



**HAL**  
open science

# Effective integrable dynamics for some nonlinear wave equation

Patrick Gerard, Sandrine Grellier

► **To cite this version:**

Patrick Gerard, Sandrine Grellier. Effective integrable dynamics for some nonlinear wave equation. Analysis & PDE, 2012, 5-5, pp.1139-1155. hal-00635686

**HAL Id: hal-00635686**

**<https://hal.science/hal-00635686>**

Submitted on 25 Oct 2011

**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.

# EFFECTIVE INTEGRABLE DYNAMICS FOR SOME NONLINEAR WAVE EQUATION

PATRICK GÉRARD AND SANDRINE GRELLIER

ABSTRACT. We consider the following degenerate half wave equation on the one dimensional torus

$$i\partial_t u - |D|u = |u|^2 u, \quad u(0, \cdot) = u_0.$$

We show that, on a large time interval, the solution may be approximated by the solution of a completely integrable system—the cubic Szegő equation. As a consequence, we prove an instability result for large  $H^s$  norms of solutions of this wave equation.

## 1. INTRODUCTION

Let us consider, on the one dimensional torus  $\mathbb{T}$ , the following “half-wave” equation

$$(1) \quad i\partial_t u - |D|u = |u|^2 u, \quad u(0, \cdot) = u_0.$$

Here  $|D|$  denotes the pseudo-differential operator defined by

$$|D|u = \sum_k |k| u_k e^{ikx}, \quad u = \sum_k u_k e^{ikx}.$$

This equation can be seen as a toy model for non linear Schrödinger equation on degenerate geometries leading to lack of dispersion. For instance, it has the same structure as the cubic non linear Schrödinger equation on the Heisenberg group, or associated with the Grušin operator. We refer to [4] and [5] for more detail.

We endow  $L^2(\mathbb{T})$  with the symplectic form

$$\omega(u, v) = \text{Im}(u|v).$$

Equation (1) may be seen as the Hamiltonian system related to the energy function  $H(u) := \frac{1}{2}(|D|u, u) + \frac{1}{4}\|u\|_{L^4}^4$ . In particular,  $H$  is invariant by the flow which also admits the following conservation laws,

$$Q(u) := \|u\|_{L^2}^2, \quad M(u) := (Du|u).$$

However, equation (1) is a non dispersive equation. Indeed, it is equivalent to the system

$$(2) \quad i(\partial_t \pm \partial_x)u_{\pm} = \Pi_{\pm}(|u|^2 u), \quad u_{\pm}(0, \cdot) = \Pi_{\pm}(u_0),$$

---

*Date:* Oct 24, 2011.

*2010 Mathematics Subject Classification.* 35B34, 35B40, 37K55.

where  $u_{\pm} = \Pi_{\pm}(u)$ . Here,  $\Pi_{+}$  denotes the orthogonal projector from  $L^2(\mathbb{T})$  onto

$$L^2_{+}(\mathbb{T}) := \left\{ u = \sum_{k \geq 0} u_k e^{ikx}, (u_k)_{k \geq 0} \in \ell^2 \right\}$$

and  $\Pi_{-} := I - \Pi_{+}$ .

Though the scaling is  $L^2$ -critical, the first iteration map of the Duhamel formula

$$u(t) = e^{-it|D|}u_0 - i \int_0^t e^{-i(t-\tau)|D|} (|u(\tau)|^2 u(\tau)) d\tau$$

is not bounded on  $H^s$  for  $s < \frac{1}{2}$ . Indeed, such boundedness would require the inequality

$$\int_0^1 \|e^{-it|D|} f\|_{L^4(\mathbb{T})}^4 dt \lesssim \|f\|_{H^{s/2}}^4.$$

However, testing this inequality on functions localized on positive modes for instance, shows that this fails if  $s < \frac{1}{2}$  (see the appendix for more detail).

Proceeding as in the case of the cubic Szegő equation (see [5], Theorem 2.1),

$$(3) \quad i\partial_t w = \Pi_{+}(|w|^2 w),$$

one can prove the global existence and uniqueness of solutions of (1) in  $H^s$  for any  $s \geq 1/2$ . The proof uses in particular the a priori bound of the  $H^{1/2}$ -norm provided by the energy conservation law.

**Proposition 1.** *Given  $u_0 \in H^{\frac{1}{2}}(\mathbb{T})$ , there exists  $u \in C(\mathbb{R}, H^{\frac{1}{2}}(\mathbb{T}))$  unique such that*

$$i\partial_t u - |D|u = |u|^2 u, \quad u(0, x) = u_0(x).$$

*Moreover if  $u_0 \in H^s(\mathbb{T})$  for some  $s > \frac{1}{2}$ , then  $u \in C(\mathbb{R}, H^s(\mathbb{T}))$ .*

Notice that similarly to the cubic Szegő equation, the proof of Proposition 1 provides only bad large time estimates,

$$\|u(t)\|_{H^s} \lesssim e^{e^{C_s t}}.$$

This naturally leads to the question of the large time behaviour of solutions of (1). In order to answer this question, a fundamental issue is the decoupling of non negative and negative modes in system (2). Assuming that initial data are small and spectrally localized on non negative modes, a first step in that direction is given by the next simple proposition, which shows that  $u_{-}(t)$  remains smaller in  $H^{1/2}$  uniformly in time.

**Proposition 2.** *Assume*

$$\Pi_+ u_0 = u_0 = O(\varepsilon) \text{ in } H^{\frac{1}{2}}(\mathbb{T}).$$

*Then, the solution  $u$  of (1) satisfies*

$$\sup_{t \in \mathbb{R}} \|\Pi_- u(t)\|_{H^{\frac{1}{2}}} = O(\varepsilon^2) .$$

*Proof.* By the energy and momentum conservation laws, we have

$$\begin{aligned} (|D|u, u) + \frac{1}{2} \|u\|_{L^4}^4 &= (|D|u_0, u_0) + \frac{1}{2} \|u_0\|_{L^4}^4 , \\ (Du, u) &= (Du_0, u_0) . \end{aligned}$$

Substracting these equalities, we get

$$2(|D|u_-, u_-) + \frac{1}{2} \|u\|_{L^4}^4 = \frac{1}{2} \|u_0\|_{L^4}^4 = O(\varepsilon^4) ,$$

hence

$$\|u_-\|_{H^{\frac{1}{2}}}^2 = O(\varepsilon^4) .$$

□

This decoupling result suggests to neglect  $u_-$  in system (2) and hence to compare the solutions of (1) to the solutions of

$$i\partial_t v - Dv = \Pi_+( |v|^2 v ),$$

which can be reduced to (3) by the transformation  $v(t, x) = w(t, x - t)$ .

Our main result is the following.

**Theorem 1.1.** *Let  $s > 1$  and  $u_0 = \Pi_+(u_0) \in L_+^2(\mathbb{T}) \cap H^s(\mathbb{T})$  with  $\|u_0\|_{H^s} = \varepsilon$ ,  $\varepsilon > 0$  small enough. Denote by  $v$  the solution of the cubic Szegö equation*

$$(4) \quad i\partial_t v - Dv = \Pi_+( |v|^2 v ), \quad v(0, \cdot) = u_0.$$

*Then, for any  $\alpha > 0$ , there exists a constant  $c = c_\alpha < 1$  so that*

$$(5) \quad \|u(t) - v(t)\|_{H^s} = \mathcal{O}(\varepsilon^{3-\alpha}) \text{ for } t \leq \frac{c_\alpha}{\varepsilon^2} \log \frac{1}{\varepsilon} .$$

*Furthermore, there exists  $c > 0$  such that*

$$(6) \quad \forall t \leq \frac{c}{\varepsilon^3} , \quad \|u(t)\|_{L^\infty} = \mathcal{O}(\varepsilon) .$$

Theorem 1.1 calls for several remarks. Firstly, if we rescale  $u$  as  $\varepsilon u$ , equation (1) becomes

$$i\partial_t u - |D|u = \varepsilon^2 |u|^2 u, \quad u(0, \cdot) = u_0$$

with  $\|u_0\|_{H^s} = 1$ . On the latter equation, it is easy to prove that  $u(t) = e^{-it|D|} u_0 + o(1)$  for  $t \ll \frac{1}{\varepsilon^2}$ , so that non linear effects only start for  $\frac{1}{\varepsilon^2} \lesssim t$ . Rescaling  $v$  as  $\varepsilon v$  in equation (4), Theorem 1.1 states that the cubic Szegö dynamics appear as the effective dynamics of equation (1) on a time interval where non linear effects are taken into account.

Secondly, as pointed out before, (4) reduces to (3) by a simple Galilean transformation. Equation (3) has been studied in [4], [5] and [6] where its complete integrability is established together with an explicit formula for its generic solutions. Consequently, the first part of Theorem 1.1 provides an accurate description of solutions of equation (1) for a reasonably large time. Moreover, the second part of Theorem 1.1 claims an  $L^\infty$  bound for the solution of (1) on an even larger time. This latter bound is closely related to a special conservation law of equation (3), namely, some Besov norm of  $v$  –see section 2 below.

Our next observation is that, in the case of small Cauchy data localized on non negatives modes, system (2) can be reformulated as a — singular — perturbation of the cubic Szegö equation (3). Indeed, write  $u_0 = \varepsilon w_0$  and  $u(t, x) = \varepsilon w(\varepsilon^2 t, x - t)$ , then  $w = w_+ + w_-$  solves the system

$$(7) \quad \begin{cases} i\partial_t w_+ & = \Pi_+(|w|^2 w) \\ i(\varepsilon^2 \partial_t - 2\partial_x)w_- & = \varepsilon^2 \Pi_-(|w|^2 w) \end{cases}$$

Notice that, for  $\varepsilon = 0$  and  $\Pi_+ w_0 = w_0$ , the solution of this system is exactly the solution of (3). It is therefore natural to ask how much, for  $\varepsilon > 0$  small, the solution of system (7) stays close to the solution of equation (3). Since equation (3) turns out to be completely integrable, this problem appears as a perturbation of a completely integrable infinite dimensional system. There is a lot of literature on this subject (see e.g. the books by Kuksin [13], Craig [3] and Kappeler–Pöschel [12] for the KAM theory). In the case of the 1D cubic NLS equation and of the modified KdV equation, with special initial data such as solitons or 2-solitons, we refer to recent papers by Holmer-Zworski ([8], [9]), Holmer-Marzuola-Zworski ([10]), Holmer-Perelman-Zworski ([11]) and to references therein. Here we emphasize that our perturbation is more singular and that we deal with general Cauchy data.

Finally, let us mention that the proof of Theorem 1.1 is based on a Poincaré-Birkhoff normal form approach, similarly to [1] and [7] for instance. More specifically, we prove that equation (4) turns out to be a Poincaré-Birkhoff normal form of equation (1), for small initial data with only non negative modes.

As a corollary of Theorem 1.1, we get the following instability result.

**Corollary 1.** *Let  $s > 1$ . There exists a sequence of data  $u_0^n$  and a sequence of times  $t^{(n)}$  such that, for any  $r$ ,*

$$\|u_0^n\|_{H^r} \rightarrow 0$$

*while the corresponding solution of (1) satisfies*

$$\|u^n(t^{(n)})\|_{H^s} \simeq \|u_0^n\|_{H^s} \left( \log \frac{1}{\|u_0^n\|_{H^s}} \right)^{2s-1}.$$

It is interesting to compare this result to what is known about cubic NLS. In the one dimensional case, the cubic NLS is integrable [17] and admits an infinite number of conservation laws which control the regularity of the solution in Sobolev spaces. As a consequence, no such norm inflation occurs. This is in contrast with the 2D cubic NLS case for which Colliander, Keel, Staffilani, Takaoka, Tao exhibited in [2] small initial data in  $H^s$  which give rise to large  $H^s$  solutions after a large time.

In our case, the situation is different. Although the cubic Szegő equation is completely integrable, its conservation laws do not control the regularity of the solutions, which allows a large time behavior similar to the one proved in [2] for 2D cubic NLS (see [5] section 6, corollary 5). Unfortunately, the time interval on which the approximation (5) holds does not allow to infer large solutions for (1), but only solutions with large relative size with respect to their Cauchy data –see section 3 below. A time interval of the form  $[0, \frac{1}{\varepsilon^{2+\beta}}]$  for some  $\beta > 0$  would be enough to construct large solutions for (1) for some  $H^s$ -norms.

We close this introduction by mentioning that O. Pocovnicu solved a similar problem for equation (1) on the line by using the renormalization group method instead of the Poincaré-Birkhoff normal form method. Moreover, she improved the approximation in Theorem 1.1 by introducing a quintic correction to the Szegő cubic equation [15].

The paper is organized as follows. In section 2 we recall some basic facts about the Lax pair structure for the cubic Szegő equation (3). In section 3, we deduce Corollary 1 from Theorem 1.1. Finally, the proof of Theorem 1.1 is given in section 4.

## 2. THE LAX PAIR FOR THE CUBIC SZEGÖ EQUATION AND SOME OF ITS CONSEQUENCES

In this section, we recall some basic facts about equation (3) (see [5] for more detail). Given  $w \in H^{1/2}(\mathbb{T})$ , we define (see *e.g.* Peller [16], Nikolskii [14]), the Hankel operator of symbol  $w$  by

$$H_w(h) = \Pi_+(w\bar{h}) , \quad h \in L_+^2 .$$

It is easy to check that  $H_w$  is a  $\mathbb{C}$ -antilinear Hilbert-Schmidt operator. In [5], we proved that the cubic Szegő flow admits a Lax pair in the following sense. For simplicity let us restrict ourselves to the case of  $H^s$  solutions of (3) for  $s > \frac{1}{2}$ . From [5] Theorem 3.1, there exists a mapping  $w \in H^s \mapsto B_w$ , valued into  $\mathbb{C}$ -linear bounded skew-symmetric operators on  $L_+^2$ , such that

$$(8) \quad H_{-i\Pi_+(|w|^2w)} = [B_w, H_w] .$$

Moreover,

$$B_w = \frac{i}{2}H_w^2 - iT_{|w|^2} ,$$

where  $T_b$  denotes the Toeplitz operator of symbol  $b$  given by  $T_b(h) = \Pi_+(bh)$ . Consequently,  $w$  is a solution of (3) if and only if

$$(9) \quad \frac{d}{dt}H_w = [B_w, H_w] .$$

An important consequence of this structure is that the cubic Szegő equation admits an infinite number of conservation laws. Indeed, denoting  $W(t)$  the solution of the operator equation

$$\frac{d}{dt}W = B_w W , \quad W(0) = I ,$$

the operator  $W(t)$  is unitary for every  $t$ , and

$$W(t)^* H_{w(t)} W(t) = H_{w(0)} .$$

Hence, if  $w$  is a solution of (3), then  $H_{w(t)}$  is unitarily equivalent to  $H_{w(0)}$ . Consequently, the spectrum of the  $\mathbb{C}$ -linear positive self adjoint trace class operator  $H_w^2$  is conserved by the evolution. In particular, the trace norm of  $H_w$  is conserved by the flow. A theorem by Peller, see [16], Theorem 2, p. 454, states that the trace norm of a Hankel operator  $H_w$  is equivalent to the norm of  $w$  in the Besov space  $B_{1,1}^1(\mathbb{T})$ . Recall that the Besov space  $B^1 = B_{1,1}^1(\mathbb{T})$  is defined as the set of functions  $w$  so that  $\|w\|_{B_{1,1}^1}$  is finite where

$$\|w\|_{B_{1,1}^1} = \|S_0(w)\|_{L^1} + \sum_{j=0}^{\infty} 2^j \|\Delta_j w\|_{L^1} ,$$

here  $w = S_0(w) + \sum_{j=0}^{\infty} \Delta_j w$  stands for the Littlewood-Paley decomposition of  $w$ . It is standard that  $B^1$  is an algebra included into  $L^\infty$  (in fact into the Wiener algebra). The conservation of the trace norm of  $H_w$  therefore provides an  $L^\infty$  estimate for solutions of (3) with initial data in  $B^1$ .

The space  $B^1$  and formula (8) will play an important role in the proof of Theorem 1.1. In particular, the last part will follow from the fact that  $\|u(t)\|_{B^1}$  remains bounded by  $\varepsilon$  for  $t \ll \frac{1}{\varepsilon^3}$ . The fact that  $H^s(\mathbb{T}) \subset B^1$  for  $s > 1$ , explains why we assume  $s > 1$  in the statement.

### 3. PROOF OF COROLLARY 1

As observed in [5], section 6.1, Proposition 7, and section 6.2, Corollary 5, the equation

$$i\partial_t w = \Pi_+( |w|^2 w ) , \quad w(0, x) = \frac{a_0 e^{ix} + b_0}{1 - p_0 e^{ix}}$$

with  $a_0, b_0, p_0 \in \mathbb{C}, |p_0| < 1$  can be solved as

$$w(t, x) = \frac{a(t) e^{ix} + b(t)}{1 - p(t) e^{ix}}$$

where  $a, b, p$  satisfy an ODE system explicitly solvable.

In the particular case when

$$a_0 = \varepsilon, \quad b_0 = \varepsilon \delta, \quad p_0 = 0, \quad w_\varepsilon(0, x) = \varepsilon(e^{ix} + \delta),$$

one finds

$$1 - \left| p \left( \frac{\pi}{2\varepsilon^2\delta} \right) \right|^2 \simeq \delta^2,$$

so that, for  $s > \frac{1}{2}$ ,

$$\left\| w_\varepsilon \left( \frac{\pi}{2\varepsilon^2\delta} \right) \right\|_{H^s} \simeq \frac{\varepsilon}{\delta^{2s-1}}.$$

Let  $v_\varepsilon$  be the solution of

$$i(\partial_t + \partial_x)v_\varepsilon = \Pi_+( |v_\varepsilon|^2 v_\varepsilon ), \quad v_\varepsilon(0, x) = \varepsilon(e^{ix} + \delta)$$

then  $v_\varepsilon(t, x) = w_\varepsilon(t, x - t)$  so that

$$\left\| v_\varepsilon \left( \frac{\pi}{2\varepsilon^2\delta} \right) \right\|_{H^s} \simeq \frac{\varepsilon}{\delta^{2s-1}}.$$

Choose

$$\varepsilon = \frac{1}{n}, \quad \delta = \frac{C}{\log n}$$

with  $C$  large enough so that if  $t^{(n)} := \frac{\pi}{2\varepsilon^2\delta}$  then  $t^{(n)} < c \frac{\log(1/\varepsilon)}{\varepsilon^2}$ , where  $c = c_\alpha$  in Theorem 1.1 for  $\alpha = 1$ , say. Denote by  $u_0^n := v_\varepsilon(0, \cdot)$ . As  $\|u_0^n\|_{H^s} \simeq \varepsilon$ , the previous estimate reads

$$\left\| v_\varepsilon \left( \frac{\pi}{2\varepsilon^2\delta} \right) \right\|_{H^s} \simeq \|u_0^n\|_{H^s} \left( \log \frac{1}{\|u_0^n\|_{H^s}} \right)^{2s-1}.$$

Applying Theorem 1.1, we get the same information about  $\|u_n(t^{(n)})\|_{H^s}$ .

#### 4. PROOF OF THEOREM 1.1

First of all, we rescale  $u$  as  $\varepsilon u$  so that equation (1) becomes

$$(10) \quad i\partial_t u - |D|u = \varepsilon^2 |u|^2 u, \quad u(0, \cdot) = u_0$$

with  $\|u_0\|_{H^s} = 1$ .



4.1. **Study of the resonances.** We write the Duhamel formula as

$$u(t) = e^{-it|D|}\underline{u}(t)$$

with

$$\hat{\underline{u}}(t, k) = \hat{u}_0(k) - i\varepsilon^2 \sum_{k_1 - k_2 + k_3 - k = 0} I(k_1, k_2, k_3, k),$$

where

$$I(k_1, k_2, k_3, k) = \int_0^t e^{-i\tau\Phi(k_1, k_2, k_3, k)} \hat{\underline{u}}(\tau, k_1) \overline{\hat{\underline{u}}(\tau, k_2)} \hat{\underline{u}}(\tau, k_3) d\tau ,$$

and

$$\Phi(k_1, k_2, k_3, k_4) := |k_1| - |k_2| + |k_3| - |k_4| .$$

If  $\Phi(k_1, k_2, k_3, k_4) \neq 0$ , an integration by parts in  $I(k_1, k_2, k_3, k_4)$  provides an extra factor  $\varepsilon^2$ , hence the set of  $(k_1, k_2, k_3, k_4)$  such that  $\Phi(k_1, k_2, k_3, k_4) = 0$  is expected to play a crucial role in the analysis. This set is described in the following lemma.

**Lemma 1.** *Given  $(k_1, k_2, k_3, k_4) \in \mathbb{Z}^4$ ,*

$$k_1 - k_2 + k_3 - k_4 = 0 \text{ and } |k_1| - |k_2| + |k_3| - |k_4| = 0$$

*if and only if at least one of the following properties holds :*

- (1)  $\forall j, k_j \geq 0$  .
- (2)  $\forall j, k_j \leq 0$  .
- (3)  $k_1 = k_2$  ,  $k_3 = k_4$  .
- (4)  $k_1 = k_4$  ,  $k_3 = k_2$  .

*Proof.* Consider  $(k_1, k_2, k_3, k_4) \in \mathbb{Z}^4$  such that  $k_1 - k_2 + k_3 - k_4 = 0$ ,  $|k_1| - |k_2| + |k_3| - |k_4| = 0$ , and the  $k_j$ 's are not all non negative or all non positive. Let us prove in that case that either  $k_1 = k_2$  and  $k_3 = k_4$ , or  $k_1 = k_4$  and  $k_3 = k_2$ . Without loss of generality, we can assume that at least one of the  $k_j$  is positive, for instance  $k_1$ . Then, subtracting both equations, we get that  $|k_3| - k_3 = |k_2| - k_2 + |k_4| - k_4$ . If  $k_3$  is non negative, then, necessarily both  $k_2$  and  $k_4$  are non negative and hence all the  $k_j$ 's are non negative. Assume now that  $k_3$  is negative. At least one among  $k_2, k_4$  is negative. If both of them are negative, then  $k_3 = k_2 + k_4$  but this would imply  $k_1 = 0$  which is impossible by assumption. So we get either that  $k_3 = k_2$  (and so  $k_1 = k_4$ ) or  $k_3 = k_4$  (and so  $k_1 = k_2$ ). This completes the proof of the lemma.  $\square$

4.2. **First reduction.** We get rid of the resonances corresponding to cases (3) and (4) by applying the transformation

$$(11) \quad u(t) \mapsto e^{2it\varepsilon^2\|u_0\|_{L^2}^2} u(t)$$

which, since the  $L^2$  norm of  $u$  is conserved, leads to the equation

$$(12) \quad i\partial_t u - |D|u = \varepsilon^2(|u|^2 - 2\|u\|_{L^2}^2)u, \quad u(0, \cdot) = u_0 .$$

Notice that this transformation does not change the  $H^s$  norm. The Hamiltonian function associated to the equation (12) is given by

$$H(u) = \frac{1}{2}(|D|u, u) + \frac{\varepsilon^2}{4}(\|u\|_{L^4}^4 - 2\|u\|_{L^2}^4) = H_0(u) + \varepsilon^2 R(u) ,$$

where

$$H_0(u) := \frac{1}{2}(|D|u, u) , \quad R(u) := \frac{1}{4}(\|u\|_{L^4}^4 - 2\|u\|_{L^2}^4) = \frac{1}{4} \sum_{\substack{k_1 - k_2 + k_3 - k_4 = 0, \\ k_1 \neq k_2, k_4}} u_{k_1} \overline{u_{k_2}} u_{k_3} \overline{u_{k_4}} .$$

**4.3. The Poincaré-Birkhoff normal form.** We claim that under a suitable canonical transformation on  $u$ ,  $H$  can be reduced to the following Hamiltonian

$$\tilde{H}(u) = H_0(u) + \varepsilon^2 \tilde{R}(u) + O(\varepsilon^4)$$

where

$$\tilde{R}(u) = \frac{1}{4} \sum_{\mathbf{k} \in \mathcal{R}} u_{k_1} \overline{u_{k_2}} u_{k_3} \overline{u_{k_4}}$$

with

$$\mathcal{R} = \{ \mathbf{k} = (k_1, k_2, k_3, k_4) : k_1 - k_2 + k_3 - k_4 = 0, \\ k_1 \neq k_2, k_1 \neq k_4, \quad \forall j, k_j \geq 0 \text{ or } \forall j, k_j \leq 0 \} .$$

We look for a canonical transformation as the value at time 1 of some Hamiltonian flow. In other words, we look for a function  $F$  such that its Hamiltonian vector field is smooth on  $H^s$  and on  $B^1$ , so that our canonical transformation is  $\varphi_1$ , where  $\varphi_\sigma$  is the solution of

$$(13) \quad \frac{d}{d\sigma} \varphi_\sigma(u) = \varepsilon^2 X_F(\varphi_\sigma(u)), \quad \varphi_0(u) = u.$$

Recall that, given a smooth real valued function  $F$ , its Hamiltonian vector field  $X_F$  is defined by

$$dF(u).h =: \omega(h, X_F(u)) ,$$

and, given two functions  $F, G$  admitting Hamiltonian vector fields, the Poisson bracket of  $F, G$  is defined by

$$\{F, G\}(u) = \omega(X_F(u), X_G(u)) .$$

Let us make some preliminary remarks about the Poisson brackets.

In view of the expression of  $\omega$ , we have

$$\{F, G\} := dG.X_F = \frac{2}{i} \sum_k (\partial_k F \partial_k G - \partial_{\bar{k}} G \partial_k F)$$

where  $\partial_k F$  stands for  $\frac{\partial F}{\partial u_k}$  and  $\partial_{\bar{k}} F$  for  $\frac{\partial F}{\partial \overline{u_k}}$ . In particular, if  $F$  and  $G$  are respectively homogeneous of order  $p$  and  $q$ , then their Poisson bracket is homogeneous of order  $p + q - 2$ .

We prove the following lemma.

**Lemma 2.** *Set*

$$F(u) := \sum_{k_1 - k_2 + k_3 - k_4 = 0} f_{k_1, k_2, k_3, k_4} u_{k_1} \overline{u_{k_2}} u_{k_3} \overline{u_{k_4}},$$

where

$$f_{k_1, k_2, k_3, k_4} = \begin{cases} \frac{i}{4(|k_1| - |k_2| + |k_3| - |k_4|)} & \text{if } |k_1| - |k_2| + |k_3| - |k_4| \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then  $X_F$  is smooth on  $H^s$ ,  $s > \frac{1}{2}$ , as well as on  $B^1$ , and

$$\{F, H_0\} + R = \tilde{R},$$

$$\|DX_F(u)h\| \lesssim \|u\|^2 \|h\|,$$

where the norm is taken either in  $H^s$ ,  $s > \frac{1}{2}$ , or in  $B^1$ .

*Proof.* First we make a formal calculation with  $F$  given by

$$F(u) := \sum_{k_1 - k_2 + k_3 - k_4 = 0} f_{k_1, k_2, k_3, k_4} u_{k_1} \overline{u_{k_2}} u_{k_3} \overline{u_{k_4}}$$

for some coefficients  $f_{k_1, k_2, k_3, k_4}$  to be determined later. We compute

$$\{F, H_0\} = \frac{1}{i} \sum_{k_1 - k_2 + k_3 - k_4 = 0} (-|k_1| + |k_2| - |k_3| + |k_4|) f_{k_1, k_2, k_3, k_4} u_{k_1} \overline{u_{k_2}} u_{k_3} \overline{u_{k_4}}$$

so that equality  $\{F, H_0\} + R = \tilde{R}$  requires

$$f_{k_1, k_2, k_3, k_4} = \begin{cases} \frac{i}{4(|k_1| - |k_2| + |k_3| - |k_4|)} & \text{if } |k_1| - |k_2| + |k_3| - |k_4| \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

One can easily check that the function  $F$  is explicitly given by

$$F(u) = \frac{1}{2} \text{Im} \left( (D_0^{-1} u_- |u_+|^2 u_+) - (D_0^{-1} u_+ |u_-|^2 u_-) - (D_0^{-1} |u_+|^2 |u_-|^2) \right)$$

where  $D_0^{-1}$  is the operator defined by

$$D_0^{-1} u(x) = \sum_{k \neq 0} \frac{u_k}{k} e^{ikx}.$$

In view of the above formula, the Hamiltonian vector field  $X_F(u)$  is a sum and products of terms involving the following maps  $f \mapsto \bar{f}$ ,  $f \mapsto D_0^{-1} f$ ,  $f \mapsto \Pi_{\pm} f$ ,  $(f, g) \mapsto fg$ . These maps are continuous on  $H^s$  and on  $B^1$ . Hence,  $X_F$  is smooth and its differential satisfies the claimed estimate on  $H^s$ ,  $s > \frac{1}{2}$ , and  $B^1$ .  $\square$

For further reference, we state the following technical lemma, which is based on straightforward calculations.

**Lemma 3.** *The function  $\tilde{R}$  and its Hamiltonian vector field are given by*

$$\begin{aligned}\tilde{R}(u) &= \frac{1}{4}(\|\tilde{u}_+\|_{L^4}^4 + \|\tilde{u}_-\|_{L^4}^4) + \operatorname{Re}((u|1)(u_-|u_-^2)) - \frac{1}{2}(\|u_+\|_{L^2}^4 + \|u_-\|_{L^2}^4), \\ iX_{\tilde{R}}(u) &= \Pi_+(|u_+|^2 u_+) + \Pi_-(|u_-|^2 u_-) - 2\|u_+\|_{L^2}^2 u_+ - 2\|u_-\|_{L^2}^2 u_- \\ &\quad + (u_-|u_-^2) + 2(1|u)\Pi_-(|u_-|^2) + (1|u)u_-^2,\end{aligned}$$

where we have set  $u_{\pm} := \Pi_{\pm}(u)$ .

The maps  $X_{\{F,R\}}$  and  $X_{\{F,\tilde{R}\}}$  are smooth homogeneous polynomials of degree five on  $B^1$  and on  $H^s$  for every  $s > \frac{1}{2}$ .

We now perform the canonical transformation

$$\chi_{\varepsilon} := \exp(\varepsilon^2 X_F).$$

**Lemma 4.** *Set  $\varphi_{\sigma} := \exp(\varepsilon^2 \sigma X_F)$  for  $-1 \leq \sigma \leq 1$ . There exist  $m_0 > 0$  and  $C_0 > 0$  so that, for any  $u \in B^1$  so that  $\varepsilon\|u\|_{B^1} \leq m_0$ ,  $\varphi_{\sigma}(u)$  is well defined for  $\sigma \in [-1, 1]$  and*

$$\begin{aligned}\|\varphi_{\sigma}(u)\|_{B^1} &\leq \frac{3}{2}\|u\|_{B^1} \\ \|\varphi_{\sigma}(u) - u\|_{B^1} &\leq C_0 \varepsilon^2 \|u\|_{B^1}^3 \\ \|D\varphi_{\sigma}(u)\|_{B^1 \rightarrow B^1} &\leq e^{C_0 \varepsilon^2 \|u\|_{B^1}^2}\end{aligned}$$

Moreover, the same estimates hold in  $H^s$ ,  $s > \frac{1}{2}$ , with some constants  $m(s)$  and  $C(s)$ .

*Proof.* Write  $\varphi_{\sigma}$  as the integral of its derivative and use Lemma 2 to get

$$(14) \quad \sup_{|\sigma| \leq \tau} \|\varphi_{\sigma}(u)\|_{B^1} \leq \|u\|_{B^1} + C\varepsilon^2 \sup_{|\sigma| \leq \tau} \|\varphi_{\sigma}(u)\|_{B^1}^3, \quad 0 \leq \tau \leq 1$$

We now use the following standard bootstrap lemma.

**Lemma 5.** *Let  $a, b, T > 0$  and  $\tau \in [0, T] \mapsto M(\tau) \in \mathbb{R}_+$  be a continuous function satisfying*

$$\forall \tau \in [0, T], M(\tau) \leq a + bM(\tau)^3.$$

*Assume*

$$\sqrt{3b} M(0) < 1, \quad \sqrt{3b} a < \frac{2}{3}.$$

*Then*

$$\forall \tau \in [0, T], M(\tau) \leq \frac{3}{2}a.$$

*Proof.* For the convenience of the reader, we give the proof of Lemma 5. The function  $f : z \geq 0 \mapsto z - bz^3$  attains its maximum at  $z_c = \frac{1}{\sqrt{3b}}$ , equal to  $f_m = \frac{2}{3\sqrt{3b}}$ . Consequently, since  $a$  is smaller than  $f_m$  by the second inequality,

$$\{z \geq 0 : f(z) \leq a\} = [0, z_-] \cup [z_+, +\infty)$$

with  $z_- < z_c < z_+$  and  $f(z_-) = a$ . Since  $M(\tau)$  belongs to this set for every  $\tau$  and since  $M(0)$  belongs to the first interval by the first inequality, we conclude by continuity that  $M(\tau) \leq z_-$  for every  $\tau$ . By concavity of  $f$ ,  $f(z) \geq \frac{2}{3}z$  for  $z \in [0, z_c]$ , hence  $z_- \leq \frac{3}{2}a$ .  $\square$

Let us come back to the proof of Lemma 4. If  $\varepsilon \|u\|_{B^1} < \frac{2}{3\sqrt{3}C}$ , equation (14) and Lemma 5 imply that

$$(15) \quad \sup_{|\sigma| \leq 1} \|\varphi_\sigma(u)\|_{B^1} \leq \frac{3}{2} \|u\|_{B^1},$$

which is the first estimate. For the second one, we write for  $|\sigma| \leq 1$ ,

$$\|\varphi_\sigma(u) - u\|_{B^1} = \|\varphi_\sigma(u) - \varphi_0(u)\|_{B^1} \leq |\sigma| \sup_{|s| \leq |\sigma|} \left\| \frac{d}{ds} \varphi_s(u) \right\|_{B^1} \leq C_0 \varepsilon^2 \|u\|_{B^1}^3,$$

where the last inequality comes from Lemma 2 and estimate (15).

It remains to prove the last estimate. We differentiate the equation satisfied by  $\varphi_\sigma$  and use again Lemma 2 to obtain

$$\begin{aligned} \|D\varphi_\sigma(u)\|_{B^1 \rightarrow B^1} &\leq 1 + \varepsilon^2 \left| \int_0^\sigma \|DX_F(\varphi_\tau(u))\|_{B^1 \rightarrow B^1} \|D\varphi_\tau(u)\|_{B^1 \rightarrow B^1} d\tau \right| \\ &\leq 1 + C_0 \varepsilon^2 \|u\|_{B^1}^2 \left| \int_0^\sigma \|D\varphi_\tau(u)\|_{B^1 \rightarrow B^1} d\tau \right|, \end{aligned}$$

and Gronwall's lemma yields the result. Analogous proofs give the estimates in  $H^s$ .  $\square$

Let  $u$  satisfy the assumption of Lemma 4 in  $B^1$  or in  $H^s$  for some  $s > \frac{1}{2}$ .

Let us compute  $H \circ \chi_\varepsilon = H \circ \varphi_1$  as the Taylor expansion of  $H \circ \varphi_\sigma$  at time 1 around 0. One gets

$$\begin{aligned}
H \circ \chi_\varepsilon &= H \circ \varphi_1 = H_0 \circ \varphi_1 + \varepsilon^2 R \circ \varphi_1 \\
&= H_0 + \frac{d}{d\sigma} [H_0 \circ \varphi_\sigma]_{\sigma=0} + \varepsilon^2 R + \\
&+ \int_0^1 \left( (1-\sigma) \frac{d^2}{d\sigma^2} [H_0 \circ \varphi_\sigma] + \varepsilon^2 \frac{d}{d\sigma} [R \circ \varphi_\sigma] \right) d\sigma \\
&= H_0 + \varepsilon^2 (\{F, H_0\} + R) + \varepsilon^4 \int_0^1 ((1-\sigma)\{F, \{F, H_0\}\} + \{F, R\}) \circ \varphi_\sigma d\sigma \\
&= H_0 + \varepsilon^2 \tilde{R} + \varepsilon^4 \int_0^1 \left( (1-\sigma)\{F, \tilde{R}\} + \sigma\{F, R\} \right) \circ \varphi_\sigma d\sigma \\
&:= H_0 + \varepsilon^2 \tilde{R} + \varepsilon^4 \int_0^1 G(\sigma) \circ \varphi_\sigma d\sigma .
\end{aligned}$$

By Lemma 3, one gets

$$\sup_{0 \leq \sigma \leq 1} \|X_{G(\sigma)}(w)\| \leq C \|w\|^5$$

where the norm stands for the  $B^1$  norm or the  $H^s$  norm. Since

$$X_{G(\sigma) \circ \varphi_\sigma}(u) = D\varphi_{-\sigma}(\varphi_\sigma(u)) \cdot X_{G(\sigma)}(\varphi_\sigma(u)) ,$$

we conclude from Lemma 4 that, if  $\varepsilon \|u\|_{B^1} \leq m_0$ ,

$$\|X_{G(\sigma) \circ \varphi_\sigma}(u)\|_{B^1} \leq C \|u\|_{B^1}^5 .$$

As a consequence, one can write

$$X_{H \circ \chi_\varepsilon} = X_{H_0} + \varepsilon^2 X_{\tilde{R}} + \varepsilon^4 Y ,$$

where, if  $\varepsilon \|u\|_{B^1} \leq m_0$ , then

$$\|Y(u)\|_{B^1} \lesssim \|u\|_{B^1}^5 .$$

An analogous estimate holds in  $H^s$ ,  $s > \frac{1}{2}$ .

**4.4. End of the proof.** We first deal with the  $B^1$ -norm of  $u$ , solution of equation (12). We are going to prove that  $\|u(t)\|_{B^1} = \mathcal{O}(1)$  for  $t \ll \frac{1}{\varepsilon^3}$  by the following bootstrap argument. We assume that for some  $K$  large enough with respect to  $\|u_0\|_{B^1}$ , for some  $T > 0$ , for all  $t \in [0, T]$ ,  $\|u(t)\|_{B^1} \leq 10K$ , and we prove that if  $T \ll \frac{1}{\varepsilon^3}$ ,  $\|u(t)\|_{B^1} \leq K$  for  $t \in [0, T]$ . This will prove the result by continuity.

Set, for  $t \in [0, T]$ ,

$$\tilde{u}(t) := \chi_\varepsilon^{-1}(u(t)) ,$$

so that  $\tilde{u}$  is solution of

$$i\partial_t \tilde{u} - |D|\tilde{u} = \varepsilon^2 X_{\tilde{R}}(\tilde{u}) + \varepsilon^4 Y(\tilde{u}) .$$

Moreover, by Lemma 4,

$$\|\tilde{u}(t) - u(t)\|_{B^1} \lesssim \varepsilon^2 \|u\|_{B^1}^3$$

and so by the hypothesis,  $\|\tilde{u}(t)\|_{B^1} \leq 11K$  if  $\varepsilon$  is small enough. In view of the expression of the Hamiltonian vector field of  $\tilde{R}$  in Lemma 3, the equation for  $\tilde{u}$  reads

$$\begin{cases} i\partial_t \tilde{u}_+ - D\tilde{u}_+ &= \varepsilon^2 (\Pi_+(|\tilde{u}_+|^2 \tilde{u}_+) - 2\|\tilde{u}_+\|_{L^2}^2 \tilde{u}_+ + \int_{\mathbb{T}} |\tilde{u}_-|^2 \tilde{u}_-) \\ &+ \varepsilon^4 Y_+(\tilde{u}) , \\ i\partial_t \tilde{u}_- + D\tilde{u}_- &= \varepsilon^2 (\Pi_-(|\tilde{u}_-|^2 \tilde{u}_-) - 2\|\tilde{u}_-\|_{L^2}^2 \tilde{u}_- + 2(1|\tilde{u})\Pi_-(|\tilde{u}_-|^2) + (1|\tilde{u})\tilde{u}_-) \\ &+ \varepsilon^4 Y_-(\tilde{u}) . \end{cases}$$

Notice that all the Hamiltonian functions we have dealt with so far are invariant by multiplication by complex numbers of modulus 1, hence their Hamiltonian vector fields satisfy

$$X(e^{i\theta} z) = e^{i\theta} z ,$$

so that the corresponding Hamiltonian flows conserve the  $L^2$  norm. Hence  $\tilde{u}$  has the same  $L^2$  norm as  $u$ , which is the  $L^2$  norm of  $u_0$ . In particular,  $|(1|\tilde{u})| \leq \|u_0\|_{L^2}$ .

Moreover, as  $\|u_0\|_{B^1} \lesssim \|u_0\|_{H^s} = \mathcal{O}(1)$  since  $s > 1$ ,  $\tilde{u}_0$  satisfies

$$\|\tilde{u}_0 - u_0\|_{B^1} \lesssim \varepsilon^2$$

by Lemma 4 so that, as  $u_{0-} = 0$ , we get  $\|\tilde{u}_{0-}\|_{B^1} = \mathcal{O}(\varepsilon^2)$ . Then we obtain from the second equation

$$\sup_{0 \leq \tau \leq t} \|\tilde{u}_-(\tau)\|_{B^1} \lesssim \varepsilon^2 + \varepsilon^2 t \left( \sup_{0 \leq \tau \leq t} \|\tilde{u}_-(\tau)\|_{B^1}^3 + \sup_{0 \leq \tau \leq t} \|\tilde{u}_-(\tau)\|_{B^1}^2 \right) + \varepsilon^4 t K^5 .$$

Let  $M(t) = \frac{1}{\varepsilon} \sup_{0 \leq \tau \leq t} \|\tilde{u}_-(\tau)\|_{B^1}$  so that, if  $t \leq T$ ,

$$M(t) \lesssim \varepsilon + \varepsilon^3 T M(t)^2 (1 + \varepsilon M(t)) + \varepsilon^3 T .$$

As  $3m^2 \leq 1 + 2m^3$  for any  $m \geq 0$ , we get

$$M(t) \lesssim \varepsilon + \varepsilon^3 T M(t)^3 + \varepsilon^3 T .$$

Using Lemma 5, we conclude that, if  $T \ll \frac{1}{\varepsilon^3}$ ,

$$\sup_{0 \leq \tau \leq T} \|\tilde{u}_-(\tau)\|_{B^1} \ll \varepsilon .$$

For further reference, notice that, if  $T \lesssim \frac{1}{\varepsilon^2} \log \frac{1}{\varepsilon}$ , this estimate can be improved as

$$\sup_{0 \leq \tau \leq T} \|\tilde{u}_-(\tau)\|_{B^1} \lesssim \varepsilon^{2-\alpha} , \quad \forall \alpha > 0 .$$

We come back to the case  $T \ll \frac{1}{\varepsilon^3}$ . From the estimate on  $\tilde{u}_-$ , we infer

$$\|\tilde{u}_+\|_{L^2}^2 = \|\tilde{u}\|_{L^2}^2 + \mathcal{O}(\varepsilon^2) = \|u_0\|_{L^2}^2 + \mathcal{O}(\varepsilon^2) ,$$

and the equation for  $\tilde{u}_+$  reads

$$i\partial_t \tilde{u}_+ - D\tilde{u}_+ = \varepsilon^2 (\Pi_+ (|\tilde{u}_+|^2 \tilde{u}_+) - 2\|u_0\|_{L^2}^2 \tilde{u}_+) + \varepsilon^4 Y_+(\tilde{u}) + \mathcal{O}(\varepsilon^5) + \mathcal{O}(\varepsilon^4) \tilde{u}_+ .$$

Since  $\tilde{u}_{0+}$  is not small in  $B^1$ , we have to use a different strategy to estimate  $\tilde{u}_+$ . We use the complete integrability of the cubic Szegő equation, especially its Lax pair and the conservation of the  $B^1$ -norm.

At this stage it is of course convenient to cancel the linear term  $\|u_0\|_{L^2}^2 \tilde{u}_+$  by multiplying  $\tilde{u}_+(t)$  by  $e^{2i\varepsilon^2 t \|u_0\|_{L^2}^2}$ . As pointed out before, this change of unknown is completely transparent to the above system. This leads to

$$i\partial_t \tilde{u}_+ - D\tilde{u}_+ = \varepsilon^2 \Pi_+ (|\tilde{u}_+|^2 \tilde{u}_+) + \varepsilon^4 Y_+(\tilde{u}) + \mathcal{O}(\varepsilon^5) + \mathcal{O}(\varepsilon^4) \tilde{u}_+ .$$

We now appeal to the results recalled in section 2. We introduce the unitary family  $U(t)$  defined by

$$i\partial_t U - DU = \varepsilon^2 (T_{|\tilde{u}_+|^2} - \frac{1}{2} H_{\tilde{u}_+}^2) U , \quad U(0) = I ,$$

so that, using formula (8),

$$i\partial_t (U(t)^* H_{\tilde{u}_+(t)} U(t)) = \varepsilon^4 U(t)^* H_{Y_+(\tilde{u}) + \mathcal{O}(\varepsilon) + \mathcal{O}(1)\tilde{u}_+} U(t) .$$

Then, we use Peller's theorem [16] which states, as recalled in section 2, that the trace norm of a Hankel operator of symbol  $b$  is equivalent to the  $B^1$ -norm of  $b$  to obtain

$$\begin{aligned} \|\tilde{u}_+(t)\|_{B^1} &\simeq \operatorname{Tr} |H_{\tilde{u}_+(t)}| \\ &\lesssim \operatorname{Tr} |H_{\tilde{u}_{0+}}| + \varepsilon^4 \int_0^t (\operatorname{Tr} |H_{Y_+(\tilde{u})}(\tau)| + \operatorname{Tr} |H_{\tilde{u}_+}(\tau)| + \varepsilon) d\tau \\ &\lesssim \|\tilde{u}_{0+}\|_{B^1} + \varepsilon^4 \int_0^t (\|\tilde{u}(\tau)\|_{B^1}^5 + \|\tilde{u}_+(\tau)\|_{B^1} + \varepsilon) d\tau \end{aligned}$$

so that as  $\|\tilde{u}(t)\|_{B^1} \leq 11K$ ,

$$\|\tilde{u}_+(t)\|_{B^1} \lesssim \|\tilde{u}_{0+}\|_{B^1} + \varepsilon^4 t (11K)^5 ,$$

and, if  $t \ll \frac{1}{\varepsilon^3}$  and  $\varepsilon$  is small enough,

$$\|\tilde{u}(t)\|_{B^1} \lesssim \|\tilde{u}_{0+}\|_{B^1} .$$

Using again the second estimate in Lemma 4, we infer

$$\|u(t)\|_{B^1} \leq K .$$

Finally, using the inverse of transformation (11) and multiplying  $u$  by  $\varepsilon$ , we obtain estimate (6) of Theorem 1.1.



We now estimate the difference between the solution of the wave equation and the solution of the cubic Szegö equation. Since we have applied transformation (11), we have to compare in  $B^1$  the solution  $u$  of equation (12) to the solution  $v$  of equation

$$i\partial_t v - Dv = \varepsilon^2(\Pi_+(|v|^2 v) - 2\|u_0\|_{L^2}^2 v), \quad v(0) = u_0.$$

Notice that, as  $u_0$  is bounded in  $H^s$ ,  $s > 1$ , and as the  $B^1$  norm is conserved by the cubic Szegö flow,

$$\|v(t)\|_{B^1} \simeq \|u_0\|_{B^1} \lesssim \|u_0\|_{H^s} = \mathcal{O}(1).$$

We shall prove that, for every  $\alpha > 0$ , there exists  $c_\alpha > 0$  such that,

$$\forall t \leq \frac{c_\alpha}{\varepsilon^2} \log \frac{1}{\varepsilon}, \quad \|u(t) - v(t)\|_{B^1} \leq \varepsilon^{2-\alpha}.$$

In view of the previous estimates, it is enough to prove that, on the same time interval,

$$\|\tilde{u}_+(t) - v(t)\|_{B^1} \leq \varepsilon^{2-\alpha},$$

where  $\tilde{u}_+$  satisfies

$$(16) \quad \begin{cases} i\partial_t \tilde{u}_+ - D\tilde{u}_+ &= \varepsilon^2 (\Pi_+(|\tilde{u}_+|^2 \tilde{u}_+) - 2\|u_0\|_{L^2}^2 \tilde{u}_+) + \mathcal{O}(\varepsilon^4), \\ \tilde{u}_+(0) &= \tilde{u}_{0,+}. \end{cases}$$

As  $\|\tilde{u}(t)\|_{B^1} \lesssim 1$ ,  $\|v(t)\|_{B^1} \lesssim 1$ ,  $\|\tilde{u}_{0,+} - u_0\|_{B^1} \leq \varepsilon^2 \|u_0\|_{B^1} \lesssim \varepsilon^2$  and  $(i\partial_t - D)(\tilde{u}_+ - v) = \varepsilon^2 \Pi_+(|\tilde{u}_+|^2 \tilde{u}_+ - |v|^2 v - 2\|u_0\|_{L^2}^2 (\tilde{u}_+ - v)) + \mathcal{O}(\varepsilon^4)$ , we get, using that  $B^1$  is an algebra on which  $\Pi_+$  acts,

$$\|\tilde{u}_+(t) - v(t)\|_{B^1} \lesssim \varepsilon^2 + \varepsilon^4 t + \varepsilon^2 \int_0^t \|\tilde{u}_+(\tau) - v(\tau)\|_{B^1} d\tau.$$

This yields

$$\|\tilde{u}_+(t) - v(t)\|_{B^1} \lesssim (\varepsilon^2 + \varepsilon^4 t) e^{\varepsilon^2 t},$$

hence, for  $t \leq \frac{c_\alpha}{\varepsilon^2} \log \frac{1}{\varepsilon}$ ,

$$\|\tilde{u}_+(t) - v(t)\|_{B^1} \leq \varepsilon^{2-\alpha}.$$

We now turn to the estimates in  $H^s$  for  $s > 1$ .

From the equation on  $v$  and the a priori estimate in  $B^1$ , it follows that  $\|v(t)\|_{H^s} \leq A e^{A\varepsilon^2 t}$ ,  $t > 0$ , so that  $\|v(t)\|_{H^s} \leq N(\varepsilon)$  for  $t \leq \frac{c}{\varepsilon^2} \log(\frac{1}{\varepsilon})$ ,  $0 < c \ll 1$  where  $N(\varepsilon) := A\varepsilon^{-cA}$ .

Let us assume that for some  $T > 0$ ,

$$\forall t \in [0, T], \quad \|u(t)\|_{H^s} \leq 10N(\varepsilon).$$

We are going to prove that, for every  $\alpha > 0$ , there exists  $c_\alpha > 0$  such that, if

$$T \leq \frac{c_\alpha}{\varepsilon^2} \log \frac{1}{\varepsilon} ,$$

then

$$\forall t \in [0, T], \|u(t) - v(t)\|_{H^s} \leq \varepsilon^{2-\alpha} ,$$

Since  $\|v(t)\|_{H^s} \leq N(\varepsilon)$  for  $t \leq \frac{c}{\varepsilon^2} \log \frac{1}{\varepsilon}$ , this will prove the result by a bootstrap argument.

As before, we perform the same canonical transformation

$$\tilde{u}(t) := \chi_\varepsilon^{-1}(u(t)) ,$$

to get the solution of

$$i\partial_t \tilde{u} - |D|\tilde{u} = \varepsilon^2 X_{\tilde{R}}(\tilde{u}) + \varepsilon^4 Y(\tilde{u}) .$$

By Lemma 4,

$$\|\tilde{u}(t) - u(t)\|_{H^s} \lesssim \varepsilon^2 N(\varepsilon)^3$$

and so  $\|\tilde{u}(t)\|_{H^s} \lesssim N(\varepsilon)$ . Therefore it suffices to prove that

$$\forall t \in [0, T], \|\tilde{u}(t) - v(t)\|_{H^s} \leq \varepsilon^{2-\alpha} .$$

We first deal with  $\tilde{u}_-$ . A similar argument as the one developed in  $B^1$  gives that for, for  $0 \leq t \lesssim \frac{1}{\varepsilon^2} \log \frac{1}{\varepsilon}$ ,

$$\sup_{0 \leq \tau \leq t} \|\tilde{u}_-(\tau)\|_{H^s} \leq C_\alpha \varepsilon^{2-\alpha}$$

for every  $\alpha > 0$ .

It remains to estimate the  $H^s$  norm of  $\tilde{u}_+ - v$ . Notice that

$$\|\tilde{u}_{0,+} - u_0\|_{H^s} \leq \varepsilon^2$$

by Lemma 4. We use the following inequality — recall that  $B^1 \subset L^\infty$ ,

$$\begin{aligned} \|\Pi_+( |u|^2 u - |v|^2 v )\|_{H^s} &\lesssim (\|u\|_{B^1}^2 + \|v\|_{B^1}^2) \|u - v\|_{H^s} + \\ &+ (\|v\|_{H^s} + \|u - v\|_{H^s}) (\|u\|_{B^1} + \|v\|_{B^1}) \|u - v\|_{B^1} . \end{aligned}$$

Plugging this into a Gronwall inequality, in view of the previous estimates, we finally get

$$\|\tilde{u}_+(t) - v(t)\|_{H^s} \leq \varepsilon^{2-\alpha}$$

for  $t \leq \frac{c_\alpha}{\varepsilon^2} \log \frac{1}{\varepsilon}$ . This completes the proof.

## 5. APPENDIX: A NECESSARY CONDITION FOR WELLPOSEDNESS

In this section, we justify that the boundedness in  $H^s$  of the first iteration map of the Duhamel formula

$$F(t) = e^{-it|D|}f - i \int_0^t e^{-i(t-\tau)|D|}(|F(\tau)|^2 F(\tau))d\tau$$

implies

$$\int_0^1 \|e^{-it|D|}f\|_{L^4(\mathbb{T})}^4 dt \lesssim \|f\|_{H^{s/2}}^4.$$

Indeed, assume the following inequality

$$\left\| \int_0^1 e^{-i(1-\tau)|D|}(|e^{-i\tau|D|}f|^2 e^{-i\tau|D|}f)d\tau \right\|_{H^s} \lesssim \|f\|_{H^s}^3.$$

We compute the scalar product of the expression in the left hand side with  $e^{-i|D|}f$  and we get

$$\int_0^1 \|e^{-i\tau|D|}f\|_{L^4}^4 d\tau \lesssim \|f\|_{H^s}^3 \|f\|_{H^{-s}}.$$

If we assume first that  $f$  is spectrally supported, that is if  $f = \Delta_N f$  for some  $N$ , then  $\|f\|_{H^{\pm s}} \simeq N^{\pm s} \|f\|_{L^2}$  and the preceding inequality reads

$$\int_0^1 \|e^{-i\tau|D|}f\|_{L^4}^4 d\tau \lesssim N^{2s} \|f\|_{L^2}^4.$$

Eventually, for general  $f = \sum_N \Delta_N(f)$ , we used the Littlewood-Paley estimate

$$\|g\|_{L^4}^4 \lesssim \sum_N \|\Delta_N g\|_{L^4}^4$$

to get

$$\int_0^1 \|e^{-i\tau|D|}f\|_{L^4}^4 d\tau \lesssim \|f\|_{H^{s/2}}^4.$$

## REFERENCES

- [1] Bambusi D., *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Physics 234 (2003), 253–283.
- [2] Colliander J., Keel M., Staffilani G., Takaoka H., Tao, T., *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrodinger equation*, Inventiones Math. 181 (2010), 39–113.
- [3] Craig, W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses, 9. Société Mathématique de France, Paris, 2000.

- [4] Gérard, P., Grellier, S., *The cubic Szegő equation*, Séminaire X-EDP, 20 octobre 2008, École polytechnique, Palaiseau.
- [5] Gérard, P., Grellier, S., *The cubic Szegő equation*, Ann. Scient. Éc. Norm. Sup. 43 (2010), 761-810.
- [6] Gérard, P., Grellier, S., *Invariant Tori for the cubic Szegő equation*, to appear in *Inventiones Mathematicae*, DOI 10.1007/s00222-011-0342-7.
- [7] Grébert, B., *Birkhoff Normal Form and Hamiltonian PDEs*, arXiv math.AP (2006), 55 pages.
- [8] Holmer, J., Zworski M., *Slow soliton interaction with delta impurities*, J. Mod. Dyn. 1, no. 4 (2007), 689–718.
- [9] Holmer, J., Zworski M., *Soliton interaction with slowly varying potentials*, Int. Math. Res. Not., (2008), no. 10, Art. ID rnn026, 36 pp.
- [10] Holmer, J., Marzuola J., Zworski M., *Fast soliton scattering by delta impurities* Comm. Math. Phys. 274, Number 1 (2007), 187-216.
- [11] Holmer, J., Perelman, G., Zworski M., *Effective dynamics of double solitons for perturbed mKdV*, with an appendix by Bernd Sturmfels, Comm. Math. Phys. 305(3)(2011), 363–425.
- [12] Kappeler, T., Pöschel, J., *KdV & KAM*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], 45. Springer-Verlag, Berlin, 2003.
- [13] Kuksin, S., *Nearly integrable infinite-dimensional Hamiltonian systems*. Lecture Notes in Mathematics, 1556. Springer-Verlag, Berlin, 1993.
- [14] Nikolskii, N. K., *Operators, functions, and systems: an easy reading. Vol. 1. Hardy, Hankel, and Toeplitz*. Translated from the French by Andreas Hartmann. Mathematical Surveys and Monographs, 92. American Mathematical Society, Providence, RI, 2002.
- [15] Pocovnicu, O., *First and second order approximations for a non linear equation*, in *Etude d'une équation non linéaire, non dispersive et complètement intégrable et de ses perturbations* PhD thesis, université Paris-sud, 2011, and paper to appear.
- [16] Peller, V. V., *Hankel operators of class  $\mathfrak{S}_p$  and their applications (rational approximation, Gaussian processes, the problem of majorization of operators)*, Math. USSR Sb. 41, 443-479 (1982).
- [17] Zakharov, V. E., Shabat, A. B., *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media*. Soviet Physics JETP 34 (1972), no. 1, 62–69.

UNIVERSITÉ PARIS-SUD XI, LABORATOIRE DE MATHÉMATIQUES D'ORSAY,  
CNRS, UMR 8628

*E-mail address:* Patrick.Gerard@math.u-psud.fr

MAPMO-UMR 6628, DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ D'ORLEANS,  
45067 ORLÉANS CEDEX 2, FRANCE

*E-mail address:* Sandrine.Grellier@univ-orleans.fr