Semi-discretization in time for Nonlinear Zakharov Waves Equations.

T. Colin[2], G. Ebrard[1,2], G. Gallice[1]

1 *CEA CESTA, SIS, BP 2 33114 Le Barp, France.*
2 *IMB University Bordeaux 1, Inria Futurs project MC2*
351 cours de la Libération, 33405 Talence, France

Abstract

In this paper we construct and study discretizations of a nonlinear Zakharov-wave system occurring in plasma physics. These systems are generalizations of the Zakharov system that can be recovered by taking a singular limit. We introduce two numerical schemes that take into account this singular limit process. One of the scheme is conservative but sensible to the polarization of the initial data while the other one is able to handle ill-prepared initial data. We prove some convergence results and we perform some numerical tests.

Key words: Zakharov systems, Langmuir turbulence, Crank-Nicholson, polarization.

AMS subject classification: 65M06, 65M12, 78A60

1 Introduction

1.1 Setting of the problem and motivation

The strong Langmuir turbulence phenomenon is well-known problem in plasma physics [13]. The key mechanism of this physical phenomenon is governed by the so-called Zakharov equations. These equations describe the slowly varying motions of the complex-valued envelope of the Langmuir electric field E^0 which nonlinearly interacts with large-scale fluctuation n^0 through the ponderomotrice force. They read in a dimensionless scalar form

$$\begin{cases} 2i\partial_t E^0 + \Delta E^0 = n^0 E^0, \\ \\ (\partial_t^2 - \Delta)n^0 = 2\Delta|E^0|^2. \end{cases} \tag{1.1}$$

This system can be derived from the Klein-Gordon-wave system [4] or directly from the Euler-Maxwell system [15] using a small parameter $\varepsilon = \frac{1}{\omega_{pe}T}$ where ω_{pe} is the electronic plasma pulsation and T a characteristic time of the pulse. The size of ε governs the regime of evolution of the solution.

System (1.1) is only an approximation that can be sometimes not accurate enough to describe the full physics corresponding to the large variety of plasmas created by laser pulses. In [2], an intermediate system (between (1.1) and Euler-Maxwell in the hierarchy) has been introduced. It is less stiff than Euler-Maxwell but more precise than (1.1). In particular, it takes into account more deeply the oscillations of the electric field and allows "larger" values of ε. Its range of validity is expected to be wider than that of Zakharov. It reads, in a dimensionless form :

$$\begin{cases} \varepsilon\partial_t^2\mathcal{E} - 2i\partial_t\mathcal{E} - \Delta\mathcal{E} = -\mathcal{N}\mathcal{E}, \\ \\ \partial_t^2\mathcal{N} - \Delta\mathcal{N} = 2\Delta|\mathcal{E}|^2 \end{cases} \tag{1.2}$$

1

endowed with the initial data

$$\mathcal{E}(t=0) = \mathcal{E}_0, \partial_t \mathcal{E}(t=0) = \mathcal{E}_1, \mathcal{N}(t=0) = \mathcal{N}_0, \partial_t \mathcal{N}(t=0) = \mathcal{N}_1. \tag{1.3}$$

System (1.2) is intermediate between Zakharov system (1.1) and Klein-Gordon-waves system, namely

$$\begin{cases} (\partial_t^2 - \frac{1}{\varepsilon}\Delta + \frac{1}{\varepsilon^2})E = -\frac{1}{\varepsilon}nE, \\[2mm] (\partial_t^2 - \Delta)n = \Delta|E|^2. \end{cases} \tag{1.4}$$

Indeed, because of the presence of the parameter ε in (1.4), the electric field E is highly oscillatory whereas n is low frequency. Mathematically, these oscillations can be described through the change of function $E(x,t) = \mathcal{E}(x,t)e^{-\frac{it}{\varepsilon}} + c.c.$. For the low frequency ion fluctuation we set $n(x,t) = \mathcal{N}(x,t)$. By injecting (E,n) in system (1.4) and neglecting oscillatory terms in the nonlinear term of the second equation, one obtains system (1.2). Moreover in [2], it is shown that for any initial value $(\mathcal{E}_0, \mathcal{E}_1, \mathcal{N}_0, \mathcal{N}_1) \in H^{s+2}(\mathbb{R}^d) \times H^s(\mathbb{R}^d) \times H^{s+1}(\mathbb{R}^d) \times H^s(\mathbb{R}^d)$ with s large enough and $\varepsilon|\mathcal{E}_1| \to 0$ as $\varepsilon \to 0$, there exists an unique solution to (1.2) on a maximal existence time $[0,T^\varepsilon[$ such that $T^\varepsilon > 0$. Furthermore, if (E^0, n^0) denotes the maximal solution of Zakharov equations (1.1) with initial value $(\mathcal{E}_0, \mathcal{N}_0, \mathcal{N}_1)$ defined on $[0,T_0[$, then on one hand, $\lim_{\varepsilon \to 0} \inf T^\varepsilon \geq T^0$, and on the other hand, for any $T < T_0$, $(\mathcal{E}, \mathcal{N})$ converges to (E^0, n^0) on $[0,T]$ in $\mathcal{C}([0,T]; H^{s+2}(\mathbb{R}^d) \times H^{s+1}(\mathbb{R}^d))$.

Remark 1.1. *The convergence result is easier than the one established in [4]. Indeed the high frequency term in the non-linear force $\Delta|\mathcal{E}|^2$ have been filtered out.*

In the non-compatible case (that is when $\mathcal{E}_1 \neq -\frac{i}{2}\Delta\mathcal{E}_0 + \mathcal{E}_1 + \frac{i}{\varepsilon}\mathcal{N}_0\mathcal{E}_0$), $\partial_t\mathcal{E}$ oscillates, with a $O(1)$ amplitude, at frequency $\frac{2}{\varepsilon}$ around the value $\partial_t E^0$. In the compatible case (that is when $\mathcal{E}_1 = -\frac{i}{2}\Delta\mathcal{E}_0 + \mathcal{E}_1 + \frac{i}{\varepsilon}\mathcal{N}_0\mathcal{E}_0$), the quantity $\mathcal{E} - E^0$ oscillates at frequency $\frac{2}{\varepsilon}$ and these oscillations are of order of magnitude ε. In the last section, we will consider numerically larger initial data (cases where $\partial_t\mathcal{E} = O(1/\varepsilon)$ and show how to handle them. We can divide $\mathcal{E}$ into two components F and G involving the propagators $e^{\frac{i}{\varepsilon}(1\mp\sqrt{1-\varepsilon\Delta})t}$ respectively. Analogously, we decompose $\mathcal{N}$ into $H + I$. The functions (F,G,H,I) are solution of the split-system, namely

$$\begin{cases} \partial_t F = \frac{i}{\varepsilon}(1 - \sqrt{1-\varepsilon\Delta})F + \dfrac{1}{2i\sqrt{1-\varepsilon\Delta}}(H+I)(F+G), \\[4mm] \partial_t G = \frac{i}{\varepsilon}(1 + \sqrt{1-\varepsilon\Delta})G - \dfrac{1}{2i\sqrt{1-\varepsilon\Delta}}(H+I)(F+G), \\[4mm] \partial_t H = i\sqrt{-\Delta}H + \dfrac{1}{\sqrt{\Delta}}\Delta|F+G|^2, \\[4mm] \partial_t I = -i\sqrt{-\Delta}I - \dfrac{1}{\sqrt{\Delta}}\Delta|F+G|^2. \end{cases} \tag{1.5}$$

The initial data for F,G,H,I are recovered from $\mathcal{E}_0, \mathcal{E}_1, \mathcal{N}_0, \mathcal{N}_1$ (see [2] for details). Note that this splitting is exact. Finally, system (1.2) owns the two following invariants

$$\int_{\mathbb{R}^d}|\mathcal{E}|^2 - \varepsilon\mathcal{I}m\int_{\mathbb{R}^d}\bar{\mathcal{E}}\partial_t\mathcal{E} = const \tag{1.6}$$

and

$$\int_{\mathbb{R}^d}\left(\varepsilon|\partial_t\mathcal{E}|^2 + |\nabla\mathcal{E}|^2 + \mathcal{N}|\mathcal{E}|^2 + \frac{1}{4}(\nabla\psi)^2 + \frac{1}{4}\mathcal{N}^2\right) = const \tag{1.7}$$

where ψ is defined by $\partial_t\mathcal{N} = \Delta\psi$.
• Setting of the problem : It is clear that $(\mathcal{E}, \mathcal{N})$ converges to $(F, \mathcal{N})$, as $\varepsilon \to 0$, while G is the oscillatory

part of $\mathcal{E}$. The period of these oscillations is $\pi\varepsilon$ while their amplitude is of order ε. It implies that the more G oscillates, the more the contribution of G in $\mathcal{E}$ is negligible. From the numerical point of view, it means that we need a small time step in order to compute G precisely while it is not necessary since its contribution in $E = F + G$ is small. However, for ε large enough it is essential to compute both the average and the oscillatory parts of the solution in view of the error estimate established in [4].

• Motivation : The aim of this paper is to propose a numerical approach for solving problem (1.2)-(1.3). In fact, we investigate a numerical scheme which is precise and not expensive for any ε. In particular we study two schemes of Crank-Nicolson type for (1.2) using the splitting (1.5) for the associated discrete operators. We show that the non-centered in time scheme (called S_1) owns dissipative and polarization properties contrary to the centered in time scheme (called S_2). We will precise later what is polarization. Moreover we numerically check the error estimate established in [4] and confirm that system (1.2) compute precisely the propagation of short or ultra-short waves through a plasma. For a precise description of short or ultra-short waves, one can see [1].

1.2 Numerical schemes

We recall that the Crank-Nicolson scheme for (1.1) has been introduced by Glassey [6]. It reads

$$
\begin{cases}
2i\dfrac{E_j^{m+1} - E_j^m}{\delta t} + \left(D_+D_-\left(\dfrac{E^{m+1} + E^m}{2}\right)\right)_j = \dfrac{1}{4}(n_j^{m+1} + n_j^m)(E_j^{m+1} + E_j^m), \\[4mm]
\dfrac{n_j^{m+1} - 2n_j^m + n_j^{m-1}}{\delta t^2} - \left(D_+D_-\left(\dfrac{n^{m+1} + n^{m-1}}{2}\right)\right)_j = 2(D_+D_-|E^m|^2)_j
\end{cases}
$$

where $E_j^m \approx E^0(t = m\delta t, x = x_j), n_j^m \approx n^0(t = m\delta t, x = x_j)$ and D_+, D_- are the forward and backward non-centered finite difference operators respectively. Several authors have studied numerical schemes for the Zakharov system (see for example [3], [14] and recently [7]). Analogously, we introduce the two following schemes for (1.2) :

⋆ the non-centered in time scheme

$$
\begin{cases}
\varepsilon\left(\dfrac{\mathcal{E}^{m+1} - 2\mathcal{E}^m + \mathcal{E}^{m-1}}{\delta t^2}\right) - 2i\left(\dfrac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right) - D_+D_-\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^m}{2}\right) = -\mathcal{N}^m\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^m}{2}\right), \\[4mm]
\dfrac{\mathcal{N}^{m+1} - 2\mathcal{N}^m + \mathcal{N}^{m-1}}{\delta t^2} - D_+D_-\left(\dfrac{\mathcal{N}^{m+1} + \mathcal{N}^{m-1}}{2}\right) = 2D_+D_-|\mathcal{E}^m|^2, \\[4mm]
\mathcal{E}^0 = f, \dfrac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} = g, \mathcal{N}^0 = \mathcal{N}_0, \dfrac{\mathcal{N}^0 - \mathcal{N}^{-1}}{\delta t} = \mathcal{N}_1,
\end{cases}
$$

⋆ the centered in time scheme

$$
\begin{cases}
\varepsilon\left(\dfrac{\mathcal{E}^{m+1} - 2\mathcal{E}^m + \mathcal{E}^{m-1}}{\delta t^2}\right) - i\left(\dfrac{\mathcal{E}^{m+1} - \mathcal{E}^{m-1}}{\delta t}\right) - D_+D_-\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^{m-1}}{2}\right) = -\mathcal{N}^m\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^{m-1}}{2}\right), \\[4mm]
\dfrac{\mathcal{N}^{m+1} - 2\mathcal{N}^m + \mathcal{N}^{m-1}}{\delta t^2} - D_+D_-\left(\dfrac{\mathcal{N}^{m+1} + \mathcal{N}^{m-1}}{2}\right) = 2D_+D_-|\mathcal{E}^m|^2, \\[4mm]
\dfrac{\mathcal{E}^0 + \mathcal{E}^{-1}}{2} = f, \dfrac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} = g, \mathcal{N}^0 = \mathcal{N}_0, \dfrac{\mathcal{N}^0 - \mathcal{N}^{-1}}{\delta t} = \mathcal{N}_1.
\end{cases}
$$

The aim of this paper is to provide some mathematical results for semi-discretized version of these schemes, that is when D_+D_- is replaced by the Laplace operator Δ. The schemes then read :

$\star$ the non-centered in time scheme (S_1)

$$(S_1)\begin{cases} \varepsilon\left(\dfrac{\mathcal{E}^{m+1} - 2\mathcal{E}^m + \mathcal{E}^{m-1}}{\delta t^2}\right) - 2i\left(\dfrac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right) - \Delta\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^m}{2}\right) = -\mathcal{N}^m\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^m}{2}\right), \\[4mm] \dfrac{\mathcal{N}^{m+1} - 2\mathcal{N}^m + \mathcal{N}^{m-1}}{\delta t^2} - \Delta\left(\dfrac{\mathcal{N}^{m+1} + \mathcal{N}^{m-1}}{2}\right) = 2\Delta|\mathcal{E}^m|^2, \\[4mm] \mathcal{E}^0 = f, \dfrac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} = g, \mathcal{N}^0 = \mathcal{N}_0, \dfrac{\mathcal{N}^0 - \mathcal{N}^{-1}}{\delta t} = \mathcal{N}_1, \end{cases}$$

$\star$ the centered in time scheme (S_2)

$$(S_2)\begin{cases} \varepsilon\left(\dfrac{\mathcal{E}^{m+1} - 2\mathcal{E}^m + \mathcal{E}^{m-1}}{\delta t^2}\right) - i\left(\dfrac{\mathcal{E}^{m+1} - \mathcal{E}^{m-1}}{\delta t}\right) - \Delta\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^{m-1}}{2}\right) = -\mathcal{N}^m\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^{m-1}}{2}\right), \\[4mm] \dfrac{\mathcal{N}^{m+1} - 2\mathcal{N}^m + \mathcal{N}^{m-1}}{\delta t^2} - \Delta\left(\dfrac{\mathcal{N}^{m+1} + \mathcal{N}^{m-1}}{2}\right) = 2\Delta|\mathcal{E}^m|^2, \\[4mm] \dfrac{\mathcal{E}^0 + \mathcal{E}^{-1}}{2} = f, \dfrac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} = g, \mathcal{N}^0 = \mathcal{N}_0, \dfrac{\mathcal{N}^0 - \mathcal{N}^{-1}}{\delta t} = \mathcal{N}_1. \end{cases}$$

We prove existence, uniqueness and convergence of the discrete sequences $(\mathcal{E}^m, \mathcal{N}^m)$ for both schemes for fixed ε. We show that scheme (S_2) is conservative while scheme (S_1) is dissipative. Moreover scheme (S_1) polarizes the solution whatever the initial data we take. We will precise later on what we mean by polarize.

The aim of this paper is to investigate the effect of the dissipation and the polarization phenomenon on the discrete solution given by (S_1).

The paper is organized as follows. In section 2, we state existence, uniqueness and convergence results of the discrete solution given by both schemes. We introduce the splitting of schemes (S_1) and (S_2). Then we give some details of the proof of theorem 2.1, that is related to scheme (S_1). For scheme (S_2) we just check the key points of theorem 2.2.

In section 3, we discuss the behavior of the discrete solution given by both schemes. We study dissipative and polarization phenomenon of scheme (S_1). We compare the solution of (1.2) given by both schemes with the solution of Zakharov system (1.1). At the end we illustrate the transparency property [11] of system (1.4) established in [4].

Our mathematical result mainly depend on the study of the linear part of the schemes. Therefore a lot of technical tools are imported from [5]. The main difference between this paper and [5] is the well known loss of derivative involve in the Zakharov system. This loss of derivative give rise to weaker convergence and stability results than in [5].

2 Theoretical results

Following [2] in the continuous case, we introduce an exact splitting for our schemes. Using these splitting, we prove existence, uniqueness and convergence results for the discrete solution for fixed ε.

2.1 Splitting of the schemes

We recall that in the continuous case, one divides $\mathcal{E}$ into two components F and G having the modes of propagation $e^{\frac{i}{\varepsilon}(1\mp\sqrt{1-\varepsilon\Delta})t}$ respectively. Analogously, we write $\mathcal{N} = H + I$. The functions (F, G, H, I) are

solution of the split system (1.5). One can see that G contains all the oscillatory part of the solution and the order of magnitude is of order $|G(0)|$. Moreover $(F, \mathcal{N})$ converges to the solution of (1.1) with the suitable initial data. The aim of this section is to present a similar splitting for (S_1) and (S_2) which will have the same behavior like the continuous case. In analogy with [5], we look for a splitting for (S_1) and (S_2) of the form $\mathcal{E}^m = F^m + G^m$ and $\mathcal{N}^m = H^m + I^m$ such that

$$
\begin{cases}
\dfrac{F^{m+1} - F^m}{\delta t} = i\alpha_1 \dfrac{F^{m+1} + F^m}{2} + \dfrac{\gamma_1}{\delta t}(1 + i\dfrac{\alpha_1 \delta t}{2})f^{m+1}, \\[4mm]
\dfrac{G^{m+1} - G^m}{\delta t} = i\alpha_2 \dfrac{G^{m+1} + G^m}{2} + \dfrac{\gamma_2}{\delta t}(1 + i\dfrac{\alpha_2 \delta t}{2})f^{m+1}, \\[4mm]
\dfrac{H^{m+1} - H^m}{\delta t} = i\beta_1 \dfrac{H^{m+1} + H^m}{2} + \dfrac{\nu_1}{\delta t}(1 + i\dfrac{\beta_1 \delta t}{2})h^{m}, \\[4mm]
\dfrac{I^{m+1} - I^m}{\delta t} = i\beta_2 \dfrac{I^{m+1} + I^m}{2} + \dfrac{\nu_2}{\delta t}(1 + i\dfrac{\beta_2 \delta t}{2})h^{m}
\end{cases}
\tag{2.1}
$$

where f^{m+1}, h^m denote the right-hand side of (S_1) or (S_2). For the initial data of (2.1) we refer to [5]. The operators α_j, β_j are given by $\alpha_j = \frac{2i(1-X_j)}{\delta t(1+X_j)}, \beta_j = \frac{2i(1-Y_j)}{\delta t(1+Y_j)}$ where X_j, Y_j are the roots of the characteristic equations of (S_1) and (S_2) for $j = 1, 2$. The nature of the operators α_j and β_j give us some information on the behavior of the schemes. We investigate the properties of each scheme separately.

- Splitting for (S_1) : One gets the following result (see [5] for details) :

Proposition 2.1. *It is equivalent to solve (S_1) with $\mathcal{E}_0, \mathcal{E}_1, \mathcal{N}_0, \mathcal{N}_1$ given or to solve*

$$
\begin{cases}
F^{m+1} = X_1 F^m + \gamma X_1 f^{m+1}, \\[3mm]
G^{m+1} = X_2 G^m - \gamma X_2 f^{m+1}, \\[3mm]
H^{m+1} = Y_1 H^m + \nu Y_1 h^m, \\[3mm]
I^{m+1} = Y_2 I^m - \nu Y_2 h^m
\end{cases}
\tag{2.2}
$$

with $\mathcal{E}^m = F^m + G^m, \mathcal{N}^m = H^m + I^m, f^{m+1} = -(H^m + I^m)(\frac{F^{m+1} + G^{m+1} + F^m + G^m}{2}), h^m = 2|F^m + G^m|^2$ and the suitable initial values. Operators (X_1, X_2, Y_1, Y_2) are solution of the characteristic equations

$$
X^2\left(\frac{\varepsilon}{\delta t^2} - \frac{2i}{\delta t} - \frac{\Delta}{2}\right) + X\left(-\frac{2\varepsilon}{\delta t^2} + \frac{2i}{\delta t} - \frac{\Delta}{2}\right) + \frac{\varepsilon}{\delta t^2} = 0,
$$

$$
Y^2\left(\frac{1}{\delta t^2} - \frac{\Delta}{2}\right) + Y\left(-\frac{2}{\delta t^2}\right) + \left(\frac{1}{\delta t^2} - \frac{\Delta}{2}\right) = 0.
$$

They are given by

$$
X_{1,2} = \left(-\left(-\frac{2\varepsilon}{\delta t^2} + \frac{2i}{\delta t} - \frac{\Delta}{2}\right) \pm \left(\sqrt{\frac{\lambda + \sqrt{\lambda^2 + \frac{4\Delta^2}{\delta t^2}}}{2}} + i\sqrt{\frac{-\lambda + \sqrt{\lambda^2 + \frac{4\Delta^2}{\delta t^2}}}{2}}\right)\right)\left(2\left(\frac{\varepsilon}{\delta t^2} - \frac{2i}{\delta t} - \frac{\Delta}{2}\right)\right)^{-1}
$$

where

$$
\lambda = \frac{4\varepsilon\Delta}{\delta t^2} - \frac{4}{\delta t^2} + \frac{\Delta^2}{4}
$$

and

$$Y_{1,2} = \left(1 \pm i\sqrt{\frac{\Delta^2 \delta t^4}{4} - \Delta \delta t^2}\right)\left(1 - \frac{\Delta \delta t^2}{2}\right)^{-1}.$$

Finally operators γ and ν are given by

$$\gamma = \left(\frac{\varepsilon}{\delta t^2} - \frac{2i}{\delta t} - \frac{\Delta}{2}\right)^{-1}(X_1 - X_2)^{-1} \text{ and } \nu = \left(\frac{1}{\delta t^2} - \frac{\Delta}{2}\right)^{-1}(Y_1 - Y_2)^{-1}\Delta.$$

The proof is the same than the one of proposition 4.1. in [5]. We omit it. Moreover it is clear that the Crank-Nicolson scheme (S_1) satisfies the two following approximate conservation laws :

$$\int |\mathcal{E}^{m+1}|^2 - \varepsilon \mathcal{I}m \int \left(\overline{\mathcal{E}}^{m+1}\left(\frac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right)\right) - \frac{\varepsilon \delta t}{2}\mathcal{I}m \int \sum_{k=0}^{m}\left(\frac{\mathcal{E}^{k+1} - \mathcal{E}^k}{\delta t}\right)\left(\frac{\overline{\mathcal{E}}^k - \overline{\mathcal{E}}^{k-1}}{\delta t}\right)$$

$$= \int |\mathcal{E}^0|^2 - \varepsilon \mathcal{I}m \int \left(\overline{\mathcal{E}}^0\left(\frac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t}\right)\right) \quad (2.3)$$

and

$$\varepsilon \int \left|\frac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right|^2 + \int |\nabla \mathcal{E}^{m+1}|^2 + \delta t \int \sum_{k=0}^{m}|\mathcal{E}^k|^2\left(\frac{\mathcal{N}^{k+1} - \mathcal{N}^{k-1}}{2\delta t}\right) + \delta t \int \sum_{k=0}^{m}\mathcal{N}^k\left(\frac{|\mathcal{E}^{k+1}|^2 - |\mathcal{E}^k|^2}{\delta t}\right)$$

$$+ \frac{1}{4}\int (\nabla \psi^{m+\frac{1}{2}})^2 + \frac{1}{8}\int ((\mathcal{N}^{m+1})^2 + (\mathcal{N}^m)^2) + \varepsilon \delta t^2 \int \sum_{k=0}^{m}\left|\frac{\mathcal{E}^{k+1} - 2\mathcal{E}^k + \mathcal{E}^{k-1}}{\delta t^2}\right|^2$$

$$= \varepsilon \int \left|\frac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t}\right|^2 + \int |\nabla \mathcal{E}^0|^2 + \frac{1}{4}\int (\nabla \psi^{-\frac{1}{2}})^2 + \frac{1}{8}\int ((\mathcal{N}^0)^2 + (\mathcal{N}^{-1})^2) \quad (2.4)$$

where $\Delta \psi^{m+\frac{1}{2}} = \dfrac{\mathcal{N}^{m+1} - \mathcal{N}^m}{\delta t}$ for $m \geq 0$.

- Splitting for (S_2) : We have a similar result for (S_2). The function f^{m+1} becomes $f^{m+1} = -(H^m + I^m)\left(\frac{F^{m+1} + G^{m+1} + F^{m-1} + G^{m-1}}{2}\right)$. However, the operators (X_1, X_2), which are solution of the following characteristic equation

$$X^2\left(\frac{\varepsilon}{\delta t^2} - \frac{i}{\delta t} - \frac{\Delta}{2}\right) + X\left(-\frac{2\varepsilon}{\delta t^2}\right) + \left(\frac{\varepsilon}{\delta t^2} + \frac{i}{\delta t} - \frac{\Delta}{2}\right) = 0,$$

are different from the previous scheme. They read

$$X_{1,2} = \left(\varepsilon \mp i\sqrt{\frac{\Delta^2 \delta t^4}{4} - \varepsilon \Delta \delta t^2 + \delta t^2}\right)\left(\varepsilon - \frac{\Delta \delta t^2}{2} - i\delta t\right)^{-1}.$$

The discrete conservative quantities corresponding to (1.6) and (1.7) read

$$\frac{1}{2}\int (|\mathcal{E}^{m+1}|^2 + |\mathcal{E}^m|^2) - \varepsilon \int \mathcal{I}m\left(\overline{\mathcal{E}}^{m+1}\left(\frac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right)\right)$$

$$= \frac{1}{2}\int (|\mathcal{E}^0|^2 + |\mathcal{E}^{-1}|^2) - \varepsilon \int \mathcal{I}m\left(\overline{\mathcal{E}}^0\left(\frac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t}\right)\right) \quad (2.5)$$

and

$$\varepsilon \int \left| \frac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t} \right|^2 + \frac{1}{2} \int (|\nabla \mathcal{E}^{m+1}|^2 + |\nabla \mathcal{E}^m|^2) + \frac{1}{2} \int (\mathcal{N}^m |\mathcal{E}^{m+1}|^2 + |\mathcal{E}^m|^2 \mathcal{N}^{m+1})$$

$$+ \frac{1}{4} \int (\nabla \psi^{m+\frac{1}{2}})^2 + \frac{1}{8} \int ((\mathcal{N}^{m+1})^2 + (\mathcal{N}^m)^2)$$

$$= \varepsilon \int \left| \frac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} \right|^2 + \frac{1}{2} \int (|\nabla \mathcal{E}^0|^2 + |\nabla \mathcal{E}^{-1}|^2) + \frac{1}{2} \int (|\mathcal{E}^0|^2 \mathcal{N}^{-1} + \mathcal{N}^0 |\mathcal{E}^{-1}|^2)$$

$$+ \frac{1}{4} \int (\nabla \psi^{-\frac{1}{2}})^2 + \frac{1}{8} \int ((\mathcal{N}^0)^2 + (\mathcal{N}^{-1})^2). \quad (2.6)$$

Remark 2.2. *Operators X_1 and X_2 are unitary. It implies that $i\alpha_1$ and $i\alpha_2$ are skew-adjoint and generate unitary groups on any H^s. Of course, it is not the case for (S_1).*

We are now able to state the main theoretical results of this paper.

2.2 Existence and convergence

We give the two following results related to (S_1) and (S_2) respectively.
For (S_1) one gets :

Theorem 2.1. • *Let $(F^0, G^0, H^0, I^0) \in (H^s(\mathbb{R}^d))^2 \times (H^{s-1}(\mathbb{R}^d))^2 (s > \frac{d}{2} + 1)$ and ε fixed in $]0,1[$. There exists $\delta t_1 > 0$ depending on initial data and ε such that for any $\delta t \leq \delta t_1$ there exists $T > 0$ and a unique solution $(\mathcal{E}^m, \mathcal{N}^m)$ of (S1) belonging to $l^\infty(\{0..M\}; H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d))$, where $T = M\delta t$.*
• *Let $(F^{\delta t}, G^{\delta t}, H^{\delta t}, I^{\delta t})$ be the piecewise constant functions which value at time $t \in [m\delta t, (m+1)\delta t[$ is (F^m, G^m, H^m, I^m) and (F, G, H, I) be the solution of (1.5) with the initial data $F(0) = F^0, G(0) = G^0, H(0) = H^0, I(0) = I^0$ and T_∞ its existence time, then for ε fixed in $]0,1[$ and for all $T < T_\infty$ (with $T = M\delta t$),*

$$(F^{\delta t}, G^{\delta t}, H^{\delta t}, I^{\delta t}) \to (F, G, H, I) \text{ in } L^\infty([0,T]; (H^s(\mathbb{R}^d))^2 \times (H^{s-1}(\mathbb{R}^d))^2) \text{ when } \delta t \to 0.$$

For (S_2), one gets :

Theorem 2.2. • *Let $(F^0, G^0, H^0, I^0) \in (H^s(\mathbb{R}^d))^2 \times (H^{s-1}(\mathbb{R}^d))^2 (s > \frac{d}{2} + 1)$. There exists $\delta t_2 > 0$ depending on the initial data such that for any $\delta t \leq \delta t_2$ and $\varepsilon > 0$ there exists an unique solution $(\mathcal{E}^m, \mathcal{N}^m)$ of (S_2) belonging to $l^\infty([0,T]; H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d))$ with $T = M\delta t$.*
• *Let (F, G, H, I) be the solution of (1.5) with the initial conditions $F(0) = F^0, G(0) = G^0, H(0) = H^0, I(0) = I^0$ and T_∞ its existence time. For ε fixed in $]0,1[$, one gets for all $T < T_\infty$,*

$$(F^{\delta t}, G^{\delta t}, H^{\delta t}, I^{\delta t}) \to (F, G, H, I) \text{ in } L^\infty([0,T]; (H^s(\mathbb{R}^d))^2 \times (H^{s-1}(\mathbb{R}^d))^2) \text{ when } \delta t \to 0$$

where $(F^{\delta t}, G^{\delta t}, H^{\delta t}, I^{\delta t})$ are the piecewise constant functions which value at time $t \in [m\delta t, (m+1)\delta t[$ is (F^m, G^m, H^m, I^m).

Remark 2.3. *Note that these results are not "uniform" with respect to ε. This means that one can not perform a priori both limit $\varepsilon \to 0, \delta t \to 0$ contrary to the result of [5]. This is certainly linked to the fact that, using the same kind of method in [2], the authors where not able to perform the limit $\omega \to \infty$ and $c \to \infty$ in the Klein-Gordon-Wave system, namely*

$$\begin{cases} (\frac{1}{\omega^2}\partial_t^2 - 2i\partial_t - \Delta)E = nE, \\ (\frac{1}{c^2}\partial_t^2 - \Delta)n = \Delta(|E|^2). \end{cases}$$

Note that this problem has been solved in [12] using transparency type properties also used in [4]. The transposition of such kind of properties in a semi-discrete framework is out of the scope of this paper.

2.3 Proof of theorems 2.1 and 2.2

We will give some details for the proof of theorem 2.1 related to (S_1). For theorem 2.2, we will just check the main steps.

Proof. of theorem 2.1. • We first prove existence and uniqueness of the discrete solution given by (S_1). We will use a fixed point procedure as in the continuous case. We write a Duhamel formula (2.8)-(2.9) with the four groups $\mathcal{U}_0^{\delta t}, \mathcal{U}_1^{\delta t}, \mathcal{V}_0^{\delta t}, \mathcal{V}_1^{\delta t}$ generated by the discrete equivalent of $-\frac{i}{\varepsilon}(1 \mp \sqrt{1 - \varepsilon\Delta}), \mp i\sqrt{-\Delta}$ and defined below. The main point is to obtain uniform estimates with respect to δt.

We denote by $(\mathcal{U}_0^{\delta t}(t), \mathcal{U}_1^{\delta t}(t), \mathcal{V}_0^{\delta t}(t), \mathcal{V}_1^{\delta t}(t))$ the piecewise constant operator value functions which value on $[m\delta t, (m+1)\delta t[$ is $(X_1^m, X_2^m, Y_1^m, Y_2^m)$ and $(\mathcal{E}^{\delta t}(t), \mathcal{N}^{\delta t}(t))$ the piecewise constant functions which value on $[m\delta t, (m+1)\delta t[$ is $(\mathcal{E}^m, \mathcal{N}^m)$. One has

$$X_1^{m+1} = X_1 S_0^{\delta t}(m\delta t) - \frac{1 - X_1}{\delta t} S_1^{\delta t}(m\delta t) \text{ and } X_2^{m+1} = X_2 S_0^{\delta t}(m\delta t) - \frac{1 - X_2}{\delta t} S_1^{\delta t}(m\delta t) \qquad (2.7)$$

where $S_0^{\delta t}(m\delta t)f$ is the solution of

$$(S_1^f) \begin{cases} \varepsilon\left(\dfrac{\mathcal{E}^{m+1} - 2\mathcal{E}^m + \mathcal{E}^{m-1}}{\delta t^2}\right) - 2i\left(\dfrac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right) - \Delta\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^m}{2}\right) = 0, \\[3mm] \mathcal{E}^0 = f, \dfrac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} = 0 \end{cases}$$

and $S_1^{\delta t}(m\delta t)g$ the one of

$$(S_1^g) \begin{cases} \varepsilon\left(\dfrac{\mathcal{E}^{m+1} - 2\mathcal{E}^m + \mathcal{E}^{m-1}}{\delta t^2}\right) - 2i\left(\dfrac{\mathcal{E}^{m+1} - \mathcal{E}^m}{\delta t}\right) - \Delta\left(\dfrac{\mathcal{E}^{m+1} + \mathcal{E}^m}{2}\right) = 0, \\[3mm] \mathcal{E}^0 = 0, \dfrac{\mathcal{E}^0 - \mathcal{E}^{-1}}{\delta t} = g. \end{cases}$$

The equivalent of (2.7) is

$$Y_1^{m+1} = Y_1 T_0^{\delta t}(m\delta t) - \frac{1 - Y_1}{\delta t} T_1^{\delta t}(m\delta t) \text{ and } Y_2^{m+1} = Y_2 T_0^{\delta t}(m\delta t) - \frac{1 - Y_2}{\delta t} T_1^{\delta t}(m\delta t)$$

where $T_0^{\delta t}(m\delta t)p$ is the solution of

$$\begin{cases} \dfrac{\mathcal{N}^{m+1} - 2\mathcal{N}^m + \mathcal{N}^{m-1}}{\delta t^2} - \Delta\left(\dfrac{\mathcal{N}^{m+1} + \mathcal{N}^{m-1}}{2}\right) = 0, \\[3mm] \mathcal{N}^0 = p, \dfrac{\mathcal{N}^0 - \mathcal{N}^{-1}}{\delta t} = 0 \end{cases}$$

and $T_1^{\delta t}(m\delta t)q$ the one of

$$\begin{cases} \dfrac{\mathcal{N}^{m+1} - 2\mathcal{N}^m + \mathcal{N}^{m-1}}{\delta t^2} - \Delta\left(\dfrac{\mathcal{N}^{m+1} + \mathcal{N}^{m-1}}{2}\right) = 0, \\[3mm] \mathcal{N}^0 = 0, \dfrac{\mathcal{N}^0 - \mathcal{N}^{-1}}{\delta t} = q. \end{cases}$$

In view of these notations, the discrete Duhamel formulation for (S_1) reads

$$\mathcal{E}^{\delta t}((m+1)\delta t) = \mathcal{U}_0^{\delta t}((m+1)\delta t)F^0 + \mathcal{U}_1^{\delta t}((m+1)\delta t)G^0 + \delta t \sum_{k=0}^{m}[\mathcal{U}_0^{\delta t}((k+1)\delta t) - \mathcal{U}_1^{\delta t}((k+1)\delta t)]\frac{\gamma_1}{\delta t}f^{m+1-k},$$

$$(2.8)$$

$$\mathcal{N}^{\delta t}((m+1)\delta t) = \mathcal{V}_0^{\delta t}((m+1)\delta t) + \mathcal{V}_1^{\delta t}((m+1)\delta t)I^0 + \delta t \sum_{k=0}^{m}[\mathcal{V}_0^{\delta t}((k+1)\delta t) - \mathcal{V}_1^{\delta t}((k+1)\delta t)]\frac{\nu_1}{\delta t}h^{m-k} \quad (2.9)$$

with $f^{m+1-k} = -\mathcal{N}^{m-k}\dfrac{\mathcal{E}^{m+1-k} + \mathcal{E}^{m-k}}{2}$ and $h^{m-k} = 2|\mathcal{E}^{m-k}|^2$.

We introduce the spaces $X_T = l^\infty(\{0..M\}; H^s(\mathbb{R}^d)), Y_T = l^\infty(\{0..M\}; H^{s-1}(\mathbb{R}^d))$ and B_R the ball of radius R in $X_T \times Y_T$ defined by :

$$B_R = \{(u,v) \in X_T \times Y_T / |(u,v)|_{X_T \times Y_T} := (|u|^2_{X_T} + |v|^2_{Y_T})^{\frac{1}{2}} \leq R\}$$

with $R = (6c(K)(|F^0|^2_{H^s} + |G^0|^2_{H^s}) + 6(|H^0|^2_{H^{s-1}} + |I^0|^2_{H^{s-1}}))^{\frac{1}{2}}$ where $c(K)$ is a constant to be determined later on. The time T is given by $T = M\delta t$, so that for fixed T, $\delta t \to 0$ is equivalent to $M \to +\infty$. For $(\mathcal{E}^{m+1})_{0 \leq m \leq M-1}$ and $(\mathcal{N}^{m+1})_{0 \leq m \leq M-1}$ we construct the two following series $(\mathcal{T}(\mathcal{E}^{m+1}, \mathcal{N}^{m+1}))_{0 \leq m \leq M-1}$ and $(\mathcal{S}(\mathcal{E}^{m+1}, \mathcal{N}^{m+1}))_{0 \leq m \leq M-1}$, namely

$$\mathcal{T}(\mathcal{E}^{m+1}, \mathcal{N}^{m+1}) = X_1^{m+1}F^0 + X_2^{m+1}G^0 + \delta t \sum_{k=0}^{m}(X_1^{k+1} - X_2^{k+1})\frac{\gamma_1}{\delta t}f^{m+1-k},$$

$$\mathcal{S}(\mathcal{E}^{m+1}, \mathcal{N}^{m+1}) = Y_1^{m+1}H^0 + Y_2^{m+1}I^0 + \delta t \sum_{k=0}^{m}(Y_1^{k+1} - Y_2^{k+1})\frac{\nu_1}{\delta t}h^{m-k}.$$

Finally we define the application

$$\mathcal{H} : H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d) \quad \to \quad H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$$
$$(\mathcal{E}, \mathcal{N}) \quad \longmapsto \quad (\mathcal{T}(\mathcal{E}, \mathcal{N}), \mathcal{S}(\mathcal{E}, \mathcal{N})).$$

In view of the fixed point theorem, we will show that $\mathcal{H}$ is a contraction on B_R.

In order to perform these estimate, we need to control the norm of the following operators :

• X_1^{m+1} and X_2^{m+1} : see propositions 2.4, 2.5;
• Y_1^{m+1} and Y_2^{m+1} : see proposition 2.7;
• $\gamma_1(X_1^{m+1} - X_2^{m+1})$ and $\nu_1(Y_1^{m+1} - Y_2^{m+1})$: see lemma 2.1.

The first steps are recalled from [5] :

Proposition 2.4. *Operators X_1, X_2 are bounded uniformly with respect to ε and δt in $H^s(\mathbb{R}^d)$.*

Proposition 2.5. *Let K such that $\delta t^{\frac{1}{2}}\varepsilon^{-1} \leq K$. There exists a constant $c(K)$ such that*

$$|X_1^{m+1}| + |X_2^{m+1}| \leq c(K).$$

Remark 2.6. *If $S_0^{\delta t}(m\delta t)f$ is a solution of (S_1^f), it implies that for all α and any regular function f, $\partial^{|\alpha|}S_0^{\delta t}(m\delta t)f$ is a solution of (S_1^f) too.*

For operators Y_1^{m+1} and Y_2^{m+1} we have the following result :

Proposition 2.7. *Operators Y_1 et Y_2 are unitary in $H^{s-1}(\mathbb{R}^d)$.*

Proof.: We recall that operators Y_1 and Y_2 are given by

$$Y_{1,2} = \left(1 \pm i\sqrt{\frac{\Delta^2\delta t^4}{4} - \Delta\delta t^2}\right)\left(1 - \frac{\Delta\delta t^2}{2}\right)^{-1}.$$

Computation of $|\hat{Y}_{1,2}|^2$ leads to the result of proposition 2.7. $\qquad\square$

For operators $\gamma_1(X_1^{m+1} - X_2^{m+1})$ and $\nu_1(Y_1^{m+1} - Y_2^{m+1})$ corresponding to the nonlinear terms of (2.8)-(2.9), one gets :

Lemma 2.1. *Let $f \in H^{s-1}(\mathbb{R}^d)$ and $g \in H^s(\mathbb{R}^d)$. One has*

   i) $\forall \varepsilon > 0 \ \exists c(\varepsilon)$ *such that* $|\gamma_1(X_1^{m+1} - X_2^{m+1})f|_{H^s(\mathbb{R}^d)} \leq c(\varepsilon)\delta t|f|_{H^{s-1}(\mathbb{R}^d)}$,

   ii) $|\nu_1(Y_1^{m+1} - Y_2^{m+1})g|_{H^{s-1}(\mathbb{R}^d)} \leq \delta t|g|_{H^s(\mathbb{R}^d)}$.

Proof.

   i) We recall that the operator $\gamma_1(X_1^m - X_2^m)$ is given in the Fourier space by

$$
\begin{aligned}
\hat{\gamma}_1(\hat{X}_1^{m+1} - \hat{X}_2^{m+1}) &= \left(\frac{\varepsilon}{\delta t^2} - \frac{2i}{\delta t} + \frac{|\xi|^2}{2}\right)^{-1}(\hat{X}_1 - \hat{X}_2)^{-1}(\hat{X}_1^{m+1} - \hat{X}_2^{m+1})\\
&= \left(\frac{\varepsilon}{\delta t^2} - \frac{2i}{\delta t} + \frac{|\xi|^2}{2}\right)^{-1}\mathcal{F}(S_0^{\delta t}(m\delta t)) + \left(\frac{\varepsilon}{\delta t} - 2i + \frac{|\xi|^2\delta t}{2}\right)^{-1}\mathcal{F}(S_1^{\delta t}(m\delta t))\\
&:= P + Q
\end{aligned}
$$

thanks to relation (2.7).

• Estimate of P: We get

$$
\begin{aligned}
|P|^2 = |\mathcal{F}(S_0^{\delta t}(m\delta t))|^2\left(\left(\frac{\varepsilon}{\delta t^2} + \frac{|\xi|^2}{2}\right)^2 + \frac{4}{\delta t^2}\right)^{-1} &\leq \delta t^2|\mathcal{F}(S_0^{\delta t}(m\delta t))|^2(1 + \varepsilon|\xi|^2)^{-1}\\
&\leq 2\delta t^2\left((1 + \varepsilon|\xi|^2)^{-2} + (1 + \varepsilon|\xi|^2)^{-1}\right)\\
&\leq 4\delta t^2\left(1 + \frac{1}{\varepsilon^2}\right)(1 + |\xi|^2)^{-2} + 2\delta t^2\left(1 + \frac{1}{\varepsilon}\right)(1 + |\xi|^2)^{-1}
\end{aligned}
$$

and so $|Pf|_{H^s} \leq 2(1 + \frac{1}{\varepsilon})\delta t|f|_{H^{s-2}} + \sqrt{2}(1 + \frac{1}{\sqrt{\varepsilon}})\delta t|f|_{H^{s-1}}$.

• Estimate of Q : One has

$$
\begin{aligned}
|Q|^2 = |\mathcal{F}(S_1^{\delta t}(m\delta t))|^2\left(\left(\frac{\varepsilon}{\delta t} + \frac{|\xi|^2\delta t}{2}\right)^2 + 4\right)^{-1} &\leq \delta t^2|\mathcal{F}(S_1^{\delta t}(m\delta t))|^2\varepsilon^{-2}\\
&\leq \delta t^2(1 + \varepsilon|\xi|^2)^{-1}\\
&\leq \delta t^2\left(1 + \frac{1}{\varepsilon}\right)(1 + |\xi|^2)^{-1}
\end{aligned}
$$

and therefore $|Qf|_{H^s} \leq \delta t\left(1 + \frac{1}{\sqrt{\varepsilon}}\right)|f|_{H^{s-1}}$.

By getting $c(\varepsilon) = 2(1 + \frac{1}{\varepsilon}) + (1 + \sqrt{2})(1 + \frac{1}{\sqrt{\varepsilon}})$ we establish the result *i)* of lemma 2.1.

   ii) We first give the relationship between Y_1^{m+1} et Y_2^{m+1} and $Y_1, Y_2, T_0^{\delta t}, T_1^{\delta t}$ where $T_0^{\delta t}(m\delta t)p$ and $T_1^{\delta t}(m\delta t)q$ are the groups corresponding to $H + I$ introduced in section 1. As above, one gets

$$
Y_1^{m+1} = Y_1 T_0^{\delta t}(m\delta t) - \frac{1 - Y_1}{\delta t}T_1^{\delta t}(m\delta t) \text{ and } Y_2^{m+1} = Y_2 T_0^{\delta t}(m\delta t) - \frac{1 - Y_2}{\delta t}T_1^{\delta t}(m\delta t).
$$

In view of these previous formulas, operator $\hat{\nu}_1(\hat{Y}_1^{m+1} - \hat{Y}_2^{m+1})$ is given by

$$
\hat{\nu}_1(\hat{Y}_1^{m+1} - \hat{Y}_2^{m+1}) = -|\xi|^2\mathcal{F}(T_0^{\delta t}(m\delta t))\left(\frac{1}{\delta t^2} + \frac{|\xi|^2}{2}\right)^{-1} - |\xi|^2\mathcal{F}(T_1^{\delta t}(m\delta t))\left(\frac{1}{\delta t} + \frac{|\xi|^2\delta t}{2}\right)^{-1}.
$$

10

So we get

$$
\begin{aligned}
|\hat{\nu}_1(\hat{Y}_1^{m+1} - \hat{Y}_2^{m+1})| \;\leq\; & |\xi|^2|\mathcal{F}(T_0^{\delta t}(m\delta t))|\delta t^2(1 + \frac{|\xi|^2\delta t^2}{2})^{-1} + |\xi|^2|\mathcal{F}(T_1^{\delta t}(m\delta t))|\delta t(1 + \frac{|\xi|^2\delta t^2}{2})^{-1} \\
\leq\; & |\xi|^2\delta t^2(1 + \frac{|\xi|^2\delta t^2}{2})^{-1} + |\xi|\delta t(1 + \frac{|\xi|^2\delta t^2}{2})^{-1} \\
\leq\; & |\xi|\delta t(\tfrac{1}{\sqrt{2}} + 1)
\end{aligned}
$$

thanks to the expression of $\mathcal{F}(T_0^{\delta t}(m\delta t))$ and $\mathcal{F}(T_1^{\delta t}(m\delta t))$. This leads to result $ii)$ of lemma 2.1.

In view of the fixed point procedure and the previous estimates of the operators that occur in the Duhamel formulation (2.8)-(2.9), the local existence of the solution of (S_1) in B_R is straightforward. Note the well-known loss of derivative encountered in the Zakharov system is handled by i) of lemma 2.1. The consequence is a loss of uniformity with respect to ε.

For the convergence of the solution of (S_1) towards the continuous solution of (1.2)-(1.3), we write the Duhamel formula for $(\mathcal{E}(t), \mathcal{N}(t))$ solution of (1.2)-(1.3) between 0 and $(m+1)\delta t$:

$$
\mathcal{E}((m+1)\delta t) = \mathcal{U}_0((m+1)\delta t)F^0 + \mathcal{U}_1((m+1)\delta t)G^0
$$

$$
- \int_0^{(m+1)\delta t} \frac{[\mathcal{U}_0((m+1)\delta t - \tau) - \mathcal{U}_1((m+1)\delta t - \tau)]}{2i\sqrt{1 - \varepsilon\Delta}} \mathcal{N}(\tau)\mathcal{E}(\tau)d\tau \quad (2.10)
$$

and

$$
\mathcal{N}((m+1)\delta t) = \mathcal{V}_0((m+1)\delta t)H^0 + \mathcal{V}_1((m+1)\delta t)I^0
$$

$$
+ 2\int_0^{(m+1)\delta t} \frac{[\mathcal{V}_0((m+1)\delta t - \tau) - \mathcal{V}_1((m+1)\delta t - \tau)]}{2i\sqrt{-\Delta}} \Delta|\mathcal{E}(\tau)|^2 d\tau. \quad (2.11)
$$

Substracting these identities from (2.8)-(2.9) respectively and taking the H^s and H^{s-1} norms leads to

$$
\begin{aligned}
|\mathcal{E}((m+1)\delta t) \;-\; & \mathcal{E}^{\delta t}((m+1)\delta t)|_{H^s} \leq |[\mathcal{U}_0((m+1)\delta t) - \mathcal{U}_0^{\delta t}((m+1)\delta t)]F^0|_{H^s} \\
+\; & |[\mathcal{U}_1((m+1)\delta t) - \mathcal{U}_1^{\delta t}((m+1)\delta t)]G^0|_{H^s} \\
+\; & \sum_{k=0}^m \Big| \int_{k\delta t}^{(k+1)\delta t} \frac{\mathcal{U}_0((m+1)\delta t - \tau) - \mathcal{U}_1((m+1)\delta t - \tau)}{2i\sqrt{1 - \varepsilon\Delta}} \mathcal{N}(\tau)\mathcal{E}(\tau)d\tau \\
-\; & \int_{k\delta t}^{(k+1)\delta t} \Big[\mathcal{U}_0^{\delta t}((m+1)\delta t - k\delta t) - \mathcal{U}_1^{\delta t}((m+1)\delta t - k\delta t)\Big] \frac{\gamma_1}{\delta t} \mathcal{N}^{\delta t}(k\delta t) \\
\times\; & \frac{\mathcal{E}^{\delta t}((k+1)\delta t) + \mathcal{E}^{\delta t}(k\delta t)}{2} d\tau \Big|_{H^s}
\end{aligned}
$$

$$(2.12)$$

11

and

$$
\begin{aligned}
|\mathcal{N}((m+1)\delta t) \quad - \quad & \mathcal{N}^{\delta t}((m+1)\delta t)|_{H^{s-1}} \leq |[\mathcal{V}_0((m+1)\delta t) - \mathcal{V}_0^{\delta t}((m+1)\delta t)]H^0|_{H^{s-1}} \\
+ \quad & |[\mathcal{V}_1((m+1)\delta t) - \mathcal{V}_1^{\delta t}((m+1)\delta t)]I^0|_{H^{s-1}} \\
+ \quad & 2\sum_{k=0}^{m}\left|\int_{k\delta t}^{(k+1)\delta t} \frac{\mathcal{V}_0((m+1)\delta t - \tau) - \mathcal{V}_1((m+1)\delta t - \tau)}{2i\sqrt{-\Delta}}\Delta|\mathcal{E}(\tau)|^2 d\tau \right. \\
- \quad & \left.\int_{k\delta t}^{(k+1)\delta t}[\mathcal{V}_0^{\delta t}((m+1)\delta t - k\delta t) - \mathcal{V}_1^{\delta t}((m+1)\delta t - k\delta t)]\frac{\nu_1}{\delta t}|\mathcal{E}^{\delta t}(k\delta t)|^2 d\tau\right|_{H^{s-1}}.
\end{aligned}
\tag{2.13}
$$

In order to control the r.h.s. of (2.12), we decompose the integral into three terms, namely

$$
\left|\int_{k\delta t}^{(k+1)\delta t}\left(\frac{\mathcal{U}_0((m+1)\delta t - k\delta t) - \mathcal{U}_1((m+1)\delta t - k\delta t)}{2i\sqrt{1 - \varepsilon\Delta}}\right.\right.
$$

$$
\left.\left. -(\mathcal{U}_0^{\delta t}((m+1)\delta t - k\delta t) - \mathcal{U}_1^{\delta t}((m+1)\delta t - k\delta t))\frac{\gamma_1}{\delta t}\right)\mathcal{N}^{\delta t}(k\delta t)\frac{\mathcal{E}^{\delta t}((k+1)\delta t) + \mathcal{E}^{\delta t}(k\delta t)}{2}d\tau\right|_{H^s}
$$

$$
+\left|\int_{k\delta t}^{(k+1)\delta t}\left(\frac{\mathcal{U}_0((m+1)\delta t - \tau) - \mathcal{U}_1((m+1)\delta t - \tau)}{2i\sqrt{1 - \varepsilon\Delta}}\right.\right.
$$

$$
\left.\left. -\frac{\mathcal{U}_0((m+1)\delta t - k\delta t) - \mathcal{U}_1((m+1)\delta t - k\delta t)}{2i\sqrt{1 - \varepsilon\Delta}}\right)\mathcal{N}(\tau)\mathcal{E}(\tau)d\tau\right|_{H^s}
$$

$$
+\left|\int_{k\delta t}^{(k+1)\delta t}\frac{\mathcal{U}_0((m+1)\delta t - k\delta t) - \mathcal{U}_1((m+1)\delta t - k\delta t)}{2i\sqrt{1 - \varepsilon\Delta}}\left(\mathcal{N}(\tau)\mathcal{E}(\tau)\right.\right.
$$

$$
\left.\left. -\mathcal{N}^{\delta t}(k\delta t)\frac{\mathcal{E}^{\delta t}((k+1)\delta t) + \mathcal{E}^{\delta t}(k\delta t)}{2}\right)d\tau\right|_{H^s}
$$

$$
:= I_1^{k,m} + I_2^{k,m} + I_3^{k,m}.
$$

We decompose the integral term of (2.13) into three similar terms as

$$
\left| \int_{k\delta t}^{(k+1)\delta t} \left(\frac{\mathcal{V}_0((m+1)\delta t - k\delta t) - \mathcal{V}_1((m+1)\delta t - k\delta t)}{2i\sqrt{-\Delta}} \Delta \right. \right.
$$

$$
\left. \left. - (\mathcal{V}_0^{\delta t}((m+1)\delta t - k\delta t) - \mathcal{V}_1^{\delta t}((m+1)\delta t - k\delta t)) \frac{\nu_1}{\delta t} \right) |\mathcal{E}^{\delta t}(k\delta t)|^2 d\tau \right|_{H^{s-1}}
$$

$$
+ \left| \int_{k\delta t}^{(k+1)\delta t} \left(\frac{\mathcal{V}_0((m+1)\delta t - \tau) - \mathcal{V}_1((m+1)\delta t - \tau)}{2i\sqrt{-\Delta}} \right. \right.
$$

$$
\left. \left. - \frac{\mathcal{V}_0((m+1)\delta t - k\delta t) - \mathcal{V}_1((m+1)\delta t - k\delta t)}{2i\sqrt{-\Delta}} \right) \Delta |\mathcal{E}(\tau)|^2 d\tau \right|_{H^{s-1}}
$$

$$
+ \left| \int_{k\delta t}^{(k+1)\delta t} \frac{\mathcal{V}_0((m+1)\delta t - k\delta t) - \mathcal{V}_1((m+1)\delta t - k\delta t)}{2i\sqrt{-\Delta}} \Delta \left(|\mathcal{E}(\tau)|^2 - |\mathcal{E}^{\delta t}(k\delta t)|^2 \right) d\tau \right|_{H^{s-1}}
$$

$$
:= J_1^{k,m} + J_2^{k,m} + J_3^{k,m}.
$$

In view of these previous notations, inequalities (2.12) and (2.13) are bounded by

$$
|\mathcal{E} - \mathcal{E}^{\delta t}|_{X_T} \le |(\mathcal{U}_0 - \mathcal{U}_0^{\delta t})F^0|_{X_T} + |(\mathcal{U}_1 - \mathcal{U}_1^{\delta t})G^0|_{X_T} + \sup_{m \in \{0..M-1\}} \sum_{k=0}^{m} I_1^{k,m} + I_2^{k,m} + I_3^{k,m} \quad (2.14)
$$

and

$$
|\mathcal{N} - \mathcal{N}^{\delta t}|_{Y_T} \le |(\mathcal{V}_0 - \mathcal{V}_0^{\delta t})H^0|_{Y_T} + |(\mathcal{V}_1 - \mathcal{V}_1^{\delta t})I^0|_{Y_T} + 2 \sup_{m \in \{0..M-1\}} \sum_{k=0}^{m} J_1^{k,m} + J_2^{k,m} + J_3^{k,m}. \quad (2.15)
$$

Since $\mathcal{U}_{0,1}^{\delta t} \to \mathcal{U}_{0,1}$ and $\mathcal{V}_{0,1}^{\delta t} \to \mathcal{V}_{0,1}$, one has

$$
\lim_{\delta t \to 0} |(\mathcal{U}_0 - \mathcal{U}_0^{\delta t})F^0|_{X_T} = 0, \lim_{\delta t \to 0} |(\mathcal{U}_1 - \mathcal{U}_1^{\delta t})G^0|_{X_T} = 0 \text{ as } \delta t \to 0 \quad (2.16)
$$

and

$$
\lim_{\delta t \to 0} |(\mathcal{V}_0 - \mathcal{V}_0^{\delta t})H^0|_{Y_T} = 0, \lim_{\delta t \to 0} |(\mathcal{V}_1 - \mathcal{V}_1^{\delta t})I^0|_{Y_T} = 0 \text{ as } \delta t \to 0. \quad (2.17)
$$

In the same way, as $\dfrac{\gamma_1}{\delta t} \to \dfrac{1}{2i\sqrt{1-\varepsilon\Delta}}$ and $\dfrac{\nu_1}{\delta t} \to \dfrac{\Delta}{2i\sqrt{-\Delta}}$, one gets

$$
\sup_{m \in \{0..M-1\}} \sum_{k=0}^{m} I_1^{k,m} \to 0 \text{ and } \sup_{m \in \{0..M-1\}} \sum_{k=0}^{m} J_1^{k,m} \to 0 \text{ as } \delta t \to 0. \quad (2.18)
$$

On the other hand, as the groups $\mathcal{U}_{0,1}$ and $\mathcal{V}_{0,1}$ are strongly continuous,

$$
\sup_{m \in \{0..M-1\}} \sum_{k=0}^{m} I_2^{k,m} \to 0 \text{ and } \sup_{m \in \{0..M-1\}} \sum_{k=0}^{m} J_2^{k,m} \to 0 \text{ as } \delta t \to 0. \quad (2.19)
$$

For $I_3^{k,m}, J_3^{k,m}$, let

$$
\mathcal{N}(\tau)\mathcal{E}(\tau) = \mathcal{N}(k\delta t) \frac{\mathcal{E}((k+1)\delta t) + \mathcal{E}(k\delta t)}{2} + R_{1,k}(\tau),
$$

$$
|\mathcal{E}(\tau)|^2 = |\mathcal{E}(k\delta t)|^2 + R_{2,k}(\tau)
$$

13

so that

$$I_3^{k,m} \leq \left(1 + \frac{1}{\varepsilon}\right)\delta t \left|\mathcal{N}(k\delta t)\frac{\mathcal{E}((k+1)\delta t) + \mathcal{E}(k\delta t)}{2} - \mathcal{N}^{\delta t}(k\delta t)\frac{\mathcal{E}^{\delta t}((k+1)\delta t) + \mathcal{E}^{\delta t}(k\delta t)}{2}\right|_{H^{s-1}}$$

$$+ \left(1 + \frac{1}{\varepsilon}\right)\delta t \sum_{k=0}^{M-1} \sup_{\tau \in [k\delta t, (k+1)\delta t]} |R_{1,k}(\tau)|_{H^{s-1}}, \quad (2.20)$$

$$J_3^{k,m} \leq \delta t \left||\mathcal{E}(k\delta t)|^2 - |\mathcal{E}^{\delta t}(k\delta t)|^2\right|_{H^s} + \delta t \sum_{k=0}^{M-1} \sup_{\tau \in [k\delta t, (k+1)\delta t]} |R_{2,k}(\tau)|_{H^s}. \quad (2.21)$$

Moreover, $(\mathcal{E}, \mathcal{N})$ lies in $\mathcal{C}([0,T]; H^s(\mathbb{R}^d)) \times \mathcal{C}([0,T]; H^{s-1}(\mathbb{R}^d))$ for $s > \frac{d}{2} + 1$, therefore

$$\left(1 + \frac{1}{\varepsilon}\right)\delta t \sum_{k=0}^{M-1} \sup_{\tau \in [k\delta t, (k+1)\delta t]} |R_{1,k}(\tau)|_{H^{s-1}} \to 0 \text{ and } \delta t \sum_{k=0}^{M-1} \sup_{\tau \in [k\delta t, (k+1)\delta t]} |R_{2,k}(\tau)|_{H^s} \to 0 \text{ as } \delta t \to 0.$$

For the first terms in the r.h.s. of (2.20) and (2.21), we write

$$\left|\mathcal{N}(k\delta t)\frac{\mathcal{E}((k+1)\delta t)+\mathcal{E}(k\delta t)}{2} - \mathcal{N}^{\delta t}(k\delta t)\frac{\mathcal{E}^{\delta t}((k+1)\delta t)+\mathcal{E}^{\delta t}(k\delta t)}{2}\right|_{H^{s-1}} = \left|(\mathcal{N}(k\delta t) - \mathcal{N}^{\delta t}(k\delta t))\frac{\mathcal{E}((k+1)\delta t)+\mathcal{E}(k\delta t)}{2}\right.$$

$$\left.+\mathcal{N}^{\delta t}(k\delta t)\left(\frac{\mathcal{E}((k+1)\delta t) + \mathcal{E}(k\delta t)}{2} - \frac{\mathcal{E}^{\delta t}((k+1)\delta t) + \mathcal{E}^{\delta t}(k\delta t)}{2}\right)\right|_{H^{s-1}}$$

and

$$\left||\mathcal{E}(k\delta t)|^2 - |\mathcal{E}^{\delta t}(k\delta t)|^2\right|_{H^s} = \left|(\mathcal{E}(k\delta t) - \mathcal{E}^{\delta t}(k\delta t))\overline{\mathcal{E}(k\delta t)} + \mathcal{E}^{\delta t}(k\delta t)(\overline{\mathcal{E}(k\delta t)} - \overline{\mathcal{E}^{\delta t}(k\delta t)})\right|_{H^s}.$$

We recall that $H^{s-1}(\mathbb{R}^d)$ and $H^s(\mathbb{R}^d)$ are algebras for $s > \frac{d}{2} + 1$. Now according to the uniform estimates of $(\mathcal{E}, \mathcal{E}^{\delta t})$ and $(\mathcal{N}, \mathcal{N}^{\delta t})$ on $[0, M\delta t]$, one has

$$I_3^{k,m} \leq \left(1 + \tfrac{1}{\varepsilon}\right)C\delta t|\mathcal{N} - \mathcal{N}^{\delta t}|_{Y_T} + \left(1 + \tfrac{1}{\varepsilon}\right)C\delta t|\mathcal{E} - \mathcal{E}^{\delta t}|_{X_T} + o(1),$$

$$J_3^{k,m} \leq 2M\delta t C|\mathcal{E} - \mathcal{E}^{\delta t}|_{X_T} + o(1) \ .$$

Finally, inequalities (2.14)-(2.15) lead to

$$|\mathcal{E} - \mathcal{E}^{\delta t}|_{X_T} + |\mathcal{N} - \mathcal{N}^{\delta t}|_{Y_T} \leq o(1) + C(\varepsilon)M\delta t(|\mathcal{E} - \mathcal{E}^{\delta t}|_{X_T} + |\mathcal{N} - \mathcal{N}^{\delta t}|_{Y_T})$$

with $C(\varepsilon) = 2\left(1 + \frac{1}{\varepsilon}\right)C$. Taking $0 \leq \delta t_2 < (C(\varepsilon)M)^{-1}$ leads to

$$|\mathcal{E} - \mathcal{E}^{\delta t}|_{X_T} + |\mathcal{N} - \mathcal{N}^{\delta t}|_{Y_T} \leq o(1).$$

It is clear that $(\mathcal{E}^{\delta t}, \mathcal{N}^{\delta t})$ belong to $L^\infty([0,T]; H^s(\mathbb{R}^d)) \times L^\infty([0,T]; H^{s-1}(\mathbb{R}^d))$ with $T = M\delta t$. This finish the proof of theorem 2.1.

$\square$

For scheme (S_2), in view of the previous results, the proof of the first point of theorem 2.2 is based on the following proposition :

Proposition 2.8. *Operators X_1 and X_2 corresponding to scheme (S_2) are unitary on $H^s(\mathbb{R}^d)$.*

Proof. Thanks to the formulas of X_1 and X_2, it is clear that the groups $X_1^{\frac{T}{\delta t}}$ and $X_2^{\frac{T}{\delta t}}$ are bounded uniformly with respect to δt and ε. $\square$

For the convergence result of theorem 2.2, one has $X_{1,2}^{\frac{T}{\delta t}} \to e^{\frac{iT}{\varepsilon}(1\mp\sqrt{1-\varepsilon\Delta})}$ and $\frac{\gamma_1}{\delta t} \to -\frac{1}{2i\sqrt{1-\varepsilon\Delta}}$ as $\delta t \to 0$. The same arguments than the ones used for theorem 2.1 conclude the proof. $\square$

3 Numerical results and comments

3.1 Diffusion and oscillations

The aim of this part is to investigate numerically the behavior for the Zakharov-wave system. We study the behavior of the solution of each scheme (S_1) and (S_2) when ε is small enough. Let us first recall the splitting for the continuous problem (1.5) (we do not rewrite the two last equations which do not depend on ε) :

$$\begin{cases} \partial_t F = \dfrac{i}{\varepsilon}(1 - \sqrt{1 - \varepsilon\Delta})F + \dfrac{1}{2i\sqrt{1 - \varepsilon\Delta}}(H + I)(F + G), \\[3mm] \partial_t G = \dfrac{i}{\varepsilon}(1 + \sqrt{1 - \varepsilon\Delta})G - \dfrac{1}{2i\sqrt{1 - \varepsilon\Delta}}(H + I)(F + G). \end{cases}$$

It is clear that

$$e^{\frac{i}{\varepsilon}(1 - \sqrt{1-\varepsilon\Delta})t} \longrightarrow e^{i\frac{\Delta t}{2}}, \; e^{\frac{i}{\varepsilon}(1 + \sqrt{1-\varepsilon\Delta})t} - e^{\frac{2it}{\varepsilon}}e^{-i\frac{\Delta t}{2}} \longrightarrow 0 \text{ when } \varepsilon \to 0.$$

It implies that $(F, H + I)$ converge to the solution of Zakharov system (1.1) whereas G is an oscillatory part of $\mathcal{E}$, solution of (1.2). Moreover the amplitude of these oscillations are of order $|G(0)|_{H^s}$ while the frequency is $\frac{2}{\varepsilon}$. The question is : how the discrete solution of each scheme behaves? To answer this question, we recall the discrete splitting of each scheme (2.1) (we rewrite only the two first equations)

$$\begin{cases} \dfrac{F^{m+1} - F^m}{\delta t} = i\alpha_1 \dfrac{F^{m+1} + F^m}{2} + \dfrac{\gamma_1}{\delta t}(1 + i\dfrac{\alpha_1\delta t}{2})f^{m+1}, \\[3mm] \dfrac{G^{m+1} - G^m}{\delta t} = i\alpha_2 \dfrac{G^{m+1} + G^m}{2} + \dfrac{\gamma_2}{\delta t}(1 + i\dfrac{\alpha_2\delta t}{2})f^{m+1} \end{cases} \tag{3.1}$$

where $\alpha_j = \frac{2i(1-X_j)}{\delta t(1+X_j)}$ for $j = 1, 2$. We do the asymptotic expansion of operators X_1 and X_2. We investigate the discrete solution given by (S_1) for two values of ε. If δt is small enough with respect to ε, for example $\varepsilon = \delta t^{1/2}$, one gets at the first order, when $\delta t \to 0$,

$$X_1 = 1 + O(\delta t) \text{ and } X_2 = 1 + 2i\sqrt{\delta t} - 4\delta t - \frac{i\Delta}{2}\delta t + O(\delta t^2).$$

It follows that the leading part of G^m can be written as $G^M = e^{\frac{2iT}{\varepsilon}}e^{-\frac{i\Delta T}{\varepsilon}}e^{-4T}G(0)$ at time $T = M\delta t$. It means that G^m oscillates at frequency $\frac{2}{\varepsilon}$ and the amplitude of these oscillations damped exponentially in time for regular initial data $G(0)$.

On the other hand, if δt is "large" with respect to ε, for example when $\varepsilon = \delta t$, one has

$$X_1 = 1 + O(\delta t) \text{ and } X_2 = \frac{1 + 2i}{5} + O(\delta t).$$

The leading part of G^m becomes $G^M = \sqrt{5}^{-\frac{T}{\varepsilon}}e^{i\arctan(2)\frac{T}{\varepsilon}}G(0)$ at time $T = M\delta t$. The amplitude of oscillations of G^m are damped in exponentially in time even if their frequency is not the good one. In summary, when $\varepsilon \to 0$, the discrete solution of (S_1) behaves like the continuous one.

For scheme $(S2)$, one has estimates for the linear part that are independent of ε see [5]. If we take as in the preceding case $\varepsilon = \delta t$, one gets $1 - \hat{X}_{1,2}, 1 + \hat{X}_{1,2}$ as $\delta t \to 0$ with $\hat{X}_{1,2} = (1 \mp i\sqrt{1 + |\xi|^2\delta t + \frac{|\xi|^4\delta t^2}{4}})(1 - i + \frac{|\xi|^2\delta t}{2})^{-1}$ and so

$$\alpha_1 \sim -\frac{|\xi|^2}{2}, \alpha_2 \sim \frac{2}{\varepsilon},$$

15

this means that the solution behaves like the continuous case. In other words, $(F, H + I)$ converge to the solution of Zakharov system (1.1) whereas G oscillates at frequency $\frac{2}{\varepsilon}$. Moreover its amplitude is of order $|G(0)|_{H^s}$ which is taken small.

We do some numerical experiments in order to illustrate the previous comments. We compute solution to (S_1) and (S_2) in 1D and we set in the periodic case. We use a finite difference discretization of $-\Delta$. We compute on the space-time interval $[-5, 5] \times [0, 2.10^{-3}]$ with 1024 points in space. We take $f = 4e^{-20x^2}, \mathcal{N}_0 = 0, \mathcal{N}_1 = 0$ and $\varepsilon = 10^{-5}$. For g we choose two different values :

- the first one, compared to the compatible case (that is $\partial_t \mathcal{E}(0) = \frac{i}{2}(\Delta \mathcal{E}(0) - \mathcal{N}(0)\mathcal{E}(0)))$, induces $g = 40f(40x^2 - 1)$.

- for the second one, we choose an arbitrary value independent of ε, for example $g = e^{-20x^2}$.

In the two following figures, we represent the L^2 norm of $\mathcal{E}^{\delta t}$ according to time given by

- scheme (S_1) for two values of δt that are large with respect to ε (figure 1);

- scheme (S_2), in the compatible case (left curve) and in the non-compatible case (right curve) for $\delta t = 10^{-6}$ (figure 2).
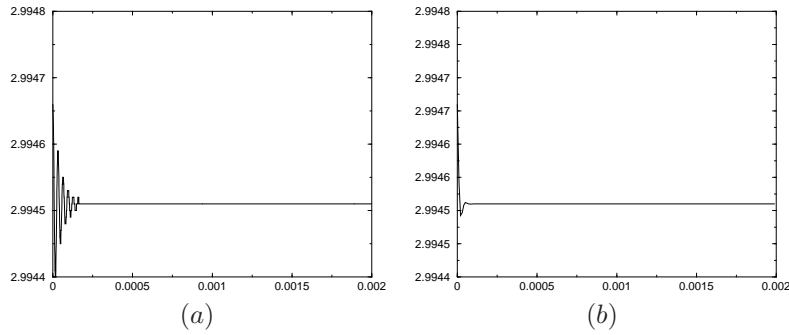

Figure 1: scheme (S_1), $\varepsilon = 10^{-5}$. $(a)|\mathcal{E}^{\delta t}|_{L^2}(t)$ for $\delta t = 10^{-6}$; $(b)|\mathcal{E}^{\delta t}|_{L^2}(t)$ for $\delta t = 10^{-5}$.

In figure 1, we notice that the oscillations are damped when the time step increases : this is the dissipative phenomenon described above. In figure 2, the only difference between the two curves concerns the amplitude of the oscillations. Indeed, in the non compatible case, the amplitude is of order ε contrary to the compatible case where they are much smaller.

3.2 Polarization

In this part, we are going to show that scheme (S_1) is able to project the numerical solution on a polarized functions space. More precisely, we will compute the solution of the Cauchy problem (1.2) with "ill-prepared" initial data. Then we will show that these solutions converge, as ε tends to zero, to the solution of the Cauchy problem (1.2) with polarized initial data (which is asymptotically the same than the one of Zakharov system (1.1)). For a precise description of this type of initial data, see below.
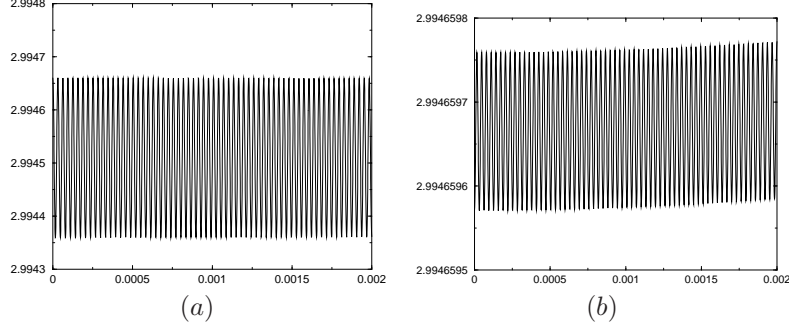
Figure 2: scheme (S_2), $\varepsilon = 10^{-5}$, $\delta t = 10^{-6}$. $(a)|\mathcal{E}^{\delta t}|_{L^2}(t)$: non-compatible case; $(b)|\mathcal{E}^{\delta t}|_{L^2}(t)$: compatible case.

3.2.1 The Klein-Gordon-Wave system

We use several notions introduced in [8], [9], [10]. Let us first recall Klein-Gordon-wave system (1.4) in 1D :

$$\begin{cases} \partial_t^2 E - \frac{1}{\varepsilon}\partial_x^2 E + \frac{1}{\varepsilon^2}E = -\frac{1}{\varepsilon}nE, \\ (\partial_t^2 - \partial_x^2)n = \partial_x^2|E|^2, \\ E(t=0) = E_0, \partial_t E(t=0) = E_1, n(t=0) = n_0, \partial_t n(t=0) = n_1. \end{cases} \tag{3.2}$$

Introducing $\varphi = \varepsilon\partial_t E, \psi = \sqrt{\varepsilon}\partial_x E$ and $U = (E, \varphi, \psi)$ such that (U, n) be a solution of the Cauchy problem for the following system :

$$\begin{cases} \partial_t U + \frac{1}{\sqrt{\varepsilon}}A_0(\partial_x)U + \frac{L_0}{\varepsilon}U = F(n, U), \\ (\partial_t^2 - \partial_x^2)n = \partial_x^2 G(U, U), \\ U(t=0) = U_0, n(t=0) = n_0, \partial_t n(t=0) = n_1 \end{cases} \tag{3.3}$$

with

$$A_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \quad L_0 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, F(n, U) = (nE, 0, 0) \text{ and } G(U, U) = |E|^2.$$

Then we expand $U(t, x)$ under the form

$$U(t, x) = (\mathcal{U}^0(t, x) + \varepsilon^{\frac{1}{2}}\mathcal{U}^1(t, x) + O(\varepsilon^{\frac{1}{2}}))e^{i\theta} + c.c. \tag{3.4}$$

with $\theta = -\frac{t}{\varepsilon}$ and where c.c. denotes the complex conjugate. The quantity n is unchanged.

Remark 3.1. *The justification of these expansion has been established in [4].*

Then using (3.4) in (3.3), we obtain the following relation at order ε^{-1}

$$-i\mathcal{U}^0 + L_0\mathcal{U}^0 = 0. \tag{3.5}$$

17

The three eigenvalues of L_0 are $\lambda_0 = 0$ and $\lambda_\pm = \pm i$. The associated eigenvectors are generated by $e_0 = (0,0,1), e_\pm = (1, \mp i, 0)$. Moreover we denote by $\Pi_0, \Pi_\pm$ the spectral projectors corresponding to the eigenvalues $\lambda_0, \lambda_\pm$ respectively. They are given by $\Pi_0 = (L_0 + i)(L_0 - i), \Pi_\pm = -\frac{L_0(L_0 \pm i)}{2}$. The non-trivial solutions of (3.5) satisfy the polarization condition $\mathcal{U}^0 = \Pi_+ \mathcal{U}^0$. If the polarization condition is satisfied at $t = 0$, the initial condition U_0 of (3.3) is called "polarized initial condition" and satisfies $U_0(x) = \Pi^+ \mathcal{U}^0(t = 0, x) + c.c.$. On the contrary, when $U_0(x) \neq \Pi^+ \mathcal{U}^0(t = 0, x) + c.c.$, the initial condition is called "ill-prepared initial condition" or "non polarized initial condition". For example, $U_0(x) = \alpha_+ e_+ + \alpha_- e_- + c.c.$ with $(\alpha_+, \alpha_-) \in \mathbb{C}^2$ such that $|\alpha_-| \neq 0$ is an ill-prepared initial condition. The question is : how can we express the polarization condition on system (3.6), namely

$$
\begin{cases}
\varepsilon \partial_t^2 \mathcal{E} - 2i \partial_t \mathcal{E} - \partial_x^2 \mathcal{E} = -\mathcal{N}\mathcal{E}, \\[2mm]
\partial_t^2 \mathcal{N} - \partial_x^2 \mathcal{N} = 2 \partial_x^2 |\mathcal{E}|^2, \\[2mm]
\mathcal{E}(t = 0) = f, \partial_t \mathcal{E}(t = 0) = g, \mathcal{N}(t = 0) = \mathcal{N}_0, \partial_t \mathcal{N}(t = 0) = \mathcal{N}_1.
\end{cases}
\tag{3.6}
$$

We recall that system (3.6) is obtained from (3.2) by getting $E(t, x) = \mathcal{E}(t, x)e^{-i\frac{t}{\varepsilon}} + c.c.$ and $n(t, x) = \mathcal{N}(t, x)$ (for a rigorous justification, one can see [4]). Therefore letting

$$
\begin{aligned}
U = (E, \varphi, \psi) &= (E, \varepsilon \partial_t E, \sqrt{\varepsilon} \partial_x E) \\[2mm]
&= (\mathcal{E}, \varepsilon \partial_t \mathcal{E} - i\mathcal{E}, \sqrt{\varepsilon} \partial_x \mathcal{E}) e^{-i\frac{t}{\varepsilon}} + c.c.
\end{aligned}
$$

and

$$
U_0 = (f, \varepsilon g - if, \sqrt{\varepsilon} \partial_x f) + c.c.
$$

- Case of polarized initial data : one has $U_0 = \alpha_+ e_+ + c.c.$ with $\alpha_+ \in \mathbb{C}$. It implies that, at the first order,
$$
\alpha_+ = f, \varepsilon g = 0, \sqrt{\varepsilon} \partial_x f = 0.
$$

In summary, any initial data which read

$$
\mathcal{E}(t = 0) = f, \partial_t \mathcal{E}(t = 0) = g, \mathcal{N}(t = 0) = \mathcal{N}_0, \partial_t \mathcal{N}(t = 0) = \mathcal{N}_1
$$

with g bounded uniformly with respect to ε such that $\lim_{\varepsilon \to 0} \varepsilon g = 0$ are polarized. The associated initial condition of Zakharov system (1.1) reads

$$
E^0(t = 0) = f, n^0(t = 0) = \mathcal{N}_0, \partial_t n^0(t = 0) = \mathcal{N}_1.
$$

Remark 3.2. *Due to the expressions of $G(0)$ one has $|G(0)|_{H^s} = O(\varepsilon)$.*

- Case of ill-prepared initial data : we choose $U_0 = \alpha_+ e_+ + \alpha_- e_- + c.c.$ with $(\alpha_+, \alpha_-) \in \mathbb{C}^2$ such that $|\alpha_-| \neq 0$. It implies that
$$
\alpha_+ = f + i\frac{\varepsilon}{2} g \quad \text{and} \quad \alpha_- = -i\frac{\varepsilon}{2} g.
$$

In opposition with the previous case, we can impose that $\lim_{\varepsilon \to 0} \varepsilon g \neq 0$ and in this case, the initial data is not bounded with respect to ε. It induces that $|G(0)|_{H^s(\mathbb{R})} = O(1)$. In summary, any initial data which read

$$
\mathcal{E}(t = 0) = f, \partial_t \mathcal{E}(t = 0) = g, \mathcal{N}(t = 0) = \mathcal{N}_0, \partial_t \mathcal{N}(t = 0) = \mathcal{N}_1
$$

with $g = O(\frac{1}{\varepsilon})$ are ill prepared. We take for example $g = i\frac{f}{\varepsilon}$ so $U_0(x) = \frac{f}{2}(e_+ + e_-) + c.c.$. The associated initial condition of Zakharov system is

$$
E^0(t = 0) = \frac{f}{2}, n^0(t = 0) = \mathcal{N}_0, \partial_t n^0(t = 0) = \mathcal{N}_1.
$$

It is recovered from $U_0(x) = \frac{f}{2}(e_+ + e_-) + c.c.$ by getting $|e_-| = 0$.

We investigate the convergence of the solution of each scheme when $\varepsilon \to 0$ for polarized and ill prepared initial data.

3.2.2 Numerical experiments

We perform the computation on the space-time domain $[-5, 5] \times [0, 1]$ with 2048 points in space and for $\delta t = 2, 5.10^{-6}$. We get some values of ε for which Zakharov model is available, that is for ε small enough. We will use the following errors involving the L^2 norm

$$e_{L^2}(t^m) := \frac{(\sum_{j=0}^{J} \delta x |\mathcal{E}_j^m - E_j^m|^2)^{\frac{1}{2}}}{(\sum_{j=0}^{J} \delta x |E_j^m|^2)^{\frac{1}{2}}} \text{ and } e_{L^2}^{pola}(t^m) := \frac{(\sum_{j=0}^{J} \delta x |F_j^m - E_j^m|^2)^{\frac{1}{2}}}{(\sum_{j=0}^{J} \delta x |E_j^m|^2)^{\frac{1}{2}}}$$

where F^m is the mean value in time of $\mathcal{E}^m$. Therefore $e_{L^2}(t^m)$ computes the error between the solution of each scheme (S_1) or (S_2) and the one of the Zakharov system whereas $e_{L^2}^{pola}(t^m)$ computes the error between the mean value in time of the solution of each scheme (S_1) or (S_2) and the one of the Zakharov system. For the four following figures, we represent $e_{L^2}(t)$ with solid lines and $e_{L^2}^{pola}(t)$ with dotted lines.

- Case of polarized initial data : we take $f = 0, 5e^{-5x^2}, g = 5i(10x^2 - 1)f, \mathcal{N}_0 = 0, \mathcal{N}_1 = 0$. The left curves represent the errors $e_{L^2}(t), e_{L_2}^{pola}(t)$ given by scheme (S_1) for $\varepsilon = 10^{-4}$ (figure 3) and $\varepsilon = 10^{-6}$ (figure 4). The right curves represent the same quantities given by scheme (S_2).
  We remark that the solution of each scheme converges to the one of Zakharov system as $\varepsilon \to 0$.
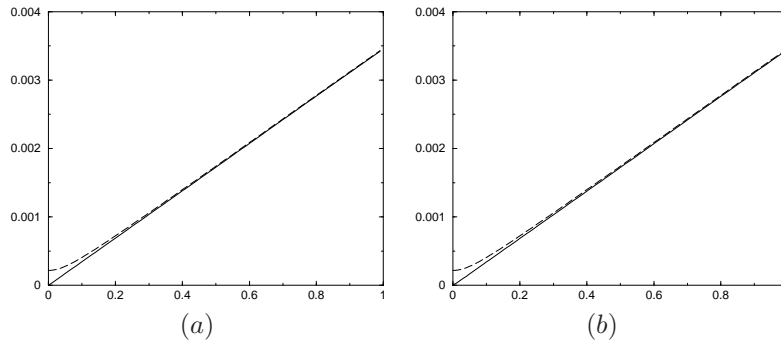

Figure 3: $\varepsilon = 10^{-4}$, errors $e_{L^2}(t)$ (solid line), $e_{L_2}^{pola}(t)$ (dotted line); (a) scheme (S_1) ; (b) scheme (S_2).

This is the result established in [2] and recalled in section 1.

- Case of ill prepared initial data : we take $f = 0.5e^{-5x^2}, g = \frac{i}{\varepsilon}f, \mathcal{N}_0 = 0, \mathcal{N}_1 = 0$. We choose $\delta t > \varepsilon^2$.
  In figures 5, 6, we represent $e_{L^2}(t)$ (solid line) and $e_{L^2}^{pola}(t)$ (dotted line) obtained by scheme (S_1) (left curves) and scheme (S_2) (right curves). We get $\varepsilon = 10^{-4}$ for figure 5 and $\varepsilon = 10^{-6}$ for figure 6. We start with two different initial conditions $\mathcal{E}(t = 0)$ and $E^0(t = 0)$. However solid and dotted lines of figures 5 (a) and 6 (a) are superposed. It means that the solution of scheme (S_1) converges to the discrete solution of Zakharov system 1.1 whatever initial data we take. This is the polarization phenomenon. On the contrary only the part of F of the solution of scheme (S_2)
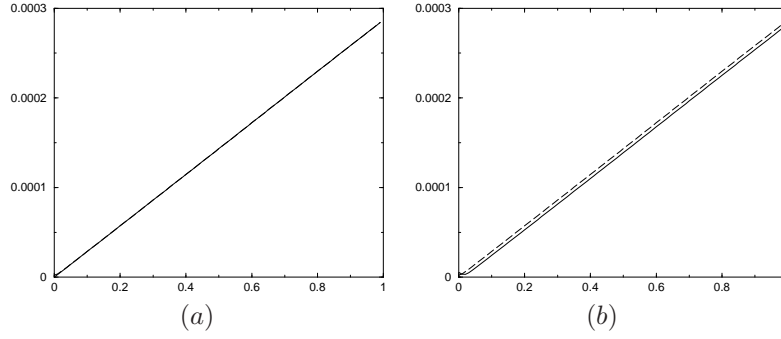
19

Figure 4: $\varepsilon = 10^{-6}$, errors $e_{L^2}(t)$ (solid line), $e_{L_2}^{pola}(t)$ (dotted line); (a) scheme (S_1) ; (b) scheme (S_2).
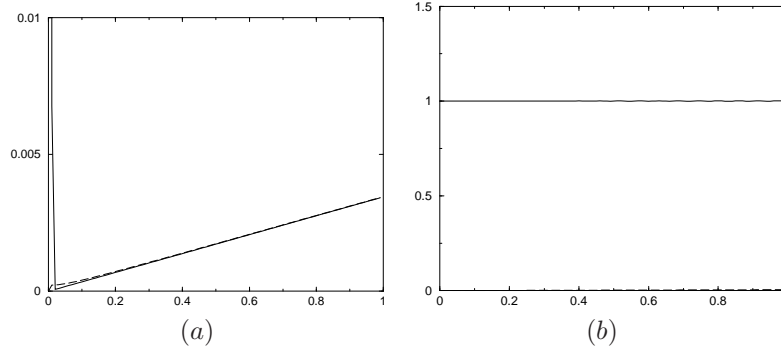

Figure 5: $\varepsilon = 10^{-4}$; $e_{L^2}(t)$ (solid line) and $e_{L^2}^{pola}(t)$ (dotted line); (a) scheme (S_1); (b) scheme (S_2).
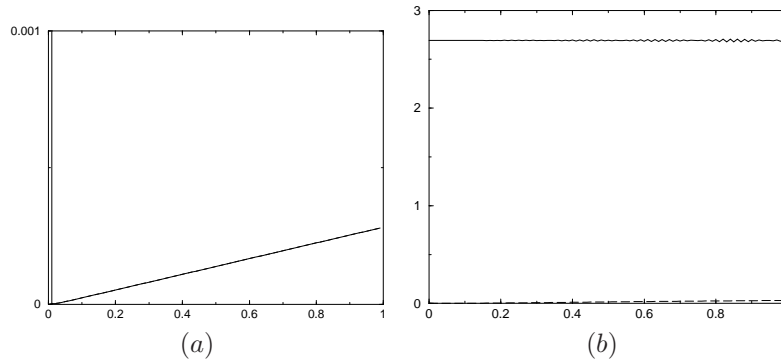

Figure 6: $\varepsilon = 10^{-6}$; $e_{L^2}(t)$ (solid line) and $e_{L^2}^{pola}(t)$ (dotted line); (a) scheme (S_1); (b) scheme (S_2).

converges to the solution of Zakharov system 1.1. We conclude that only scheme (S_1) exhibits the polarization property.

3.2.3 Large spectrum or short waves

The aim of this section is to show that system (1.2) models the propagation of short or ultra short waves [1] through a plasma contrary to Zakharov system (1.1). Moreover we illustrate the error estimate established in [4]. In order to solve system (1.4) we use a finite difference method in 1D. The numerical scheme reads

$$\varepsilon\left(\frac{E_j^{m+1} - 2E_j^m + E_j^{m-1}}{\delta t^2}\right) - \left(D_+D_-\left(\frac{E^{m+1} + E^{m-1}}{2}\right)\right)_j + \frac{1}{\varepsilon}\left(\frac{E_j^{m+1} + E_j^{m-1}}{2}\right)$$

$$= -n^m\left(\frac{E_j^{m+1} + E_j^{m-1}}{2}\right), \quad (3.7)$$

$$\frac{n_j^{m+1} - 2n_j^m + n_j^{m-1}}{\delta t^2} - \left(D_+D_-\left(\frac{n^{m+1} + n^{m-1}}{2}\right)\right)_j = \left(D_+D_-|E^m|^2\right)_j$$

endowed with the initial conditions

$$E(t = 0) = E_0, \partial_t E(t = 0) = E_1, n(t = 0) = n_0, \partial_t n(t = 0) = n_1.$$

Operators D_+ and D_- have been introduced in section 1. We introduce the errors in L^2 norm, namely

$$e_{new}(t) = \frac{|E(t,.) - (\mathcal{E}(t,.)e^{-i\frac{t}{\varepsilon}} + c.c.)|_2}{\max(|E(t,.)|_2, |\mathcal{E}(t,.)e^{-i\frac{t}{\varepsilon}} + c.c.|_2)}, n_{new}(t) = \frac{|n(t,.) - \mathcal{N}(t,.)|_2}{\max(|n(t,.)|_2, |\mathcal{N}(t,.)||_2)},$$

$$e_{zak}(t) = \frac{|E(t,.) - (E^0(t,.)e^{-i\frac{t}{\varepsilon}} + c.c.)|_2}{\max(|E(t,.)|_2, |E^0(t,.)e^{-i\frac{t}{\varepsilon}} + c.c.|_2)}, n_{zak}(t) = \frac{|n(t,.) - n^0(t,.)|_2}{\max(|n(t,.)|_2, |n^0(t,.)|_2)}.$$

We compute on the space-time interval $[-5, 5] \times [0, 1]$ with 2048 points in space and $\delta t = 2,5.10^{-6}$. We take $f = 0,5e^{-5x^2}, g = 0,5e^{-5x^2}, \mathcal{N}_0 = 0, \mathcal{N}_1 = 0$ so $E_0 = 2f, E_1 = 2g, n_0 = 0, n_1 = 0, E^0(t = 0) = f, n^0(t = 0) = 0, \partial_t n^0(t = 0) = 0$. The following figures represent errors e_{new}, n_{new} (left curves) and e_{zak}, n_{zak} (right curves) for $\varepsilon = 10^{-1}$ (figure 7), $\varepsilon = 10^{-2}$ (figure 8) and $\varepsilon = 10^{-3}$ (figure 9).
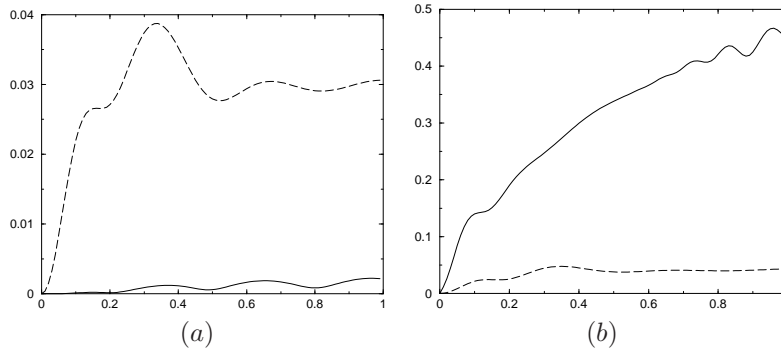

Figure 7: $\varepsilon = 10^{-1}$. $(a)e_{new}(t)$ (solid line), $n_{new}(t)$ (dotted line); $(b)e_{zak}(t)$ (solid line), $n_{zak}(t)$ (dotted line).

These results show that the Zakharov system (1.1) is not a suitable system for modeling the propagation of a short or ultra short wave in a plasma. Moreover curves of figures 7 (b), 8 (b), 9 (b) validate
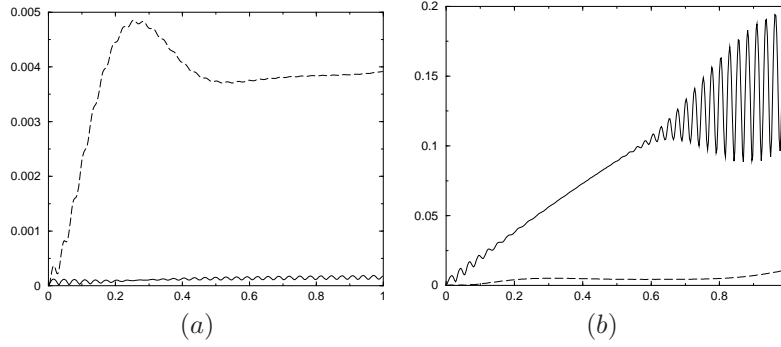
Figure 8: $\varepsilon = 10^{-2}$. $(a)e_{new}(t)$ (solid line), $n_{new}(t)$ (dotted line); $(b)e_{zak}(t)$ (solid line), $n_{zak}(t)$ (dotted line).
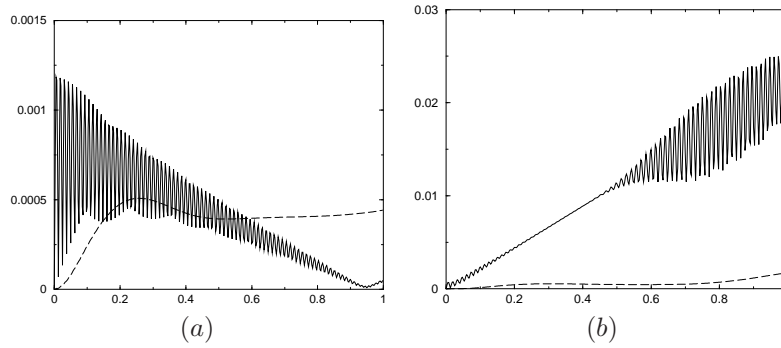

Figure 9: $\varepsilon = 10^{-3}$. $(a)e_{new}(t)$ (solid line), $n_{new}(t)$ (dotted line); $(b)e_{zak}(t)$ (solid line), $n_{zak}(t)$ (dotted line).

the error estimate established in [4]. In particular, with figure 9 (b), we check the convergence result established in [4] related to Klein-Gordon-wave system and Zakharov system when ε is small enough.

References

[1] K. Barrailh and D. Lannes. A General Framework for Diffractive Optics and its Applications to Lasers with Large Spectrum and Short Pulses. SIAM Journal of Math. Analysis **2002**.

[2] L. Bergé, B. Bidégaray, T. Colin. A perturbative analysis of the time-envelope approximation in strong Langmuir turbulence. Physica D, **1996**, *95* (3-4), 351–379.

[3] Q. Chang and H. Jiang. A Conservative Difference Scheme for the Zakharov Equations. J. Comput. Phys. **1994**, *113* (2), 309–319.

[4] T. Colin, G. Ebrard, G. Gallice and B. Texier. Justification of the Zakharov model from Klein-Gordon-Wave systems. Comm. Part. Diff. Eq. **2004**, *29* (9-10), pp 1365-1401.

[5] T. Colin and P. Fabrie. Semidiscretization in time for Nonlinear Schrödinger-Waves Equations. Discrete and Continuous Dynamical Systems **1998**, *4* (4), pp 671–690.

[6] R. Glassey. Approximate solutions to the Zakharov equations via finite differences. J. Comput. Phys. **1992**, *100* (2), 377–383.

[7] S. Jin, P. A. Markowich and C. Zheng. Numerical simulation of a generalized Zakharov system. J. Comput. Phys. **2004**, *201* (1), 376–395.

[8] J.-L. Joly, G. Métivier and J. Rauch. Diffractive nonlinear geometric optics with rectification. Indiana University Journal, **1998**, *47*, 1167–1241.

[9] J.-L. Joly, G. Métivier and J. Rauch. Recent results in non-linear geometric optics. In *Hyperbolic problems : theory, numerics, applications, Vol. II (Zürich 1998)*, volume 130 of *Internat. Ser. Numer. Math.*, pages 723-736. Birkhaüser Boston, 1999.

[10] J.-L. Joly, G. Métivier and J. Rauch. Several recent results in nonlinear geometric optics. In *Partial differential equations and mathematical physics (Copenhague, 1995; Lund, 1995)*, volume 21 of *Progr. Nonlinear Differential Equations Appl.*, pages 181-206. Birkhaüser Boston, 1996.

[11] J.-L. Joly, G. Métivier, J. Rauch. Transparent nonlinear geometric optics and Maxwell-Bloch equations. Journal of Differential Equations **2000**, *166*, 175–250.

[12] K. Nakanishi and N. Masmoudi. Uniqueness of Finite Energy Solutions for Maxwell-Dirac and Maxwell-Klein-Gordon Equations. Comm. Math. Phys. **2003**, *243* (1), 123–136.

[13] D. R. Nicholson, Introduction to Plasma Theory (Wiley, New York, 1983) pp. 177-181.

[14] G.L. Payne, D.R. Nicholson and R.M. Downie. Numerical Solution of the Zakharov Equations. J. Comput. Phys. **1983**, *50*, 482–498.

[15] B. Texier. Derivation of the Zakharov equations, Archive for Rational Mechanics and Analysis 184 **2007**, no.1, 121-183.