\mathbf{k} \in \mathbb{R}^d$, and thus diagonalizable, with real eigenvalues. We can therefore define the *characteristic variety* as

$$\mathcal{C} := \{(\omega, \mathbf{k}) \in \mathbb{R} \times \mathbb{R}^d, \omega \text{ is an eigenvalue of } A(\mathbf{k}) + \frac{E}{i}\}; \tag{13}$$

introducing

$$\mathcal{L}(\omega, \underline{\mathbf{k}}) := -\omega I + A(\mathbf{k}) + \frac{E}{i}, \tag{14}$$

one can equivalently define $\mathcal{C}$ as the set of all $(\omega, \mathbf{k}) \in \mathbb{R} \times \mathbb{R}^d$ such that $\mathcal{L}(\omega, \mathbf{k})$ is *not* invertible.

It is classical and not restrictive for our present concern to make the following assumption on $\mathcal{C}$:

Assumption 2. *There exists $m \in \mathbb{N}$ and m different smooth functions $\omega_j \in C^\infty(\mathbb{R}^d \backslash \{0\})$ $(j = 1, \dots, m)$, such that for all $\mathbf{k} \in \mathbb{R}^d \backslash \{0\}$, the eigenvalues of $A(\mathbf{k}) + \frac{E}{i}$ are exactly $\omega_j(\mathbf{k})$ $(j = 1, \dots, m)$.*

Under this Assumption 2, one can write, for all $\mathbf{k} \in \mathbb{R}^d \backslash \{0\}$,

$$A(\mathbf{k}) + \frac{E}{i} = \sum_{j=1}^{m} \omega_j(\mathbf{k}) \pi_j(\mathbf{k}), \tag{15}$$

where $\pi_j(\mathbf{k})$ denote the eigenprojector associated to $\omega_j(\mathbf{k})$ (in particular, $\pi_j \in C^\infty(\mathbb{R}^d \backslash \{0\}; \mathcal{M}_n(\mathbb{C}))$.

We finally need a last assumption on the wave number $\underline{\mathbf{k}}$ of the initial data of the ivp (10) in order to justify the asymptotic equations derived in this article.

Assumption 3. *One has $\underline{\mathbf{k}} \neq 0$ and, with $\underline{\omega} = \omega_1(\underline{\mathbf{k}})$:*

- *One has $(3\underline{\omega}, 3\underline{\mathbf{k}}) \notin \mathcal{C}$;*
- *With the notations of Assumption 2,*

$$\exists c_0 > 0, \quad \forall j = 2, \dots, m, \qquad \inf_{\mathbf{k} \in \mathbb{R}^d \backslash \{0\}} \left| \underline{\omega} - \omega_j(\mathbf{k}) \right| \geq c_0.$$

Remark 1. *The first part of the assumption excludes resonances with the third harmonic. The results presented here could easily be extended to cover such a situation, but it is not restrictive at all to make this assumption.*

Example 3. *For the Maxwell equations (11), one can check after some computations that Assumption 2 is satisfied with $m = 7$ and*

$$\begin{aligned}
\omega_1(\mathbf{k}) &= \frac{1}{2}\big(\sqrt{2(1+|\mathbf{k}|)+|\mathbf{k}|^2} + \sqrt{2(1-|\mathbf{k}|)+|\mathbf{k}|^2}\big), \\
\omega_2(\mathbf{k}) &= \sqrt{2}, \\
\omega_3(\mathbf{k}) &= \frac{1}{2}\big(\sqrt{2(1+|\mathbf{k}|)+|\mathbf{k}|^2} - \sqrt{2(1-|\mathbf{k}|)+|\mathbf{k}|^2}\big), \\
\omega_4(\mathbf{k}) &= 0,
\end{aligned}$$

and $\omega_5 = -\omega_3$, $\omega_6 = -\omega_2$, $\omega_7 = -\omega_1$ (so that $\omega_1 > \omega_2 > \cdots > \omega_7$ on $\mathbb{R}^d \backslash \{0\}$); one can therefore take $c_0 = \omega_1(\underline{\mathbf{k}}) - \sqrt{2} > 0$.

Example 4. *For the Klein-Gordon system (12) one readily checks that Assumption 2 is satisfied with $m = 2$, $\omega_1(\mathbf{k}) = \sqrt{|\mathbf{k}|^2 + |\mathbf{v}|^2}$ and $\omega_2 = -\omega_1$. One can then remark that Assumption 3 also holds for all $\underline{\mathbf{k}} \neq 0$ and $c_0 = \omega_1(\underline{\mathbf{k}}) + |\mathbf{v}|$.*

We finally end this section with some results on the Wiener algebras. First recall that $W^k(\mathbb{R}^d;\mathbb{C}^n)$ $(k,n\in\mathbb{N})$ and $W(\mathbb{R}^d_X\times\mathbb{T}_\theta;\mathbb{C}^n)$ (which will be denoted by $W(\mathbb{R}^d\times\mathbb{T};\mathbb{C}^n)$ in the sequel) are defined as

$$W^k(\mathbb{R}^d;\mathbb{C}^n):=\{f\in\mathcal{S}'(\mathbb{R}^d)^n,\forall\alpha\in\mathbb{N}^d,|\alpha|\le k,|\partial^\alpha f|_W<\infty\},\tag{16}$$

with $|f|_W:=|\widehat{f}|_{L^1(\mathbb{R}^d;\mathbb{C}^n)}$, and

$$W(\mathbb{R}^d\times\mathbb{T};\mathbb{C}^n):=\{f=\sum_{n\in\mathbb{Z}}f_n(X)e^{in\theta},|f|_{W(\mathbb{R}^d\times\mathbb{T})}:=\sum_n|f_n|_W<\infty\}$$

(when $k=0$, we write $W(\mathbb{R}^d;\mathbb{C}^n)$ instead of $W^0(\mathbb{R}^d;\mathbb{C}^n)$).
The classical properties of the Wiener algebras used in this article are recalled in the following proposition.

Proposition 1. i. *The space $W^k(\mathbb{R}^d;\mathbb{C})$, $k\in\mathbb{N}$, (resp. $W(\mathbb{R}^d\times\mathbb{T};\mathbb{C})$) is an algebra in the sense that the mapping $(f,g)\mapsto fg$ is continuous from $W^k(\mathbb{R}^d;\mathbb{C})^2$ into $W^k(\mathbb{R}^d;\mathbb{C})$ (resp. $W(\mathbb{R}^d\times\mathbb{T};\mathbb{C})^2$ into $W(\mathbb{R}^d\times\mathbb{T};\mathbb{C})$).*
ii. *If M is a skew-symmetric, real valued $n\times n$ matrix, then $\exp(-iM)$ is unitary on $W^k(\mathbb{R}^d;\mathbb{C}^n)$ (resp. $W^k(\mathbb{R}^d\times\mathbb{T};\mathbb{C}^n)$).*
iii. *If $f\in W(\mathbb{R}^d;\mathbb{C})$ then for all $\beta>0$, one has $f(\frac{\cdot}{\beta})\in W(\mathbb{R}^d;\mathbb{C})$ and $|f(\frac{\cdot}{\beta})|_W=|f|_W$.*

Remark 2. *As said in the introduction, the third point of Proposition 1 is the main motivation to work with Wiener algebra rather than Sobolev spaces, because the $W(\mathbb{R}^d;\mathbb{C}^n)$-norm of initial conditions of the form (7) remains bounded (constant) while its H^s-norm is of size $O(\beta^{d/2-s})$ as $\beta\to0$. This framework is somehow a simplified version of the functional setting of* [6] *(where the SVEA is not made) and has also proved useful in the study of wave-packets* [5] *or localized solutions* [22].

2.2. **The envelope approximation.** This section is devoted to the justification of the envelope approximation which states that the exact solution to (10) can be described at leading order by u^ε_{app} defined as

$$u^\varepsilon_{app}(t,x)=U(t,x)e^{i\frac{\mathbf{k}\cdot x-\underline{\omega}t}{\varepsilon}}+\text{c.c.},\tag{17}$$

where the *envelope* U solves the *envelope equation*

$$\begin{cases}\partial_tU+\frac{i}{\varepsilon}\mathcal{L}(\underline{\omega},\underline{\mathbf{k}}+\varepsilon D)U=\varepsilon\mathcal{T}(U)\\ U_{|_{t=0}}(x)=U^0(x),\end{cases}\tag{18}$$

with $\mathcal{L}(\cdot,\cdot)$ given by (14) and $\mathcal{T}$ is defined as

$$\mathcal{T}(U)=T(\overline{U},U,U)+T(U,\overline{U},U)+T(U,U,\overline{U}).$$

The interest of the envelope equation (18) with respect to the original ivp (10) is that the fast oscillating scale has been removed from the initial data. The main result of this section is the following theorem.

Theorem 1. *Let Assumptions 1, 2 and 3 be satisfied, and let $a,b\in W^1(\mathbb{R}^d;\mathbb{C}^n)$, and*

$$U^0=\pi_1(\underline{\mathbf{k}})a+\varepsilon b\qquad\textit{(Polarization condition).}$$

i. *There exists a time $\tau_0>0$ such that for all $0<\varepsilon<1$, there is a unique solution $U\in C([0,\frac{\tau_0}{\varepsilon});W(\mathbb{R}^d;\mathbb{C}^n))$ to (18);*

ii. *For all $0 < \tau < \tau_0$, there exists ε_0 such that for all $0 < \varepsilon < \varepsilon_0$, there is a unique solution $u^\varepsilon_{ex} \in C([0, \tau/\varepsilon] \times \mathbb{R}^d)^n$ to (10), and one has*

$$|u^\varepsilon_{ex} - u^\varepsilon_{app}|_{L^\infty([0,\tau/\varepsilon]\times\mathbb{R}^d)^n} \leq \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W),$$

where u^ε_{app} is as defined in (17).

Remark 3. *Assuming that $|U^0|_W$ and $|b|_W$ are $O(1)$ quantities –which is of course the case for short pulses–, one deduces the following result*

$$\textit{The SVEA is valid if } |\nabla U^0|_W \ll \frac{1}{\varepsilon},$$

and Theorem 1 thus provides a rigorous basis for the "practical rule" (4). When working with short pulses with initial condition (7), it is easy to check that this condition reads simply $\varepsilon \ll \beta$.

Remark 4. *Working in the more classical framework of Sobolev spaces, one could establish an error estimate similar to the one given by the theorem, but with H^s-norms ($s > d/2$) instead of W-norms in the rhs of the estimates. For short pulses with initial data (7), the control would therefore be of the form $\varepsilon C(\tau, \frac{1}{\beta^{s-d/2}})\frac{1}{\beta^{s+(1-d)/2}}$, which is obviously useless when $\beta \to 0$.*

Proof. Let us prove the following lemma, which implies the first point of the theorem.

Lemma 1. *Let $U^0 \in W(\mathbb{R}^d; \mathbb{C}^n)$, and assume that Assumption 1 is satisfied. There exists $\tau_0 > 0$ such that for all $0 < \varepsilon < 1$ one has a unique solution $U \in C([0, \frac{\tau_0}{\varepsilon}); W(\mathbb{R}^d; \mathbb{C}^n))$ to (18). For all $0 < \tau < \tau_0$, one has*

$$(a) \qquad \sup_{0\leq t\leq \tau/\varepsilon} |U(t)|_W \leq C(\tau, |U^0|_W);$$

if moreover $U^0 \in W^1(\mathbb{R}^d; \mathbb{C}^n)$ then one also has

$$(b) \qquad \sup_{0\leq t\leq \tau/\varepsilon} |\nabla U(t)|_W \leq C(\tau, |U^0|_W)|\nabla U^0|_W.$$

Proof. Uniqueness is obvious, and to prove existence, we use a classical iterative method: let $U_0 = U^0$ and, for all $n \in \mathbb{N}$,

$$U_{n+1}(t) = S(t)U^0 + \varepsilon \int_0^t S(t - t')\mathcal{T}(U_n)dt',$$

with $S(t) = \exp(-\frac{it}{\varepsilon}\mathcal{L}(\underline{\omega}, \underline{\mathbf{k}} + \varepsilon D))$.
Since $\mathcal{L}(\underline{\omega}, \underline{\mathbf{k}}+\varepsilon\xi)$ is real and skew-symmetric for all $\xi \in \mathbb{R}^d$, we can use Proposition 1 and the trilinearity of T to get

$$\sup_{[0,t]} |U_{n+1}|_W \leq |U^0|_W + \varepsilon \text{Cst } t\big(\sup_{[0,t]} |U_n|_W\big)^3,$$

and

$$\sup_{[0,t]} |U_{n+1} - U_n|_W \leq \varepsilon \text{Cst } tC\big(\sup_{[0,t]} |U_n|_W, \sup_{[0,t]} |U_{n-1}|_W\big) \sup_{[0,t]} |U_n - U_{n-1}|_W,$$

and it follows easily that the sequence $(U_n)_n$ converges in $C([0, \frac{\tau}{\varepsilon}]; W^0(\mathbb{R}^d; \mathbb{C}^n))$ (for some $\tau > 0$) to a solution U of

$$U(t) = S(t)U^0 + \varepsilon \int_0^t S(t - t')\mathcal{T}(U)dt'; \tag{19}$$

the solution can then be extended to a maximal time interval $[0, \frac{\tau_{max}(\varepsilon)}{\varepsilon})$ and one can show with classical arguments that $\tau_0 := \inf_{0<\varepsilon<1} \tau_{\max}(\varepsilon) > 0$. The estimate (a) of the lemma also follows easily from (19) and a Gronwall-type lemma.
Differentiating (19) with respect to x_j $(j = 1, \dots, d)$ and using Proposition 1 to control the $W(\mathbb{R}^d; \mathbb{C}^n)$-norm, one gets

$$|\partial_j U(t)|_W \leq |\partial_j U^0|_W + \varepsilon \mathrm{Cst} \int_0^t |U(t')|_W^2 |\partial_j U(t')|_W dt',$$

and the estimate (b) follows from Gronwall's lemma and the estimate (a). □

Before going further in the proof of the theorem, let us introduce some notation. We decompose the solution U of (18) provided by Lemma 1 as

$$U = U_1 + \cdots + U_m, \quad \text{with} \quad U_j = \pi_j(\underline{\mathbf{k}} + \varepsilon D)U,$$

and we also write $U_{II} = U_2 + \cdots + U_m$.
The first step of the proof of the second part of the theorem consists in controlling $\partial_t U_1$ uniformly in ε –which is much better than the $O(1/\varepsilon)$ estimate on $\partial_t U$ one can deduce directly from the equation (18).

Lemma 2. *If Assumptions 1 and 2 are satisfied, then for all* $0 < \tau < \tau_0$,

$$\sup_{0 \leq t \leq \tau/\varepsilon} |\partial_t U_1|_W \leq C(\tau, |U^0|_W)\big(1 + |\nabla U^0|_W\big).$$

Proof. Multiplying the envelope equation (18) by $\pi_1(\underline{\mathbf{k}} + \varepsilon D)$, one gets

$$\partial_t U_1 + \frac{i}{\varepsilon}\big(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega}\big)U_1 = \varepsilon \pi_1(\underline{\mathbf{k}} + \varepsilon D)\mathcal{T}(U). \tag{20}$$

Recalling that $\underline{\omega} = \omega_1(\underline{\mathbf{k}})$, a first order Taylor expansion shows that

$$\big|\big(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega}\big)U_1\big|_W \leq \varepsilon |\nabla \omega_1|_{L^\infty} |\nabla U_1|_W$$

(note that one infers $|\nabla \omega_1|_{L^\infty} < \infty$ from the observation that for all $\mathbf{k} \neq 0$, $\partial_j \omega_1(\mathbf{k})\pi_1(\mathbf{k}) = \pi_1(\mathbf{k})A_j\pi_1(\mathbf{k})$). It follows therefore from (20), the trilinearity of T, and Proposition 1 that

$$|\partial_t U_1|_W \leq |\nabla \omega_1|_{L^\infty} |\nabla U_1|_W + \mathrm{Cst}\ \varepsilon |U|_W^3,$$

and the result follows from Lemma 1. □

We now prove that the components U_j $(j \geq 2)$ remain of size $O(\varepsilon)$ if this is initially the case.

Lemma 3. *If Assumptions 1, 2 and 3 are satisfied, and if* $U^0 = \pi_1(\underline{\mathbf{k}})a + \varepsilon b$, *then one has, for all* $0 < \tau < \tau_0$,

$$\sup_{t \in [0, \tau/\varepsilon]} |U_{II}(t)|_W \leq \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W).$$

Proof. Multiplying (18) by $\pi_j(\underline{\mathbf{k}} + \varepsilon D)$ $(j \geq 2)$ gives

$$\begin{aligned} \partial_t U_j + \frac{i}{\varepsilon}\big(\omega_j(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega}\big)U_j &= \varepsilon \pi_j(\underline{\mathbf{k}} + \varepsilon D)\mathcal{T}(U) \\ &= \varepsilon \pi_j(\underline{\mathbf{k}} + \varepsilon D)\mathcal{T}(U_1) + \varepsilon \pi_j(\underline{\mathbf{k}} + \varepsilon D)\big(\mathcal{T}(U) - \mathcal{T}(U_1)\big). \end{aligned}$$

With $S_j(t) = \exp\big(-i\frac{t}{\varepsilon}(\omega_j(\underline{\mathbf{k}}+\varepsilon D)-\underline{\omega})\big)$, one gets therefore

$$\begin{aligned} U_j(t) &= S_j(t)U_j^0 + \varepsilon\int_0^t S_j(t-t')\pi_j(\underline{\mathbf{k}}+\varepsilon D)\mathcal{T}(U_1)dt' \\ &\quad +\varepsilon\int_0^t S_j(t-t')\pi_j(\underline{\mathbf{k}}+\varepsilon D)\big(\mathcal{T}(U)-\mathcal{T}(U_1)\big)dt'. \end{aligned} \tag{21}$$

We now bound the W-norm of the three terms of the r.h.s. of (21):

- Estimate of $S_j(t)U_j^0$. Since $S_j(t)$ is unitary on $W(\mathbb{R}^d;\mathbb{C}^n)$, one has

$$|S_j(t)U_j^0|_W = |U_j^0|_W = |\pi_j(\underline{\mathbf{k}}+\varepsilon D)U^0|_W.$$

  Since moreover one can write

$$\pi_j(\underline{\mathbf{k}}+\varepsilon D)U^0 = (\pi_j(\underline{\mathbf{k}}+\varepsilon D)-\pi_j(\underline{\mathbf{k}}))U^0 + \pi_j(\underline{\mathbf{k}})U^0,$$

  it follows from the orthogonality of the projectors π_j $(j=1,\dots,m)$ that

$$\pi_j(\underline{\mathbf{k}}+\varepsilon D)U^0 = (\pi_j(\underline{\mathbf{k}}+\varepsilon D)-\pi_j(\underline{\mathbf{k}}))U^0 + \varepsilon\pi_j(\underline{\mathbf{k}})b.$$

  Since the derivatives of $\pi_j(\cdot)$ are in general not bounded near the origin, we cannot control the first term of the rhs by a Taylor expansion and we thus write

$$\begin{aligned} \pi_j(\underline{\mathbf{k}}+\varepsilon D)U^0 &= (\pi_j(\underline{\mathbf{k}}+\varepsilon D)-\pi_j(\underline{\mathbf{k}}))1_{\{\varepsilon|D|\le|\underline{\mathbf{k}}|/2\}}U^0 \\ &+ (\pi_j(\underline{\mathbf{k}}+\varepsilon D)-\pi_j(\underline{\mathbf{k}}))1_{\{\varepsilon|D|\ge|\underline{\mathbf{k}}|/2\}}U^0 + \varepsilon\pi_j(\underline{\mathbf{k}})b, \end{aligned}$$

  where $1_{\{\varepsilon|\xi|\le|\underline{\mathbf{k}}|/2\}}=1$ if $\varepsilon|\xi|\le|\underline{\mathbf{k}}|/2$ and 0 otherwise.
  Using the fact that $\pi_j(\cdot)$ is C^∞ on the ball of center $\underline{\mathbf{k}}$ and radius $|\underline{\mathbf{k}}|/2$, we can bound the first term of the rhs in $W(\mathbb{R}^d,\mathbb{C}^n)$-norm by $\varepsilon\mathrm{Cst}\ |\nabla U^0|_W$; one can also check that a similar estimate holds for the second term of the lhs since one has $1\le\frac{2\varepsilon}{|\underline{\mathbf{k}}|}|\xi|$ for all $\varepsilon|\xi|\ge|\underline{\mathbf{k}}|/2$. We can thus conclude that

$$|S_j(t)U_j^0|_W \le \varepsilon\mathrm{Cst}\ (|b|_W + |\nabla U^0|_W). \tag{22}$$

- Estimate of $A := \varepsilon\int_0^t S_j(t-t')\pi_j(\underline{\mathbf{k}}+\varepsilon D)\mathcal{T}(U_1)dt'$. Taking the Fourier transform of this term and integrating by parts yields

$$\begin{aligned} &\varepsilon\int_0^t \exp\big(-i\frac{t-t'}{\varepsilon}(\omega_j(\underline{\mathbf{k}}+\varepsilon\xi)-\underline{\omega})\big)\pi_j(\underline{\mathbf{k}}+\varepsilon\xi)\widehat{\mathcal{T}(U_1)}dt' \\ &= -i\varepsilon\int_0^t \frac{\varepsilon\exp\big(-i\frac{t-t'}{\varepsilon}(\omega_j(\underline{\mathbf{k}}+\varepsilon\xi)-\underline{\omega})\big)}{\omega_j(\underline{\mathbf{k}}+\varepsilon\xi)-\underline{\omega}}\pi_j(\underline{\mathbf{k}}+\varepsilon\xi)\partial_t\widehat{\mathcal{T}(U_1)}dt' \\ &\quad +i\varepsilon\Big[\frac{\varepsilon\exp\big(-i\frac{t-t'}{\varepsilon}(\omega_j(\underline{\mathbf{k}}+\varepsilon\xi)-\underline{\omega})\big)}{\omega_j(\underline{\mathbf{k}}+\varepsilon\xi)-\underline{\omega}}\pi_j(\underline{\mathbf{k}}+\varepsilon\xi)\widehat{\mathcal{T}(U_1)}\Big]_0^t. \end{aligned}$$

  One deduces therefore, using Assumption 3, that

$$\sup_{t\in[0,\tau/\varepsilon]}|A(t)|_W \le \mathrm{Cst}\ \frac{\varepsilon\tau}{c_0}\sup_{[0,\tau/\varepsilon]}|U_1|_W^2\sup_{[0,\tau/\varepsilon]}|\partial_t U_1|_W + \mathrm{Cst}\ \frac{\varepsilon^2}{c_0}\sup_{[0,\tau/\varepsilon]}|U_1|_W^3,$$

  so that, owing to Lemmas 1 and 2,

$$\sup_{t\in[0,\tau/\varepsilon]}|A(t)|_W \le \varepsilon C(\tau,|U^0|_W)(1+|\nabla U^0|_W). \tag{23}$$

- Estimate of $B := \varepsilon \int_0^t S_j(t-t')\pi_j(\underline{\mathbf{k}}+\varepsilon D)\big(\mathcal{T}(U)-\mathcal{T}(U_1)\big)dt'$. First remark that owing to the trilinearity of T, one has for all $t \in [0,\tau/\varepsilon]$,
$$\big|\mathcal{T}(U)(t)-\mathcal{T}(U_1)(t)\big|_W \leq \text{Cst} \sup_{[0,\tau/\varepsilon]} |U|_W^2 |U_{II}(t)|_W;$$
using Lemma 1, we obtain therefore
$$\sup_{t\in[0,\tau/\varepsilon]} |B(t)|_W \leq \varepsilon C(\tau,|U^0|_W)\int_0^t |U_{II}(t')|_W dt'. \tag{24}$$

It is now a direct consequence of $(21)_j$ $(j=2,\dots m)$ and (22)-(24) that for all $t\in[0,\tau/\varepsilon]$,
$$|U_{II}(t)|_W \leq \varepsilon(|b|_W + C(\tau,|U^0|_W)(1+|\nabla U^0|_W) + \varepsilon C(\tau,|U^0|_W)\int_0^t |U_{II}(t')|_W dt',$$
and the result follows therefore from Gronwall's lemma. □

We are now set to conclude the proof of the theorem. We look for an exact solution u^ε_{ex} to (10) under the form
$$u^\varepsilon_{ex}(t,x) = U_{ex}(t,x,\frac{\underline{\mathbf{k}}\cdot x - \underline{\omega} t}{\varepsilon}),$$
with $U_{ex}\in W(\mathbb{R}^d\times\mathbb{T};\mathbb{C}^n)$ itself of the form
$$U_{ex}(t,x,\theta) = U_{app}(t,x,\theta) + \varepsilon V(t,x,\theta),$$
with $U_{app}(t,x,\theta) = U(t,x)e^{i\theta} + \text{c.c.}$ and V bounded in $W(\mathbb{R}^d\times\mathbb{T};\mathbb{C}^n)$. With U as given by Lemma 1, the equation that V must solve is
$$\begin{aligned}\partial_t V + \frac{i}{\varepsilon}\mathcal{L}(\underline{\omega}D_\theta,\underline{\mathbf{k}}D_\theta+\varepsilon D)V &= T(U,U,U)e^{i3\theta} + \text{c.c.}\\ &+\big(T(U_{app}+\varepsilon V,\overline{U_{app}+\varepsilon V},U_{app}+\varepsilon V) - T(U_{app},\overline{U_{app}},U_{app})\big).\end{aligned}$$
Owing to the first part of Assumption 3, we can look for V under the form
$$V(t,x,\theta) = V_0(t,x,\theta) + \varepsilon V_1(t,x)e^{i3\theta} + \text{c.c.},$$
with $V_1 = -i\mathcal{L}(3\underline{\omega},3\underline{\mathbf{k}})^{-1}T(U_1,U_1,U_1)$; the resulting equation on V_0 is
$$\partial_t V_0 + \frac{i}{\varepsilon}\mathcal{L}(\underline{\omega}D_\theta,\underline{\mathbf{k}}D_\theta+\varepsilon D)V_0 = I_1+I_2+I_3, \tag{25}$$
with
$$\begin{aligned} I_1 &= \big(T(U,U,U)-T(U_1,U_1,U_1)\big)e^{i3\theta} + \text{c.c.}\\ I_2 &= -\varepsilon\big(\partial_t + A(\partial)\big)V_1 e^{3i\theta} + \text{c.c.}\\ I_3 &= \big(T(U_{app}+\varepsilon V,\overline{U_{app}+\varepsilon V},U_{app}+\varepsilon V) - T(U_{app},\overline{U_{app}},U_{app})\big).\end{aligned}$$
Let us now bound I_j $(j=1,2,3)$ in $W(\mathbb{R}^d\times\mathbb{T};\mathbb{C}^n)$ and for all $t\in[0,\tau/\varepsilon]$:

- From Lemmas 1 and 3, one gets
$$|I_1(t)|_{W(\mathbb{R}^d\times\mathbb{T})} \leq \varepsilon C(\tau,|U^0|_W)(1+|b|_W+|\nabla U^0|_W). \tag{26}$$
- From the definition of V_1 and Lemmas 1-2, one has directly
$$|I_2(t)|_{W(\mathbb{R}^d\times\mathbb{T})} \leq \varepsilon C(\tau,|U^0|_W)(1+|\nabla U^0|_W). \tag{27}$$
- From the trilinearity of T and Lemma 1, one gets
$$|I_3(t)|_{W(\mathbb{R}^d\times\mathbb{T})} \leq \varepsilon C(\tau,|U^0|_W)(1+|V_0(t)|_W+\varepsilon|V_0(t)|_W^2+\varepsilon^2|V_0(t)|_W^3). \tag{28}$$

By Proposition 1, the semigroup $S(t) = \exp(-i\frac{t}{\varepsilon}\mathcal{L}(\underline{\omega}D_\theta, \underline{\mathbf{k}}D_\theta + \varepsilon D))$ is unitary on $W(\mathbb{R}^d \times \mathbb{T})$, so that the estimates (26)-(28) allow one to conclude to the existence of a solution $V_0 \in C([0, \tau/\varepsilon]; W(\mathbb{R}^d \times \mathbb{T})^n)$ to (25) using a fixed point formulation similar to the one used in the proof of Lemma 1. After a Gronwall argument, one also gets

$$\sup_{0\le t\le \tau/\varepsilon} |V_0(t)|_W \le C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W). \tag{29}$$

Remark 5. *In fact, one finds that V_0 exists a priori on a time interval $[0, \tau'/\varepsilon]$, with $\tau' \le \tau$. However, by a classical procedure of continuous induction (such as in the proof of Theorem 3 in* [19] *for instance), one can get $\tau' = \tau$, provided that $0 < \varepsilon < \varepsilon_0$ with ε_0 small enough.*

It follows from the above that

$$\sup_{t\in[0,\tau/\varepsilon]} |U_{ex}(t) - U_{app}(t)|_{W(\mathbb{R}^d\times\mathbb{T})} \le \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W),$$

and the theorem follows therefore from the observation that

$$|u^\varepsilon_{ex} - u^\varepsilon_{app}|_{L^\infty([0,\tau/\varepsilon]\times\mathbb{R}^d)} \le \sup_{t\in[0,\tau/\varepsilon]} |U_{ex}(t) - U_{app}(t)|_{W(\mathbb{R}^d\times\mathbb{T})}.$$

□

2.3. **Approximations by scalar equations.**

2.3.1. *The full dispersion model.* The full dispersion model consists in approximating the exact solution to (10) by $u^\varepsilon_{app,1}$ defined as

$$u^\varepsilon_{app,1}(t,x) = U_{(1)}(t,x)e^{i\frac{\underline{\mathbf{k}}\cdot x - \underline{\omega}t}{\varepsilon}} + \text{c.c.}, \tag{30}$$

where $U_{(1)}$ solves the *full dispersion* scalar equation

$$\begin{cases} \partial_t U_{(1)} + \frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega})U_{(1)} = \varepsilon\pi_1(\underline{\mathbf{k}})\mathcal{T}(U_{(1)}) \\ U_{(1)\,|_{t=0}}(x) = U^0(x) \end{cases} \tag{31}$$

and with $\omega_1(\cdot)$ as in Assumption 1.
The following corollary shows that the full dispersion scalar equation yields an approximation of same precision than the envelope equation for times $t \in [0, \tau/\varepsilon]$.

Corollary 1 (Full dispersion model)**.** *Under the assumptions of Theorem 1, and for all $0 < \varepsilon < \varepsilon_0$ ($\varepsilon_0 > 0$ small enough), there exists a unique solution $U_{(1)} \in C([0, \tau_0/\varepsilon); W(\mathbb{R}^d; \mathbb{C}^n))$ to (31).*
For all $0 < \tau < \tau_0$, one also has

$$|u^\varepsilon_{ex} - u^\varepsilon_{app,1}|_{L^\infty([0,\tau/\varepsilon]\times\mathbb{R}^d)} \le \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W),$$

where $u^\varepsilon_{app,1}$ is as defined in (30).

Remark 6. *The quantity $U_{(1)}$ remains $\mathbb{C}^n$-valued, but we call (31) a* scalar *approximation because the operator $\frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega})$ is scalar, which is not the case of $\frac{i}{\varepsilon}\mathcal{L}(\underline{\omega}, \underline{\mathbf{k}} + \varepsilon D)$ in the envelope equation (18). The interest of the FD model is that $\frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega})u$ remains bounded for spectrally localized functions u, while $\frac{i}{\varepsilon}\mathcal{L}(\underline{\omega}, \underline{\mathbf{k}} + \varepsilon D)u$ is of order $O(1/\varepsilon)$. The fast oscillations of the nonpolarized modes must therefore be taken into account with the envelope approximation, and the discretization step must therefore be much smaller in numerical computations than for the FD model.*

Remark 7. i. *Performing the same analysis as in Remark 3, one can check that the "practical rule" also applies for the FD model.*
ii. *The FD model has been derived and studied in* [12] *(it is called "intermediate model" in that reference) for the study of chirped pulses with initial data (8). However the error estimate provided in* [12] *is of the form* $\varepsilon C(\tau, |U^0|_{H^s})$*, with* s *large, which, as explained in Remark 4 is not small for short pulses (or more generally large spectrum –including chirped– pulses). Note that a variant of the FD model with the orthoganal projector* $\pi_1(\underline{\mathbf{k}})$ *replaced by* $\pi_1(\underline{\mathbf{k}} + \varepsilon D)$ *in front of the nonlinearity is also studied in* [12] *and gives very good results.*

Proof. We omit the existence/uniqueness part of the corollary, since it is obtained with the same tools as for Theorem 1 (in particular, taking a smaller ε_0 if necessary, the existence time of the envelope equation is larger than the existence time for (31) and we can thus take the same τ_0 as in Theorem 1), and we thus focus on the error estimate.
Denoting as in the proof of Theorem 1 $U_1 = \pi_1(\underline{\mathbf{k}} + \varepsilon D)U$, where U is the solution of the envelope equation, one gets from Lemma 3 that

$$\sup_{t\in[0,\tau/\varepsilon]} |U(t) - U_1(t)|_W \leq \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W),$$

so that it suffices to control $|U_1(t) - U_{(1)}(t)|_W$ to prove the corollary. Applying $\pi_1(\underline{\mathbf{k}} + \varepsilon D)$ to (18), one gets

$$\partial_t U_1 + \frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega})U_1 = \varepsilon \pi_1(\underline{\mathbf{k}} + \varepsilon D)\mathcal{T}(U),$$

so that the difference $V = U_1 - U_{(1)}$ solves

$$\left\{ \begin{array}{l} \partial_t V + \frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}} + \varepsilon D) - \underline{\omega})V = \varepsilon \pi_1(\underline{\mathbf{k}} + \varepsilon D)\mathcal{T}(U) - \varepsilon \pi_1(\underline{\mathbf{k}})\mathcal{T}(U_{(1)}) \\ V_{|t=0}(x) = \pi_1(\underline{\mathbf{k}} + \varepsilon D)U^0 - U^0. \end{array} \right. \tag{32}$$

Remark now that

$$\begin{aligned} \pi_1(\underline{\mathbf{k}} + \varepsilon D)\mathcal{T}(U) - \pi_1(\underline{\mathbf{k}})\mathcal{T}(U_{(1)}) &= \big(\pi_1(\underline{\mathbf{k}} + \varepsilon D) - \pi_1(\underline{\mathbf{k}})\big)\mathcal{T}(U) \\ &+ \pi_1(\underline{\mathbf{k}})\big(\mathcal{T}(U) - \mathcal{T}(U_1)\big) + \pi_1(\underline{\mathbf{k}})\big(\mathcal{T}(U_1) - \mathcal{T}(U_{(1)})\big). \end{aligned} \tag{33}$$

Since $|(\pi_1(\underline{\mathbf{k}} + \varepsilon D) - \pi_1(\underline{\mathbf{k}}))\mathcal{T}(U)|_W \leq \varepsilon \mathrm{Cst}\ |\nabla \mathcal{T}(U)|_W$ (see the proof of (22)), one can use Lemma 1 to bound the first component of the rhs of (33) from above by $\varepsilon C(\tau, |U^0|_W)|\nabla U^0|_W$. The second component of (33) can be estimated exactly as the term I_1 in (26), while the last one is bounded from above in $W(\mathbb{R}^d; \mathbb{C}^n)$ by $C(|U_1|_W, |U_{(1)}|_W)|V|_W$. Since moreover $|U_1|_W$ is controlled by Lemma 1 and that a similar estimates also holds obviously for $|U_{(1)}|_W$, one deduces that for all $0 \leq t \leq \tau/\varepsilon$,

$$|\mathcal{T}(U(t)) - \mathcal{T}(U_{(1)}(t))|_W \leq \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W) + C(\tau, |U^0|_W)|V(t)|_W.$$

This inequality, together with an energy estimate on (32) and a Gronwall argument shows that

$$\sup_{t\in[0,\tau/\varepsilon]} |V(t)|_W \leq \varepsilon C(\tau, |U^0|_W)(1 + |b|_W + |\nabla U^0|_W),$$

where we also used the estimate $|\pi_1(\underline{\mathbf{k}} + \varepsilon D)U^0 - U^0|_W \leq \varepsilon \mathrm{Cst}\ (1 + |b|_W + |\nabla U^0|_W)$ (which is proved with the same arguments as (22)). □

2.3.2. *The nonlinear Schrödinger equation.* In the Schrödinger approximation, the exact solution to (10) is approximated by $u^\varepsilon_{app,2}$ defined as

$$u^\varepsilon_{app,2}(t,x) = U_{(2)}(t,x)e^{i\frac{\underline{\mathbf{k}}\cdot x-\underline{\omega}t}{\varepsilon}} + \text{c.c.}, \tag{34}$$

where $U_{(2)}$ solves the *nonlinear Schrödinger* equation

$$\begin{cases} \partial_t U_{(2)} + \big(\nabla\omega_1(\underline{\mathbf{k}})\cdot\nabla\big)U_{(2)} - \varepsilon\dfrac{i}{2}\big(\nabla\cdot H_{\underline{\mathbf{k}}}(\omega_1)\nabla\big)U_{(2)} = \varepsilon\pi_1(\underline{\mathbf{k}})\mathcal{T}(U_{(2)}) \\ U_{(2)\,|_{t=0}}(x) = U^0(x), \end{cases} \tag{35}$$

and where $H_{\underline{\mathbf{k}}}(\omega_1)$ stands for the Hessian of $\omega_1(\cdot)$ at $\underline{\mathbf{k}}$. One then has

Corollary 2 (Schrödinger approximation)**.** *Under the assumptions of Theorem 1, and for all $0<\varepsilon<\varepsilon_0$ ($\varepsilon_0>0$ small enough), there exists a unique solution $U_{(2)}\in C([0,\tau_0/\varepsilon);W(\mathbb{R}^d;\mathbb{C}^n))$ to (35).*
If moreover $U^0\in W^3(\mathbb{R}^d;\mathbb{C}^n)$ then for all $0<\tau<\tau_0$, one also has

$$|u^\varepsilon_{ex}-u^\varepsilon_{app,2}|_{L^\infty([0,\tau/\varepsilon]\times\mathbb{R}^d)} \le \varepsilon C(\tau,|U^0|_W)(1+|\nabla U^0|_W+|b|_W+|\mathfrak{c}_{Schrod}|_\infty|U^0|_{W^3}),$$

where $u^\varepsilon_{app,2}$ is as defined in (34) and

$$\mathfrak{c}_{Schrod}(\xi) := \frac{\big(\omega_1(\underline{\mathbf{k}}+\varepsilon\xi)-\big(\underline{\omega}+\varepsilon\nabla\omega_1(\underline{\mathbf{k}})\cdot\xi+\varepsilon^2\frac{1}{2}\xi\cdot H_{\underline{\mathbf{k}}}(\omega_1)\xi\big)\big)}{\varepsilon^3(1+|\xi|^3)}.$$

Remark 8. i. *A third order Taylor expansion of $\omega_1(\underline{\mathbf{k}}+\varepsilon\xi)$ at $\xi=0$ shows that $|\mathfrak{c}_{Schrod}|_\infty$ is finite and can be bounded from above independently from ε.*
ii. *The component $|\mathfrak{c}_{Schrod}|_\infty|U^0|_{W^3}$ of the error estimate does not appear for the full dispersion model. It is due to the approximation of the nonlocal operator $\frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}}+\varepsilon D)-\underline{\omega})$ (lhs of (31)) by the differential operator $\nabla\omega_1(\underline{\mathbf{k}})\cdot\nabla U_{(2)}-\varepsilon\frac{i}{2}\nabla\cdot H_{\underline{\mathbf{k}}}\nabla U_{(2)}$ (lhs of (35)). This error term is thus a* linear *effect.*
iii. *This additional term is responsible for the bad behavior of the Schrödinger equation to model short pulses (and more generally large spectrum waves such as chirped pulses). For instance, for initial data like (7), the precision of the Schrödinger approximation is of order $O(\varepsilon(\frac{1}{\beta}+\frac{|\mathfrak{c}_{Schrod}|_\infty}{\beta^3}))$ when $\beta\ll 1$. In order for the Schrödinger approximation to keep the same order of precision as the full dispersion model, one needs therefore to have $\frac{|\mathfrak{c}_{Schrod}|_\infty}{\beta^3}\lesssim\frac{1}{\beta}$, which requires that $\beta^2\ge|\mathfrak{c}_{Schrod}|_\infty$. This condition is far much restrictive than the practical rule $\beta\gg\varepsilon$.*
iv. *The Schrödinger approximation has been rigorously justified for systems of the form (10) (and also for some quasilinear generalizations) by Donnat, Joly, Métivier and Rauch* [14, 15] *and Lannes* [17] *but in the Sobolev framework which we saw is not adapted for the study of large spectrum pulses.*

Proof. As in the proof of Corollary 1, we focus on the error estimate and omit the existence/uniqueness part of the proof.
The difference $V=U_{(1)}-U_{(2)}$ of the solution of the full dispersion and Schrödinger equations solves the initial value problem

$$\begin{cases} \partial_t V + \frac{i}{\varepsilon}(\omega_1(\underline{\mathbf{k}}+\varepsilon D)-\underline{\omega})V = \varepsilon\pi_1(\underline{\mathbf{k}})\big(\mathcal{T}(U_{(1)})-\mathcal{T}(U_{(2)})\big)-\varepsilon^2\mathcal{R}_2(D)U_{(2)} \\ V_{|_{t=0}}(x)=0. \end{cases} \tag{36}$$

where for all $\xi\in\mathbb{R}^d$,

$$\mathcal{R}_2(\xi) = \frac{1}{\varepsilon^3}\big(i\omega_1(\underline{\mathbf{k}}+\varepsilon\xi)-i\underline{\omega}-i\varepsilon\nabla\omega_1(\underline{\mathbf{k}})\cdot\xi-\varepsilon^2\frac{i}{2}\xi\cdot H_{\underline{\mathbf{k}}}(\omega_1)\xi\big).$$

Remark now that one has for all $0 \le t \le \tau/\varepsilon$,

$$|\mathcal{R}_2(D)U_{(2)}(t)|_W \le |\mathfrak{c}_{Schrod}|_\infty |U_2(t)|_{W^3},$$

with $\mathfrak{c}_{Schrod}(\cdot)$ as in the statement of the corollary; differentiating the Schrödinger equation (35) and estimating the W-norm of the solution, one also gets easily

$$\sup_{t\in[0,\tau/\varepsilon]} |U_{(2)}(t)|_{W^3} \le C(\tau, |U^0|_W)(1+|U^0|_{W^3}).$$

Since the first term of the r.h.s. of (36) can be bounded as in (26), one gets from Gronwall's lemma applied to (36) that

$$\sup_{t\in[0,\tau/\varepsilon]} |V(t)|_W \le C(\tau, |U^0|_W)(1+|\mathfrak{c}_{Schrod}|_\infty |U^0|_{W^3})$$

which, together with Corollary 1, yields the result. □

2.3.3. *The nonlinear Schrödinger equation with improved dispersion relation.* As said in the introduction, we propose in this paper new approximations based on a family of modified Schrödinger equations, whose dispersive properties are closer to the exact model. Such an approximation $u^\varepsilon_{app,3}$ is defined as

$$u^\varepsilon_{app,3}(t,x) = U_{(3)}(t,x)e^{i\frac{\mathbf{k}\cdot x-\underline{\omega}t}{\varepsilon}} + \text{c.c.}, \tag{37}$$

where $U_{(3)}$ solves the *nonlinear Schrödinger equation with improved dispersion relation*:

$$\begin{cases} \big(1 - i\varepsilon\mathbf{b}\cdot\nabla - \varepsilon^2\nabla\cdot B\nabla\big)\partial_t U_{(3)} \\ \qquad + \big(\nabla\omega_1(\underline{\mathbf{k}})\cdot\nabla - i\varepsilon\nabla\cdot(\frac{1}{2}H_{\underline{\mathbf{k}}}(\omega_1) + \nabla\omega_1(\underline{\mathbf{k}})\mathbf{b}^T)\nabla + \varepsilon^2\mathbf{C}(\nabla)\big)U_{(3)} \\ \qquad = \varepsilon\pi_1(\underline{\mathbf{k}})\mathcal{T}(U_{(3)}) \\ U_{(3)\,|_{t=0}}(x) = U^0(x), \end{cases} \tag{38}$$

where $\mathbf{b}\in\mathbb{C}^d$, $B\in\mathcal{M}_{d\times d}(\mathbb{R})$ and $\mathbf{C}:\mathbb{C}^d\times\mathbb{C}^d\times\mathbb{C}^d\to\mathbb{C}$ is a trilinear mapping. We assume moreover that

$$B \text{ is symmetric positive}, \quad \mathbf{b}\in \mathit{Range}(B), \quad \text{and} \quad 4-\mathbf{b}\cdot(B^{-1}\mathbf{b}) > 0 \tag{39}$$

(note that even though $B^{-1}\mathbf{b}$ is not unique when B is not definite, the scalar $\mathbf{b}\cdot(B^{-1}\mathbf{b})$ is uniquely defined). One then has the following result:

Corollary 3 (Improved Schrödinger approximation)**.** *Under the assumptions of Theorem 1, and for all $0<\varepsilon<\varepsilon_0$ ($\varepsilon_0>0$ small enough), there exists a unique solution $U_{(3)}\in C([0,\tau_0/\varepsilon);W(\mathbb{R}^d;\mathbb{C}^n))$ to (38).*
If moreover $U^0\in W^3(\mathbb{R}^d;\mathbb{C}^n)$ then for all $0<\tau<\tau_0$, one also has

$$|u^\varepsilon_{ex}-u^\varepsilon_{app,3}|_{L^\infty([0,\tau/\varepsilon]\times\mathbb{R}^d)} \le \varepsilon C(\tau,|U^0|_W)(1+|\nabla U^0|_W+|b|_W+|\mathfrak{c}_{improved}|_\infty|U^0|_{W^3}),$$

where $u^\varepsilon_{app,3}$ is as defined in (37) and

$$\mathfrak{c}_{improved}(\xi) := \frac{\omega_1(\underline{\mathbf{k}}+\varepsilon\xi) - \big(\underline{\omega} + \varepsilon\frac{\nabla\omega_1(\underline{\mathbf{k}})\cdot\xi+\varepsilon\xi\cdot(\frac{1}{2}H_{\underline{\mathbf{k}}}(\omega_1)+\nabla\omega_1(\underline{\mathbf{k}})\mathbf{b}^T)\xi-\varepsilon^2\mathbf{C}(\xi)}{1+\varepsilon\mathbf{b}\cdot\xi+\varepsilon^2\xi\cdot B\xi}\big)}{\varepsilon^3(1+|\xi|^3)}.$$

Remark 9. i. *As for the Schrödinger equation, one can check by a simple Taylor expansion that $|\mathfrak{c}_{improved}|_\infty$ is finite and uniformly bounded with respect to ε.*
ii. *Taking $\mathbf{b}=0$, $B=0$ and $\mathbf{C}=0$ (this choice satisfies (39)), one recovers the usual Schrödinger equation (35).*
iii. *The interest of (38) with respect to (35) is that one can choose $\mathbf{b}$, B and $\mathbf{C}$ such that $\mathfrak{c}_{improved}\ll\mathfrak{c}_{Schrod}$, thus improving considerably the accuracy of the*

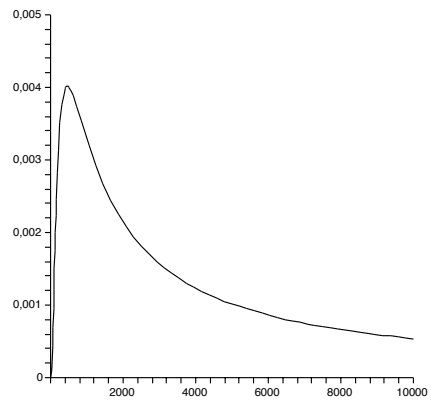

FIGURE 3. $\frac{\mathfrak{c}_{improved}(\xi)}{\mathfrak{c}_{Schrod}(\xi)}$ for $\varepsilon = 0.01$ with the coefficients (40).

approximation. In the one dimensional case $d = 1$, it is possible to choose $\mathbf{b}$, $\mathbf{B}$ and $\mathbf{C}$ in such a way that the dispersion relation for (38) is the [3,2]-Padé expansion of the dispersion relation of (31). For the case of the Klein-Gordon system (12), this leads to

$$\mathbf{b} = \frac{2\underline{\mathbf{k}}}{\mathbf{v}^2 + \underline{\mathbf{k}}^2}, \qquad \mathbf{B} = \frac{\mathbf{v}^2 + 4\underline{\mathbf{k}}^2}{4(\mathbf{v}^2 + \underline{\mathbf{k}}^2)^2}, \quad \mathbf{C} = \frac{\underline{\mathbf{k}}(3\mathbf{v}^2 + 4\underline{\mathbf{k}}^2)}{4(\mathbf{v}^2 + \underline{\mathbf{k}}^2)^{5/2}}; \tag{40}$$

we illustrate in Figure 3 how much one gains by working with (38) instead of (35) for the Klein-Gordon system (12) with $\mathbf{v} = \underline{\mathbf{k}} = 1$.

iv. *The same analysis as is Remark 8.iii shows that the approximation provided by (38) is of the same order as the envelope approximation if $\beta^2 \geq |\mathfrak{c}_{improved}|_\infty$. Since $|\mathfrak{c}_{improved}|_\infty \ll |\mathfrak{c}_{Schrod}|_\infty$, this condition is much weaker than the corresponding one for the usual Schrödinger model. In some particular cases, this condition can even be weaker than the "practical rule" $\varepsilon \ll \beta$.*

Proof. Choosing $\xi_0 \in -\frac{1}{2}B^{-1}\mathbf{b}$, one can check that

$$1 + \mathbf{b}\cdot\xi + \xi\cdot B\xi = 1 - \frac{1}{4}\mathbf{b}\cdot(B^{-1}\mathbf{b}) + (\xi - \xi_0)\cdot B(\xi - \xi_0),$$

so that it follows from assumption (39) that $1 + \mathbf{b}\cdot\xi + \xi\cdot B\xi > 0$ (uniformly with respect to $\xi \in \mathbb{R}^d$). The operator $1 - \varepsilon i\mathbf{b} - \varepsilon^2\nabla\cdot B\nabla$ is therefore invertible, and its inverse is the Fourier multiplier $(1 + \varepsilon\mathbf{b}\cdot D + \varepsilon^2 D\cdot BD)^{-1}$. The equation (38) can therefore be rewritten as

$$\begin{aligned}\partial_t U_{(3)} + i\frac{\nabla\omega_1(\underline{\mathbf{k}})\cdot D + \varepsilon D\cdot(\frac{1}{2}H_{\underline{\mathbf{k}}}(\omega_1) + \nabla\omega_1(\underline{\mathbf{k}})\mathbf{b}^T)D - \varepsilon^2\mathbf{C}(D)}{(1 + \varepsilon\mathbf{b}\cdot D + \varepsilon^2 D\cdot BD)}U_{(3)} \\ = (1 + \varepsilon\mathbf{b}\cdot D + \varepsilon^2 D\cdot BD)^{-1}\pi_1(\underline{\mathbf{k}})\mathcal{T}(U_{(3)}).\end{aligned}$$

Since $(1+\varepsilon\mathbf{b}\cdot D+\varepsilon^2 D\cdot BD)^{-1}$ is regularizing (of order -2) and acts on $W(\mathbb{R}^d;\mathbb{C}^n)$ uniformly with respect to $\varepsilon > 0$, the proof of the result follows exactly the same lines as the proof of Corollary 3 and we thus omit it. □

3. Numerical simulations

3.1. **The equations.** This section is devoted to the comparison of solutions of the different asymptotic equations derived in Section 2 with the solutions of the full system (10). In particular, we want to check numerically the results proved in Corollaries 1, 2 and 3.

We consider both short pulses (§3.3) and chirped pulses (§3.4); for the numerical computations, we use the toy model (12) with dimension $d = 1$ and $\mathbf{v} = 1$ (this is also the model used in [12] for the study of chirped pulses). Writing $u = (f, g)^T$, this model reads

$$\left\{\begin{array}{l} \partial_t f + \partial_x g - \dfrac{g}{\varepsilon} = -\varepsilon(|f|^2 + |g|^2)g, \\ \partial_t g + \partial_x f + \dfrac{f}{\varepsilon} = \varepsilon(|f|^2 + |g|^2)f. \end{array}\right. \tag{41}$$

The initial conditions are taken of the form

$$f_{|t=0}(x) = f^0(x)e^{i\frac{x}{\varepsilon}} + \text{c.c.}, \qquad g_{|t=0}(x) = \frac{1-i}{\sqrt{2}} f^0(x)e^{i\frac{x}{\varepsilon}} + \text{c.c.} \tag{42}$$

(which, with the notations of Theorem 1, corresponds to $\underline{\mathbf{k}} = 1$, $b = 0$ and $a = \pi_1(\underline{\mathbf{k}})a = (f^0, \frac{1-i}{\sqrt{2}} f^0)^T$).

The asymptotic models derived in Section 2 read in this particular case as follows:

- *The full dispersion model.* The exact solution to (41) is approximated by

$$u^\varepsilon_{app,1}(t,x) = U_{(1)}(t,x)e^{i\frac{x-\sqrt{2}t}{\varepsilon}} + \text{c.c.},$$

  where $U_{(1)} = (f_{(1)}, g_{(1)})$ solves (recall that $D = -i\partial_x$)

$$\left\{\begin{array}{l} \partial_t f_{(1)} + \frac{i}{\varepsilon}\frac{-D+\varepsilon D^2}{\sqrt{1+(1+\varepsilon D)^2}+\sqrt{2}} f_{(1)} = \frac{4i\varepsilon}{\sqrt{2}}|f_{(1)}|^2 f_{(1)}, \\ g_{(1)} = \frac{1-i}{\sqrt{2}} f_{(1)} \ \text{ (polarization condition).} \end{array}\right. \tag{43}$$

- *The nonlinear Schrödinger equation.* The exact solution to (41) is approximated by

$$u^\varepsilon_{app,2}(t,x) = U_{(2)}(t,x)e^{i\frac{x-\sqrt{2}t}{\varepsilon}} + \text{c.c.},$$

  where $U_{(2)} = (f_{(2)}, g_{(2)})$ solves

$$\left\{\begin{array}{l} \partial_t f_{(2)} + \frac{1}{\sqrt{2}}\big(\partial_x f_{(2)} - \frac{\varepsilon}{4}\partial_x^2 f_{(2)}\big) = \varepsilon\frac{4i}{\sqrt{2}}|f_{(2)}|^2 f_{(2)}, \\ g_{(2)} = \frac{1-i}{\sqrt{2}} f_{(2)} \ \text{ (polarization condition).} \end{array}\right. \tag{44}$$

- *The nonlinear Schrödinger equation with improved dispersion relation.* We approximate the solution of (41) by

$$u^\varepsilon_{app,3}(t,x) = U_{(3)}(t,x)e^{i\frac{x-\sqrt{2}t}{\varepsilon}} + \text{c.c.},$$

  where $U_{(3)} = (f_{(3)}, g_{(3)})$ solves

$$\left\{\begin{array}{l} \big(1 - i\varepsilon\partial_x - \varepsilon^2\frac{5}{16}\partial_x^2\big)\partial_t f_{(3)} + \frac{1}{\sqrt{2}}\big(\partial_x - i\varepsilon\frac{5}{4}\partial_x^2 + \varepsilon^2\frac{7}{16}\partial_x^3\big)f_{(3)} = \frac{4i\varepsilon}{\sqrt{2}}|f_{(3)}|^2 f_{(3)}, \\ g_{(3)} = \frac{1-i}{\sqrt{2}} f_{(3)} \ \text{ (polarization condition);} \end{array}\right. \tag{45}$$

  note that this modified Schrödinger equation corresponds to the set of coefficients (40).

3.2. **The numerical scheme.** We use a spectral method in space and a splitting technique in time for all the equations introduced in the previous section. We give here some details on the numerical scheme used for (41); for (43), (44) and (45), we use straightforward adaptations of this scheme.
Let us denote by $S_L(t)$ and $S_{NL}(t)$ the evolution operator associated respectively to the linear and nonlinear part of (41); namely,

$$S_L(t)u^0 = (f(t), g(t)), \text{ with } \begin{cases} \partial_t f + \partial_x g - \dfrac{g}{\varepsilon} = 0 \\ \partial_t g + \partial_x f + \dfrac{f}{\varepsilon} = 0 \end{cases} \text{ and } (f(0), g(0)) = u^0,$$

and

$$S_{NL}(t)u^0 = (f(t), g(t)), \text{ with } \begin{cases} \partial_t f = -\varepsilon(|f|^2 + |g|^2)g \\ \partial_t g = \varepsilon(|f|^2 + |g|^2)g \end{cases} \text{ and } (f(0), g(0)) = u^0$$

(and with periodic boundary conditions).
The numerical computation of $S_L(t)$ is made through an FFT-based spectral method while an explicit integration is used for $S_{NL}(t)$; we then use a second order splitting scheme to compute $u^{n+1} \sim u((n+1)\Delta t)$ in terms of $u^n \sim u(n\Delta t)$ (where Δt denotes the time step):

$$u^{n+1} = S_L(\frac{\Delta t}{2})S_{NL}(\Delta t)S_L(\frac{\Delta t}{2})u^n.$$

3.3. **Numerical results for short pulses.** In this section, we are interested in short pulses, that is we consider initial conditions for (41) of the form (42), with

$$f^0(x) = G(\frac{x - x_0}{\beta}),$$

where G is a smooth function. In the present numerical computations, the computational domain is $[0, L]$ with $L = 30\pi$, and we take $x_0 = 15$ and $G(x) = e^{-x^2}$.

The accuracy of the approximations (43), (44) and (45) is checked using the following quantity:

$$E_{(j)}(\varepsilon, \beta) = \sup_{t \in [0, \frac{1}{\varepsilon}]} \frac{|f(t, \cdot) - (f_{(j)}(t, \cdot)e^{i\frac{kx - \omega t}{\varepsilon}} + \text{c.c.})|_\infty}{|f(t, \cdot)|_\infty}, \tag{46}$$

where $j = 1$ for the full dispersion model (43), $j = 2$ for the usual Schrödinger approximation (44), and $j = 3$ for our new modified Schrödinger equation (45).
The exact solution and the difference between the exact solution with the approximation furnished by the FD, Schrödinger and improved Schrödinger models are plotted in Figure 4 for $\varepsilon = 0.01$, $\beta = 0.075$ and at time $T = 50$ on the domain $x \in [0, 30\pi]$. The following computations are also performed to test the accuracy of the approximate models:
• *Test 1:* With $\beta = 1$ fixed, we let ε vary from $\varepsilon = 0.001$ to $\varepsilon = 0.1$. This configuration corresponds to usual wave packets for which the three models should have a comparable accuracy of $O(\varepsilon)$ when ε is small enough. One can indeed observe on Figure 5 that the errors $E_{(j)}(\varepsilon, \beta)$ $(j = 1, 2, 3)$ grow linearly with ε. One will also check that when ε is too large ($\varepsilon \sim 5.10^{-2}$ for a rough precision of 20%), none of the models furnishes a good approximation.
• *Test 2:* Here, we look at the same configuration as in Test 1 but with $\beta = 0.1$, that is, we investigate here short pulses. We can observe on Figure 6 that the

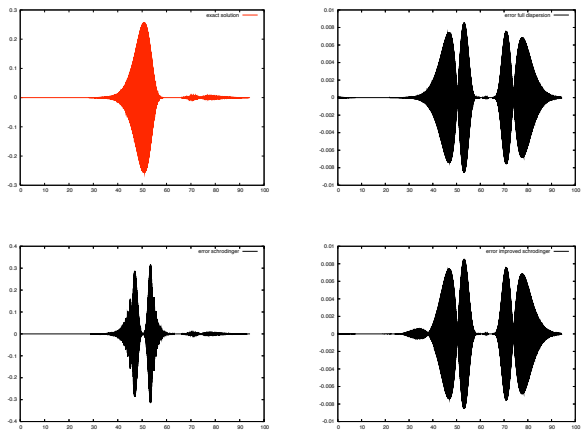

FIGURE 4. Short pulses: the exact solution, and the difference between the exact solution and the FD, Schrödinger and improved Schrödinger models (from left to right and top to bottom) with $\varepsilon = 0.01$, $\beta = 0.075$ and $T = 50$.

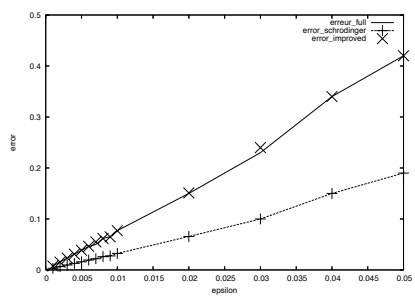

FIGURE 5. The errors $E_{(j)}(\varepsilon, \beta)$ for $\beta = 1$ and $\varepsilon \in [0.001, 0.1]$; $j = 1$ corresponds to FD, $j = 2$ to Schrödinger and $j = 3$ to the improved Schrödinger.

FD and improved Schrödinger models provide a good approximation, but that the usual Schrödinger approximation is completely inaccurate.

• *Test 3:* Here, $\varepsilon = 0.01$ is fixed and we let β vary from $\beta = 0.01$ (short pulses) to $\beta = 1$ (wave packets). It can be checked that the FD model furnishes a correct approximation for $\beta \gtrsim 0.03$ and that for such values of β, the improved Schrödinger approximation has the same precision. This is to be contrasted with the usual Schrödinger approximation which is completely inaccurate until $\beta \sim 0.2$.

3.4. **Numerical results for chirped pulses.** In this section, we are interested in chirped pulses, that is we consider initial conditions for (41) of the form (42), with

$$f^0(x) = G(x - x_0)\cos(\frac{1}{\beta}\cos(\frac{x - x_0}{\beta})),$$

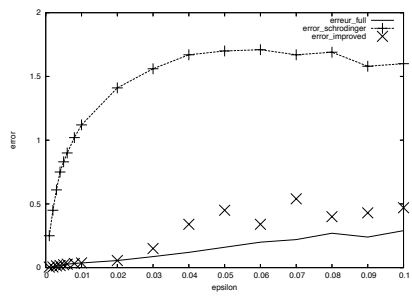

FIGURE 6. Short pulses: The errors $E_{(j)}(\varepsilon, \beta)$ for $\beta = 0.1$ and $\varepsilon \in [0.001, 0.1]$; $j = 1$ corresponds to FD, $j = 2$ to Schrödinger and $j = 3$ to the improved Schrödinger.

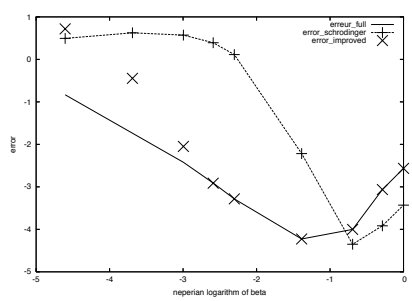

FIGURE 7. Short pulses: The errors $E_{(j)}(\varepsilon, \beta)$ for $\varepsilon = 0.01$ and $\beta \in [0.01, 1]$; $j = 1$ corresponds to FD, $j = 2$ to Schrödinger and $j = 3$ to the improved Schrödinger.

where G is a smooth function. In the present numerical computations, the computational domain is $[0, L]$ with $L = 30\pi$, and we take $x_0 = 15$ and $G(x) = e^{-x^2}$.

The accuracy of the approximations (43), (44) and (45) is checked using the quantities $E_{(j)}(\varepsilon, \beta)$ $(j = 1, 2, 3)$ defined in (46).
The exact solution and the diffenrence between the exact solution and the FD, Schrödinger and improved Schrödinger models are plotted in Figure 8 for $\varepsilon = 0.01$, $\beta = 0.3$ and at time $T = 1/\varepsilon = 100$. The following computations are also performed to test the accuracy of the approximate models:

• *Test 1:* With $\beta = 0.1$ fixed, we let ε vary from $\varepsilon = 0.001$ to $\varepsilon = 0.1$. We can observe on Figure 9 that the FD and the improved Schrödinger models are good approximations for $\varepsilon \leq 0.003$. Above this value, the approximation is no longer pertinent. Furthermore, the classical Schrödinger model is inapropriate for this range of parameters.

• *Test 2:* Here, $\varepsilon = 0.01$ is fixed and we let β vary from $\beta = 0.01$ (chirped pulses) to $\beta = 1$ (wave packets). We observe on Figure 10 that both FD and improved

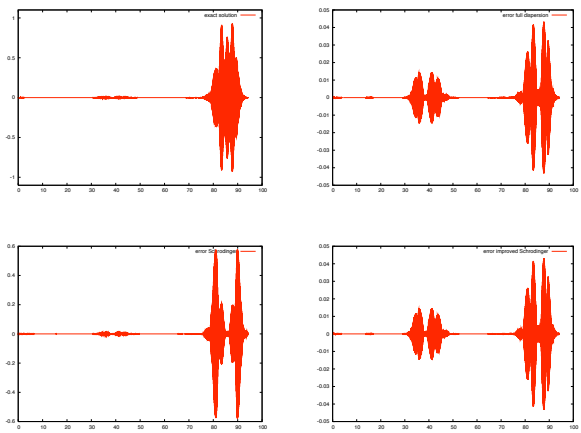

FIGURE 8. Chirped pulses: the exact solution, and the difference between the exact solution and the FD, Schrödinger and improved Schrödinger models (from left to right and top to bottom) with $\varepsilon = 0.01$, $\beta = 0.3$ and $T = 100$.

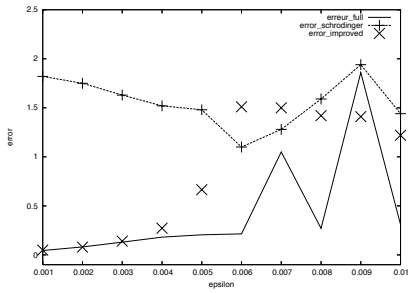

FIGURE 9. Chirped pulses: The errors $E_{(j)}(\varepsilon, \beta)$ for $\beta = 0.01$ and $\varepsilon \in [0.001, 0.1]$; $j = 1$ corresponds to FD, $j = 2$ to Schrödinger and $j = 3$ to the improved Schrödinger.

Schrödinger models become appropriate for $\beta \geq 0.1$ whereas the Schrödinger approximation is acceptable for $\beta \geq 0.4$.

REFERENCES

[1] D. ALTERMAN, J. RAUCH, *Diffractive short pulse asymptotics for nonlinear wave equations*, Physics Letters A **264** (2000),390–395.

[2] D. ALTERMAN, J. RAUCH, *The linear diffractive pulse equation*, Methods and Applications of Analysis, **7** (2000), no. 2, 263–274.

[3] D. ALTERMAN, J. RAUCH, *Diffractive Nonlinear Geometric Optics for Short Pulses*, SIAM J. Math Anal. **34** (2003) 1477-1502.

[4] A. BABIN, A. FIGOTIN, *Linear superposition in nonlinear wave dynamics*, Rev. Math. Phys. **18** (2006), no. 9, 971-1053.

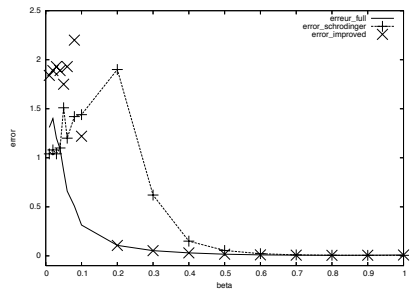

FIGURE 10. Chirped pulses: The errors $E_{(j)}(\varepsilon, \beta)$ for $\varepsilon = 0.01$ and $\beta \in [0.01, 1]$; $j = 1$ corresponds to FD, $j = 2$ to Schrödinger and $j = 3$ to the improved Schrödinger.

[5] A. Babin, A. Figotin, *Nonlinear dynamics of a system of particle-like wavepackets*, Preprint, arXiv:math/0611656.

[6] K. Barrailh and D. Lannes, *A general framework for diffractive optics and its applications to lasers with large spectrums and short pulses*, SIAM J. Math. Anal. **34** (2002), no. 3, 636-674.

[7] Benjamin T B, Bona J L and Mahony J J, *Model equations for long waves in nonlinear dispersive systems*, Phil. Trans. R. Soc. Lond. A **272** (1972) 47-78.

[8] J. L. Bona, M. Chen, J.-C. Saut, *Boussinesq Equations and Other Systems for Small-Amplitude Long Waves in Nonlinear Dispersive Media. I: Derivation and Linear Theory*, J. Nonlinear Sci. **12** (2002), 283-318.

[9] J. L. Bona, T. Colin, D. Lannes, *Long Waves Approximations for Water Waves*, Arch. Rational Mech. Anal. **178** (2005) 373-410.

[10] R. W. Boyd, *Nonlinear Optics*, Academic Press, Boston, 1992.

[11] T. Brabec and F. Krausz, *Nonlinear optical pulse propagation in the single-cycle regime*, Phys. Rev. Lett. **78** (1997), 3282-3285.

[12] T. Colin and G. Gallice and K. Laurioux, *Intermediate models in nonlinear optics*, SIAM J. Math. Anal. **36** (2005), no. 5, 1664-1688.

[13] Y. Chung and C. K. R. T. Jones and C. E. Schafer, T. and Wayne, *Ultra-short pulses in linear and nonlinear media*, Nonlinearity **18** (2005), no. 3, 1351-1374.

[14] P. Donnat, J.-L. Joly, G. Métivier, J. Rauch, *Diffractive nonlinear optics*, Séminaire Equations aux Dérivées, Partielles, Ecole Polytechnique, Exposé No. XVII (1996).

[15] J.-L. Joly, G. Métivier, J. Rauch, *Diffractive Nonlinear Geometric Optics With Rectification*, Indiana Univ. Math. Journ. **47** (1998), no. 4.

[16] H. Hasimoto and H. Ono, *Nonlinear Modulation of Gravity Waves*, Journal of the Physical Society of Japan **33** (1972), no. 3, 805-811.

[17] D. Lannes, *Dispersive effects for nonlinear geometrical optics with rectification*, Asymptot. Anal. **18** (1998), no. 1-2, 111–146.

[18] D. Lannes, *Nonlinear geometrical optics for oscillatory wave trains with a continuous oscillatory spectrum*, Adv. Differential Equations **6** (2001), no. 6, 731–768.

[19] D. Lannes and J. Rauch, *Validity of nonlinear geometric optics with times growing logarithmically*, Proc. Amer. Math. Soc. **129** (2001), no. 4, 1087-1096.

[20] H. Leblond and M. Manna, *Benjamin-Feir-type instability in a saturated ferrite: Transition between focusing and defocusing regimes for polarized electromagnetic waves*, Phys. Rev. E **50** (1994), no. 3, 2275-2286.

[21] A. C. Newell and J. V. Moloney, *Nonlinear Optics*, Addison-Wesley Reading, MA, 1991.

[22] D. Pelinovsky and G. Schneider, *Justification of the coupled-mode approximation for a nonlinear elliptic problem with a periodic potential*, Applicable Analysis **86** (2007), no. 8, 1017-1036.

[23] T. Schäfer and C. E. Wayne, *Propagation of ultra-short optical pulses in cubic nonlinear media*, Phys. D **196** (2004), no. 1-2, 90-105.
[24] V.E. Zakharov, *Stability of periodic waves of finite amplituse on the surface of a deep fluid*, J. Appl. Mech. Tech. Phys. **2** (1968), 190-194.

Université Bordeaux I; IMB et CNRS UMR 5251, 351 Cours de la Libération,, 33405 Talence Cedex, France
E-mail address: mcolin@math.u-bordeaux1.fr

Université Bordeaux I; IMB et CNRS UMR 5251, 351 Cours de la Libération,, 33405 Talence Cedex, France
E-mail address: lannes@math.u-bordeaux1.fr