The initial value problem for nonlinear Schrödinger equations

Luis Vega*

Abstract. I will review some recent work done in collaboration with C. E. Kenig, G. Ponce and C. Rolvung on a general method to solve locally in time the initial value problem for non-linear Schrödinger equations under some natural hypotheses of decay and regularity of the coefficients. Also some non-trapping conditions of the solutions of the hamiltonian flow associated to the initial data is needed. We will not assume ellipticity on the matrix of the leading order coefficients but just a non-degeneracy condition. The method is based on energy estimates which can be performed thanks to the construction of an integrating factor. This construction is of independent interest and relies on the analysis of some new pseudo-differential operators.

Mathematics Subject Classification (2000). Primary 35Q20; Secondary 35B45.

Keywords. Non-linear Schrödinger equations, ultrahyperbolic operators.

1. Introduction

In this lecture I shall describe some joint work with C. E. Kenig and G. Ponce on general non-linear Schrödinger equations built on spatial operators which are given by just non-degenerate quadratic forms. It has been and still is a great pleasure and a privilege to work with both of them.

This research was started some time ago in [14], and we could say it has come to a natural end with [16], [22] and [23], these two latter works written in collaboration with C. Rolvung. In the process we have used some fundamental work done by other authors, in particular those by Hayashi and Ozawa [10], Doi [6], [7], [8], and Craig, Kappeler and Strauss [4]. A look at the introduction of the papers [16], [22], and [23] is enough to realize that the precise results are rather technical and lengthy to write. Therefore, this lecture will be mainly expository, and I refer to the reader to the papers mentioned above and to the review given in [21] for the precise statement of the theorems.

We are interested in solving the initial value problem

$$\begin{cases} \partial_t u = i\tilde{\mathcal{L}}u + F(u, \nabla u, \bar{u}, \nabla \bar{u}), \\ u(x,0) = u_0(x), \end{cases} \tag{1}$$

*This work is partially supported by the grant MTM2004-03029.

Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006

where

- $u(x,t) \in \mathbb{C}$, $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times R^{n_2} = \mathbb{R}^n$, $n = n_1 + n_2$, and $t \in [0, T]$;
- $\tilde{\mathcal{L}}u(x) = \partial_j(\tilde{a}_{jk}(x,t,u,\nabla u,\bar{u},\nabla\bar{u})\partial_k u)$, and $\tilde{A}(u) = (\tilde{a}_{jk})_{jk}$ is a real, symmetric, invertible matrix such that

$$\lim_{|x|\to\infty} \tilde{A} = \begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & -\mathbb{I}_{n_2} \end{pmatrix}; \tag{2}$$

- F is a regular non-linear function, as is a polynomial with no linear terms.

By solving this equation I mean to find a large enough space X of initial data, and a unique solution u for each $u_0 \in X$ up to a time $T = T(u_0)$ such that u is unique in some space Y, and the map $u_0 \mapsto u$ from X to Y is continuous. Typically

$$X = X_{\alpha_0\beta_0} = \{u_0 \text{ such that } x^\alpha \partial^\beta u_0 \in L^2;\ |\alpha| \le \alpha_0,\ |\beta| \le \beta_0\},$$
$$Y = Y_{\alpha_1\beta_1} = \{\mathcal{C}([0,T] : X_{\alpha_1\beta_1})\}$$

for some finite (α_j, β_j), $j = 0, 1$, and with possibly $0 < \alpha_1 < \alpha_0$, $0 < \beta_1 < \beta_0$.

There are several reasons which motivate the study of such a general initial value problem. First of all, within this general model there are relevant equations which appear in the physics literature, as Davey–Stewartson and Zakharov–Schulman systems, Landau–Lipschitz equations, the Schrödinger map and others; see [32]. These physical models usually have a very rich algebraic structure which in many cases allows for some short cuts in their analysis; in particular regarding existence results. However, uniqueness generally involves considering the PDE solved by the difference of two solutions which need not have the same algebraic structure as the starting one, but that still is under the general setting given in (1). I will be more precise about this point in Section 3, where the particular case of the Schrödinger map is analyzed.

Another motivation is to get a better understanding of ultrahyperbolic operators including the constant coefficient case

$$\mathcal{L}_0 = \Delta_{x_1} - \Delta_{x_2}.$$

Notice that neither the heat nor the wave flow can be defined for $\mathcal{L}_0$, while the Schrödinger flow

$$e^{it\mathcal{L}_0} \tag{3}$$

makes perfect sense. Our knowledge about this operator is far behind the classical one $e^{it\Delta}$. In Section 9, I gather some information and open questions about linear and non-linear perturbations of the free propagator (3).

The rest of the paper is devoted to explain the assumptions needed to solve (1), and to exhibit the algebraic tools we use to construct the solution. The main result is given in Section 7, and some remarks about the elliptic setting can be found in Section 8.

2. Energy estimates

In order to solve equation (1) we use the so-called energy method. This is based on three steps. The first one is to add some artificial viscosity to the right-hand side of (1) depending on $\varepsilon > 0$, and to consider for some $T_\varepsilon > 0$ the equation

$$\begin{cases} \partial_t u_\varepsilon = -\varepsilon \Delta_x^2 + i\tilde{\mathcal{L}} u + F, & t \in [0, T_\varepsilon), \\ u_\varepsilon = u_0, & t = 0. \end{cases} \tag{4}$$

The existence of a solution u_ε of (4) can be easily proved by Picard iteration, using the regularity properties of the free propagator $e^{-\varepsilon t \Delta^2}$.

The second step relies on proving energy estimates for u_ε independent of ε. These estimates will give a universal time of existence valid for all u_ε, and also will allow us to pass to the limit in ε to obtain a solution of (1).

The final step is to prove uniqueness by looking at the equation satisfied by the difference $u_{\varepsilon\varepsilon'} = u_\varepsilon - u_{\varepsilon'}$. It is at this point where to work in a general setting as (1) turns out to be fundamental because the equation for $u_{\varepsilon\varepsilon'}$ is again of the same type.

In the rest of the paper I will mainly focus on the question of the energy estimates, exhibiting the algebraic tools needed to obtain them.

Recall that we have already built the solution u_ε of (4) with $\tilde{\mathcal{L}}$ as in (1). Hence we can now understand (4) as a linear equation

$$\begin{cases} \partial_t u = \mathcal{L}_\varepsilon u + i\big(\overrightarrow{b_1} \cdot \nabla u + \overrightarrow{b_2} \cdot \nabla \bar{u} + c_1 u + c_2 \bar{u}\big) + f, \\ u(x, 0) = u_0(x), \end{cases} \tag{5}$$

with

$$\mathcal{L}_\varepsilon u = \partial_t a_j(x,t)\partial_k u - \varepsilon \Delta^2 u; \;\; A = \big(a_{jk}(x,t)\big)_{jk} = \big(\tilde{a}_{jk}(x,t,u,\nabla u, \bar{u}, \nabla \bar{u})\partial_k u\big)_{jk}.$$

Therefore it seems appropriate to assume the following hypotheses regarding A and the coefficients $\overrightarrow{b_1}$, $\overrightarrow{b_2}$ and c_1, c_2:

H1. A is a regular real symmetric non-degenerate matrix, i.e. there exists $\gamma_0 \in (0, 1)$ such that

$$\gamma_0 |\xi| \leq |A\xi| \leq \gamma_0^{-1} |\xi|.$$

H2. The coefficients $\partial_x a_{jk}$, $\partial_{tx} a_{jk}$, $\partial_t a_{jk}$ have a pointwise decay for a sufficiently large $|x|$,

$$\sup_{|t| \leq T} |\partial_x^\alpha a_{jk}(x,t)| + |\partial_t \partial_x^{\alpha'} a_{jk}(x,t)| \leq \frac{C}{(1+|x|)^N},$$

with N, $|\alpha|$ and $|\alpha'|$ large enough, and so that

$$A - \begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & -\mathbb{I}_{n_2} \end{pmatrix}$$

has a similar decay for a large $|x|$.

H3. $\overrightarrow{b_1}$ and $\overrightarrow{b_2}$ are smooth complex vector fields that decay pointwise together with their derivatives at infinity.

H4. c_1 and c_2 are smooth complex scalar fields bounded together with their derivatives at infinity.

Let us start assuming that in (5) the external forces f and the potentials $\overrightarrow{b_1}$, $\overrightarrow{b_2}$ and c_1, c_2 are zero. Then because A is real and symmetric we trivially have

$$\begin{aligned}\frac{d}{dt}\langle u, u\rangle &= \langle \mathcal{L}_\varepsilon u, u\rangle + \langle u, \mathcal{L}_\varepsilon u\rangle \\ &= -\langle iA\partial_x u, \partial_x u\rangle - \langle \partial_x u, iA\partial_x u\rangle - 2\varepsilon\langle \Delta u, \Delta u\rangle \\ &= -2\varepsilon\langle \Delta u, \Delta u\rangle \le 0.\end{aligned} \tag{6}$$

Therefore

$$\sup_{0\le t\le T}\langle u, u\rangle + 2\varepsilon\int_0^T \langle \Delta u, \Delta u\rangle dt \le \langle u(0), u(0)\rangle. \tag{7}$$

If in the above calculation we add zero order terms c_1 and c_2 we will obtain after integration in time

$$\langle u, u\rangle(t) \le e^{Mt}\langle u, u\rangle(0), \tag{8}$$

with $M = \sup(|c_1| + |c_2|)$. As a conclusion from the point of view of energy estimates zero order terms are harmless.

Let us know consider that first order terms are not trivial. It is straightforward that if $\operatorname{Re}\overrightarrow{b_1}$ is zero and $\overrightarrow{b_2}$ just bounded, an integration by parts as the one given in (6) leads to the same estimate. However, the situation is completely different if $\operatorname{Re}\overrightarrow{b_1} \neq 0$. This can be easily seen even in one dimension.

Consider the model problem

$$\partial_t u = i(\partial_x^2 u + b_1\partial_x u) - \varepsilon\partial_x^4 u, \quad x \in \mathbb{R}. \tag{9}$$

Take for simplicity $b_1 = 1$. Using $\hat{u}$ the Fourier transform of u and Parseval's identity the calculation given in (6) becomes

$$\begin{aligned}\frac{d}{dt}\langle u, u\rangle &= \langle (i\partial_x^2 - \varepsilon\partial_4^x + i\partial_x)u, u\rangle + \langle u, i\partial_x^2 - \varepsilon\partial_x^4 + i\partial_x\rangle \\ &= -\varepsilon\langle \xi^4\hat{u}, \hat{u}\rangle - \langle \xi\hat{u}, \hat{u}\rangle.\end{aligned} \tag{10}$$

Therefore if $\operatorname{supp}\hat{u} \subset (-\infty, 0)$ we can not obtain a uniform bound in ε for $\langle u, u\rangle$. This is usually called the loss of derivatives obstruction, because regardless of ε there is in (10) one derivative more on the right-hand side than on the left-hand side. The final conclusion is that we should get rid, if possible, of $\operatorname{Re} b_1$.

Nevertheless there is a simple way of removing the first order term in (9) creating other ones of order zero, which is to use the integrating factor $\exp\left(\frac{1}{2}\int_x^\infty b_1\right)$. For this purpose define

$$v(x, t) = \mathcal{K}u = e^{1/2\int_x^\infty b_1(y,t)dy}u(x, t). \tag{11}$$

The equation for v becomes

$$\partial_t v = i\partial_x^2 v - \varepsilon \mathcal{K} \partial_x^4 u + c(x,t)v, \tag{12}$$

with

$$c(x,t) = \frac{i}{2}\partial_x b_1 + i\frac{b_1^2}{4} + \frac{i}{2}\int_x^\infty \partial_t b_1(y,t)dy. \tag{13}$$

Notice that $\mathcal{K}$ given in (11) is invertible so that

$$\varepsilon \mathcal{K}\partial_x^4 u = \mathcal{K}\partial_x^4 \mathcal{K}^{-1} v = \varepsilon \partial_x^4 v + \varepsilon \text{ (lower order)}.$$

This lower order terms are harmless because they are of order ε and can be easily absorbed using (7) and taking T small enough. Looking carefully at the calculation above we see that the only assumption we need on b_1 is that

$$\sup_{x,\, 0\le t\le T} \left|\operatorname{Re}\int_x^\infty b_1(y,t)\right| + \left|\operatorname{Re}\int_x^\infty \partial_t b_1(y,t)dt\right| \le M. \tag{14}$$

Notice also that because the operator $\mathcal{K}$ is given by multiplication by the integrating factor we have

$$\mathcal{K}\bar{u} = \overline{\mathcal{K}u}. \tag{15}$$

Therefore the symmetry of the good terms $\operatorname{Im} b_1$ and b_2 is not destroyed and can be added to (9) so that the computation given above works the same without any extra difficulty.

Finally let us recall that the condition given in (14) is very reminiscent of the one obtained by Mizohata in [27] (see also [11]), and suggests that H3 and H4 are natural assumptions.

3. The Schrödinger map

The above analysis rises the natural question of whether or not the integrating factor can be constructed also in higher dimensions. It is at this stage where considering a specific equation and not the general case given in (1) can make a big difference. An illustrative example is the so-called Schrödinger map given by

$$\begin{cases} \overrightarrow{u_t} = \overrightarrow{u} \wedge \Delta \overrightarrow{u}, \quad \overrightarrow{u} = \overrightarrow{u}(x,t), \quad x \in \mathbb{R}^d,\ t \in \mathbb{R},\ d = 1,2, \\ |\overrightarrow{u}|^2 = 1. \end{cases} \tag{16}$$

This equation written in coordinates (for example using the stereographic projection) involves non-linearities in the first order terms. Let us try to understand first the

one dimensional case $d = 1$. Then $\overrightarrow{u}$ is nothing but the tangent vector to a three dimensional curve $\overrightarrow{\gamma}(x,t)$ (i.e. $\overrightarrow{u} = \overrightarrow{\gamma}_x$) which satisfies

$$\begin{cases} \overrightarrow{\gamma}_t = \overrightarrow{\gamma}_x \wedge \overrightarrow{\gamma}_{xx}, \quad x \in \mathbb{R},\ t \in \mathbb{R}, \\ |\overrightarrow{\gamma}_x|^2 = 1. \end{cases} \tag{17}$$

This equation, which is sometimes called the Localized Induction Approximation, was obtained for the first time by Da Rios in 1906, see [5], as a crude approximation of the evolution of a vortex filament within Euler equations. From a geometrical point of view is better to call it the binormal flow because using Frenet equations (17) can be written as

$$\overrightarrow{\gamma}_t = c\,\overrightarrow{b},$$

where c denotes the curvature and $\overrightarrow{b}$ the binormal vector.

In 1972 Hasimoto [9] proposed the transformation

$$\psi(x,t) = c(x,t)e^{i\int_0^x \tau(y,t)\,dy}, \tag{18}$$

with τ denoting the torsion, to simplify (17) and therefore (16). After some computations he proves that if $\overrightarrow{\gamma}$ satisfies (17) then ψ solves

$$\psi_t = i\psi_{xx} + \frac{i}{2}(|\psi|^2\psi + a(t)), \tag{19}$$

for some real function $a(t)$. Therefore we could understand Hasimoto's transformation (18) as some kind of integrating factor which removes the non-linear first order terms appearing when (16) is written in local coordinates, to the expense of cubic zero order terms which are much easier to handle.

The situation for the Schrödinger map in dimension 2 is much more delicate. Hasimoto's transformation can still be used but it is not possible to remove completely the first order terms; see [1]. The equation obtained after the transformation has a good symmetry from the point of view of energy estimates which allows to prove existence even for H^1-solutions. However, much of this symmetry is lost when considering difference of solutions.

Therefore, we are back to the question we started this section with. In the one dimensional case it is possible to construct the integrating factor which greatly simplifies the equations, but a similar transformation even in the two dimensional case is far from clear.

4. The integrating factor

Let us go back to our linear equation (5) and to simplify assume in this section that $\varepsilon = 0$. Accordingly we shall write $\mathcal{L}_\varepsilon = \mathcal{L}$. We are looking for an operator $\mathcal{K}$ such that

$$(\mathcal{L}\mathcal{K} - \mathcal{K}\mathcal{L}) = \mathcal{K}\overrightarrow{b_1}\cdot\nabla + \text{ zero order}. \tag{20}$$

It turns out that this equation, except in very simple cases, does not have a solution when just the algebra of classical differential operators is considered, as it happens in the one dimensional case, where $\mathcal{K}$ was constructed using the elemental operations of multiplication and integration. Therefore we need to consider $\mathcal{K}$ to be a pseudo-differential operator $\mathcal{K} = \mathcal{K}(x, t, D)$ given by:

$$\mathcal{K}u = \frac{1}{(2\pi)^n} \int_y \int_\xi k(x, t, \xi) e^{i(x-y)\xi} u d\xi \, dy. \tag{21}$$

If in the expression above we take

$$k(x, t, \xi) = a_\alpha(x, t)(i\xi)^\alpha, \tag{22}$$

then $\mathcal{K} = a_\alpha(x, t)\partial_x^\alpha$. Notice that in this case if a_α is regular and bounded together with its derivatives we get

$$|\partial_x^{\beta_1} \partial_\xi^{\beta_2} k(x, t, \xi)| \leq C_{\beta_1\beta_2}(1 + |\xi|)^{|\alpha|-|\beta_2|}. \tag{23}$$

Assume for a moment that $\mathcal{L}$ is the constant coefficient operator $\mathcal{L}_0$. That is to say

$$A = \begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & -\mathbb{I}_{n_2} \end{pmatrix}.$$

Solving formally in (20) we get that k should be given by

$$k(x, t, \xi) = \exp\left(\frac{1}{2} \int_0^\infty \overrightarrow{b_1}(x + s\tilde{\xi}, t)\right) \cdot \xi ds, \tag{24}$$

with $\tilde{\xi} = A\xi = (\xi_1, -\xi_2)$.

Notice that in (24) $s \mapsto x + s\tilde{\xi}$ is a geodesic associated to the pseudo-metric (or metric if $A = \mathbb{I}_n$) given by A. In the variable coefficient case $A = (a_{jk})_{jk}$ we have to define

$$k(x, t, \xi) = \exp\left(\frac{1}{2} \int_0^\infty \overrightarrow{b_1}\,(X(s; x, \xi)) \cdot \Xi(s; x, \xi) ds\right) \tag{25}$$

with $(X(s; x, \xi), \Xi(s; x, \xi))$ solutions of the hamiltonian flow H_A given by

$$\begin{cases} \dfrac{d}{ds} X_j(s; x_0, \xi_0) = -2 \displaystyle\sum_{k=1}^n a_{jk}\,(X(s; x_0, \xi_0) \Xi_k(s; x_0, \xi_0)) \\ \dfrac{d}{ds} \Xi_j\,(s; x_0, \xi_0) = \displaystyle\sum_{k,l=1}^n \partial_j a_{kl}\,(X(s; x_0, \xi_0))\; \Xi_k(s; x_0, \xi_0) \Xi_l(s; x_0, \xi_0) \\ (X(0; x_0, \xi_0), \Xi(0, x_0, \xi_0)) = (x_0, \xi_0). \end{cases} \tag{26}$$

Notice that above we have dropped the dependence on time to simplify the exposition. In fact, in our main result which is given in Section 7, we will impose some hypothesis on H_A where A is determined just by the initial condition u_0.

It follows from (26) that

$$\frac{d}{ds}\langle A\Xi(s;x,\xi),\,\Xi\rangle = 0,$$

and therefore in the elliptic case (i.e. $\langle A\xi,\xi\rangle > C|\xi|^2$ with $C>0$) we get that there is a constant $\gamma_0 > 0$ such that

$$\gamma_0|\xi_0|^2 \le |\Xi|^2 \le \gamma_0^{-1}|\xi_0|^2,$$

and $(X(s;x,\xi),\,\Xi(s;x,\xi))$ are globally defined. However, the situation in the non-elliptic case is different and the following properties have to be proved, see [22]:

a) global existence;

b) continuous dependence w.r.t. (x_0,ξ_0) with just a polynomial growth on $|x_0|$;

c) the trajectories are asymptotically free.

In order to obtain these three conditions we will need $A(x,t)$ to verify the hypotheses H1, H2, H3 and H4 given in Section 2 together with the following non-trapping condition.

H5. The solutions $(X,\,\Xi)$ of the hamiltonian flow H_A associated to A verifies that given M and $(x,\xi)\in\mathbb{R}^n\times(\mathbb{R}^n-\{0\})$ there is s_0 such that

$$|X(s;x,\zeta)| > M \quad \text{for all } s > s_0.$$

It is easy to justify that H5 is necessary because otherwise the integral given in (24) will not convergence even for compactly supported $\overrightarrow{b_1}$.

This non-trapping condition is not easy to verify, although it is obviously true in the constant coefficient case. Nevertheless it is a stable condition. In [22] the following lemma is proved.

Lemma 4.1. *Consider $A(x,t)=A(x)$ such that the hypotheses* H1, H2 *and* H5 *hold. Let $B(x)$ be an $n\times n$ real matrix with entries in the Schwartz class $\mathcal{S}(\mathbb{R}^n)$, and define $A_\delta(x)=A(x)+\delta B(x)$ with δ so small that there is a constant $\gamma_0>0$ such that for all $\xi\in\mathbb{R}^n$*

$$\frac{\gamma_0}{2}|\xi| \le |A_\delta\xi| \le 2\gamma_0^{-1}|\xi|.$$

Then there is $\delta_0>0$ such that A_δ is non-trapping for all $0<\delta<\delta_0$.

Remark. In the statement of the lemma above it is sufficient to assume that the entries of B have a finite number of derivatives with a finite power like decay.

So in the particular case of IVP (1) it will be sufficient to impose the non-trapping condition H5 to

$$A_{u_0}(x) = A(x,0,u_0,\bar{u}_0,\nabla u_0,\nabla\bar{u}_0). \tag{27}$$

5. The symbols

Once the hamiltonian flow is built we can construct the "integrating factor" given by the symbol (25). Nevertheless a new problem appears. This symbol is not in any known class of pseudo differential operators even if $A = \mathbb{I}_n$ and $X(x,\xi) = x + s\xi$, $\Xi = \xi$. A model example of a symbol which behaves as the one in (24) and it is easier to handle is, in dimension two,

$$k(x,\xi) = \psi(x \cdot (A\omega)^{\perp})\chi(|\xi|), \quad \omega = \frac{\xi}{|\xi|}, \quad \xi = (\xi_1,\xi_2), \quad \xi^{\perp} = (-\xi_2,\xi_1) \tag{28}$$

with $\psi \in \mathcal{S}(\mathbb{R})$ and χ regular, $\chi(0) = 0$ and $\chi(t) = 1$ if $t > 1$. We observe that

$$|\partial_\xi^\alpha k| \le \left(\frac{1+|x|}{1+|\xi|}\right)^{|\alpha|}, \tag{29}$$

being the growth in $|x|$ a problem to handle $k(x,\xi)$ (compare this situation with the one exhibited in (23)). In [22] we prove the following result.

Proposition 5.1. *Take $\overrightarrow{b_1}$ in the vector valued Schwartz class $\overrightarrow{\mathcal{S}} \in (\mathbb{R}^n)$ and*

$$\begin{aligned} k(x,\xi) &= \exp\left(\frac{1}{2}\int_0^\infty \overrightarrow{b_1}(x + s\tilde{\xi}) \cdot \xi \, ds\right)\chi(|\xi|), \\ \tilde{\xi} &= A\xi, \quad A = \begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & -\mathbb{I}_{n_2} \end{pmatrix}, \end{aligned} \tag{30}$$

with $\chi(|\xi|)$ as in (28)*. Then $\mathcal{K}(x,D)$ given in* (21) *is bounded from L^2 into L^2.*

This proposition and a more general result was proved if $A = \mathbb{I}_n$ (i.e. $n_2 = 0$) by Craig, Kappeler and Strauss [4]. A key part of their argument is the good behaviour of the radial derivatives of the symbol k given in (30). It is easy to check that

$$\left|\left(\frac{\xi}{|\xi|} \cdot \nabla_\xi\right)^\alpha k(x,\xi)\right| \le \frac{C_\alpha}{(1+|\xi|)^\alpha}. \tag{31}$$

This property is still true for general non-degenerate matrices A but it is not sufficient to prove Proposition 5.1. In fact in [4] another geometric assumption on the "essential" support of the symbol k is needed besides (31). This property does not hold when A is not the identity matrix, see [19] for a detailed discussion of this issue.

The L^2 estimate given above is not enough. Other results regarding the composition and the computation of the adjoints of these operators have to be proved in order to do the algebraic manipulations we exhibited in the one dimensional case. The results can be seen in [22].

Now it is time to recall that to simplify the exposition we assumed that $\varepsilon = 0$. To avoid this restriction some other properties about the symbols introduced in the above proposition are needed. These were proved in [23]. However the results in [22]

and [23], although sufficient for our purposes, are quite restrictive and the algebraic manipulations we can do with these operators are very rigid.

There is also another important constraint, which is that we are able to handle just symbols as (24) but not the general cases given in (25). The reader could then ask how we overcome this difficulty. The answer is in our previous work done on the I.V.P. (1) in [14] and [15]. That paper deals with non-linearities which are small perturbations of the constant coefficient case. Smallness allows hiding the first order terms thanks to some smoothing properties of the solutions of the corresponding free propagator. These properties are the subject of the next section.

6. The local smoothing

In [14] and [15] we proved that if the first order terms in (1) are small they can be handled by the so called local smoothing property of the free operators $e^{it(\Delta_{x_1}-\Delta_{x_2})}$. Notice that this family of operators is reversible in time and leave invariant the Hilbert space L^2. Moreover they commute with differentiation, and therefore the classical L^2-Sobolev spaces H^s of distributions with s derivatives in L^2 also remain invariant under the flow. For this reason there can not be any gain of global derivatives in L^2, because otherwise making the flow go backwards we would get a contradiction. It was proved by Kato in [12] and independently by Kruzhkov and Faminskii in [24] that the solutions of the Korteweg–de Vries (KdV) equation

$$\begin{cases} u_t + u_{xxx} + uu_x = 0, \quad u = u(x,t), \quad x \in \mathbb{R},\ t \in \mathbb{R}, \\ u(x,0) = 0 \end{cases}$$

gain for almost every time and locally in space one derivative in x with respect to the initial condition u_0. This "local smoothing" is a consequence of the dispersive character of the linear part of the KdV equation, and still holds if the non-linear term uu_x is removed.

It is well known that the free Schrödinger equation is also dispersive, and therefore it should have an analogous smoothing property. In that case the solution gains $1/2$-derivative locally in x and again for a.e. time with respect to the initial data. This property was established independently in [3], [31], and [35], [36], (see also [37]), in the elliptic setting, and in [13] for the general case.

However, this 1/2 gain is not sufficient to deal with first order terms which involve a full derivative. In [14] and [15] it is proved that the solution of

$$\begin{cases} i\partial_t u + \Delta_{x_1} u - \Delta_{x_2} u = F \\ u(x,0) = 0 \end{cases}$$

gains one full derivative with respect to the right-hand side F. The proof in [14] strongly uses the Fourier transform and is not adapted to the variable coefficient

situation necessary to treat (1). This was proved later on by S. Doi in [6], [7] and [8]. The results by Doi are remarkably robust, and for example, although he proves it for the elliptic case and for scalar equations, it can be extended without any difficulty to the non-elliptic setting and for systems; see [17], [22] and [23]. The idea is to construct a classical pseudo-differential operator $b(x, D)$ such that the commutator with the general $\mathcal{L} = \partial_j(a_{jk}\partial_k)$ satisfies

$$\begin{aligned}\int_0^T \langle i\,[b(x,D)\mathcal{L} - \mathcal{L}\,b(x,D)]\,u, u\rangle\,dt \\ \geq \frac{c_0}{2}\int_0^T \left\langle \frac{(1-\Delta)^{1/2}u}{1+|x|^2}, u\right\rangle dt - \frac{2}{c_0}T\sup_{0<t<T}\langle u,u\rangle,\end{aligned} \tag{32}$$

for some universal constant $c_0 > 0$. Here I am purposely using "universal" without giving a precise definition. The full argument given in [23] depends on this constant in a crucial way, and I refer to the reader to the introduction of that paper for a more precise statement. It is also important to notice that the above inequality is useful as long as there is a control on the L^2 norm given by the term $\langle u, u\rangle$.

7. The main result

Our main result in [23] is the following.

Theorem 7.1. *Under the hypotheses* H1–H4 *there exists* $N = N(n) \in \mathbb{Z}^+$ *such that given any*

$$u_0 \in H^s(\mathbb{R}^n) \quad \text{with } \langle x\rangle^N \partial_x^\alpha u_0 \in L^2(\mathbb{R}^n),\ |\alpha| \leq s_1, \tag{33}$$

$s, s_1 \in \mathbb{Z}^+$ *sufficiently large, and* $s > s_1 + 4$, *for which the hamiltonian flow* H_A *given in* (26) *associated to the quadratic form*

$$A = A_{u_0}(x,\xi) = \sum_{j.k=1}^n a_{jk}(x, 0, u_0, \bar{u}_0, \nabla u_0, \nabla\bar{u}_0)\xi_j\xi_k \tag{34}$$

is non-trapping, there exist $T_0 > 0$, *depending on*

$$\lambda = \|u_0\|_{s,2} + \sum_{|\alpha_1|\leq s_1} \left\|\langle x\rangle^N \partial_x^\alpha u_0\right\|_2$$

the constants in H1–H4 *and on the non-trapping condition* H5, *and a unique solution* $u = u(x,t)$ *of the equation* (1) *with initial data* $u(x,0) = u_0(x)$ *on the time interval* $[0, T_0]$ *satisfying*

$$\begin{aligned}&u \in C\big([0,T_0]\,:\,H^{s-1}\big) \cap L^\infty\big([0,T_0]\,:\,H^s\big) \cap C^1\big((0,T_0)\,:\,H^{s-3}\big),\\ &\langle x\rangle^N \partial_x^\alpha u \in C\left([0,T_0]\,:\,L^2\right), \quad |\alpha| \leq s_1.\end{aligned} \tag{35}$$

Moreover, if $u_0 \in H^{s'}(\mathbb{R}^n)$ *with* $s' > s$ *then* (35) *holds with* s' *instead of* s *in the same interval* $[0, T_0]$.

We have seen in Section 2 that we can obtain energy estimates as (7) if we are able to get rid of the first order terms, and that we can achieve that using an integrating factor. But also we pointed out at the end of Section 5 that the integrating factor we were able to construct was just for a hamiltonian flow which is free outside of a ball. However as we see in the statement of Theorem 7.1, this is not the case when H_A is given as in (35). The way to bypass this obstruction is to write

$$A = A_R + \left(\begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & -\mathbb{I}_{n_2} \end{pmatrix} - A_R \right)$$

with

$$A_R = \begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & -\mathbb{I}_{n_2} \end{pmatrix}$$

if $|x| \geq R$ and $R > 0$ is a large parameter to be fixed. The error terms created by this decomposition are of first order and can be done small by taking a large enough R. Then the local smoothing inequality (32) can be used as in [14] to control them. This creates the problem of how to handle

$$\sup_{0<t<T} \langle u, u \rangle.$$

But this quantity is precisely the one we started with when doing the energy estimate (7). Notice that in (32) appears multiplied by the factor T, and therefore it can be absorbed by the left-hand side of (7) by taking a small enough T, which closes the argument.

8. The elliptic case

In this section we give some comments that illustrate the substantial differences that appear when in (1) A is the identity matrix. To start with the main result in this case is far more general, and I refer to the reader to the work [17] for a precise statement; see also [29], [28], and [26].

The key difference is the existence of a fundamental argument due to Chihara, see [2], which goes as follows. He first writes (1) as a system in $(u, \bar{u})$. The problem of doing this is that Doi's trick, explained in Section 6, can not be carried out because the pseudo-differential operator $b(x, D)$ needed to prove the estimate (32) does not have the algebraic property (15). Therefore the good structure that the terms

$$\overrightarrow{b_2} \cdot \nabla \bar{u}$$

have for the integration by parts we exhibited in Section 2 is lost after applying the operator $b(x, D)$. Notice these bad terms are off the diagonal. The observation of Chihara is that the corresponding matrix can be easily diagonalized to the expense of zero order terms that as usual are harmless.

In order to explain how this diagonalization is done let us consider the model problem

$$\begin{cases} \partial_t u = i\Delta u + \overrightarrow{b_2}\cdot\nabla\bar{u}. \\ u(x,0) = u_0(x), \end{cases} \tag{36}$$

with $\overrightarrow{b_2} \in \mathbb{C}^n$ a constant vector.

As a system this equation is written as

$$\begin{pmatrix} u \\ \bar{u} \end{pmatrix}_t = \begin{pmatrix} i\Delta & \overrightarrow{b_2}\cdot\nabla \\ \overrightarrow{\bar{b}_2}\cdot\nabla & -i\Delta \end{pmatrix}\begin{pmatrix} u \\ \bar{u} \end{pmatrix}.$$

The eigenvalues of the above matrix are

$$\pm i\big(\Delta^2 - (\overrightarrow{b_2}\cdot\nabla)\big(\overrightarrow{\bar{b}_2}\cdot\nabla\big)\big)^{1/2}.$$

Therefore we are lead to consider the system

$$\begin{pmatrix} v \\ \bar{v} \end{pmatrix}_t = i\begin{pmatrix} \big(\Delta^2 - (\overrightarrow{b_2}\cdot\nabla)\big(\overrightarrow{\bar{b}_2}\cdot\nabla\big)\big)^{1/2} & 0 \\ 0 & -\big(\Delta^2 - (\overrightarrow{b_2}\cdot\nabla)\big(\overrightarrow{\bar{b}_2}\cdot\nabla\big)\big)^{1/2} \end{pmatrix}\begin{pmatrix} v \\ \bar{v} \end{pmatrix}.$$

Notice that

$$\big(\Delta^2 - (\overrightarrow{b_2}\cdot\nabla)\big(\overrightarrow{\bar{b}_2}\cdot\nabla\big)\big)^{1/2} = \Delta\big(1 - \Delta^{-2}(\overrightarrow{b_2}\cdot\nabla)\big(\overrightarrow{\bar{b}_2}\cdot\nabla\big)\big)^{1/2}.$$

The argument ends observing that

$$\Delta\big(1 - \Delta^{-2}(\overrightarrow{b_2}\cdot\nabla)\big(\overrightarrow{\bar{b}_2}\cdot\nabla\big)\big)^{1/2} = \Delta + \text{zero order}. \tag{37}$$

However, the identity (37) is false if the laplacian is changed by an operator of the type $\Delta_{x_1} - \Delta_{x_2}$ and therefore this trick does not work in that situation. Although we have oversimplified the problem considering $\overrightarrow{b_2} \in \mathbb{C}^n$ as a constant vector, the above computations can also be carried out without major difficulty for $\overrightarrow{b_2}(x)$ regular and bounded to the expense of creating zero order terms which as usual are harmless; see [20].

Another important difference of the elliptic setting is that perturbations of the type $\Delta\bar{u}$ are also allowed. The way of seeing this is by a diagonalization argument similar to the one we have just done. Consider $a > 0$ and $b \in \mathbb{C}$ and the equation

$$u_t = ia\Delta u + ib\Delta\bar{u}.$$

Then differentiating with respect to t on both sides we get

$$u_{tt} = (-a^2 + |b|^2)\Delta^2 u$$

which is well posed as long as

$$a^2 > |b|^2.$$

This elemental algebra is much more rigid in the non-elliptic setting and works only for some trivial cases; see [23].

9. Remarks on ultrahyperbolic operators

In this final section we gather some information and open questions about linear and non-linear perturbations of the free propagator

$$e^{it\mathcal{L}_0} \quad \text{with } \mathcal{L}_0 = \Delta_{x_1} - \Delta_{x_2}. \tag{38}$$

As we saw in Section 6 one of the fundamental properties of this flow is the local smoothing effect. In fact (32) was a key ingredient to overcome the loss of derivatives obstruction which was explained in Section 2. In order to use (32) one is lead to study the following maximal function

$$\sup_t |e^{it\mathcal{L}_0} u_0|^2, \tag{39}$$

which is defined for all $x \in \mathbb{R}^n$.

In our study of (1) we did not look at the question of which is the minimal regularity to be assumed on the initial condition so that the equation is solvable. This is something which strongly depends on the specific equation one is looking at. Therefore, and for our purposes, the necessary bounds for (39) are rather simple to obtain and there is no difference between the elliptic and the non-elliptic situation at that level. The situation is completely different when looking at a specific model as for example the Schrödinger map I mentioned in Section 3. In that case having sharp bounds for the maximal function can be very useful.

It has been recently proved in [30] that the maximal function given in (39) has a different behaviour for $\mathcal{L}_0$ than for the laplacian, being worse in the former case. Also and with respect to the local smoothing it is known that $\mathcal{L}_0$ is much more sensitive to first order perturbations than the laplacian. In [20] it is proved that the $\frac{1}{2}$ derivative gain of classical Schrödinger flows I mentioned in Section 6 can be reduced to just $\frac{1}{4}$ for $\mathcal{L}_0$.

Regarding non-linear perturbations very little is known about ill-posedness results. In fact and to the best of my knowledge the only one obtained in that direction is about the semilinear equation

$$\begin{cases} \frac{1}{i}\partial_t u = \Delta_{x_1} u - \Delta_{x_2} u \pm |u|^p u \\ u(x,0) = u_0, \end{cases} \tag{40}$$

and is given in [18] with $u_0 = c\delta$. There it is proved that (40) is ill-posed if $p \geq \frac{2}{n}$ (the proof is done for Δ but it works the same for $\mathcal{L}_0$).

The question of well-posedness of (40) is related to the existence of Strichartz estimates for the free propagator $e^{it\mathcal{L}_0}u_0$ with $u_0 \in L^2$. In that case it is well known that there is no difference between a general $\mathcal{L}_0$ and the laplacian. This type of estimates are very relevant in Harmonic Analysis in order to understand the restriction properties of the Fourier transform to curved surfaces. From that point of view it is

very natural to assume initial conditions u_0 such that its Fourier transform $\hat{u}_0$ is in L^p. A lot of progress has been done when $\mathcal{L}_0 = \Delta$ but as far as I know there are no results for general $\mathcal{L}_0$ for $p > 2$; see [33], [34], and [25].

Finally let us recall the pseudo-differential operators mentioned in Section 5. As I already said the calculus we develop in [22] and [23] is quite rudimentary, and I think there are many interesting properties to be understood. For example, we do not know if the inequality proved in Proposition 5.1 can be extended to L^p for $p \neq 2$. Another limitation of our approach is that we can construct the integrating factor only for hamiltonians which are free outside of a compact set. It should be enough to assume only that the hamiltonian satisfies hypothesis H2.

References

[1] Chang, N. H., Shatah, J., Uhlenbeck, K., Schrödinger maps. *Comm. Pure Appl. Math.* **53** (2000), 590–602.

[2] Chihara, H., Local existence for semilinear Schrödinger equations. *Math. Japon.* **42** (1995), 35–51.

[3] Constantin, P., and Saut, J. C., Local smoothing properties of dispersive equations. *J. Amer. Math. Soc.* **1** (1989), 413–446.

[4] Craig, W., Kappeler, T., and Strauss, W., Microlocal dispersive smoothing for the Schrödinger equation. *Comm. Pure Appl. Math.* **48** (1995), 769–860.

[5] Da Rios, L. S. On the motion of an unbounded fluid with a vortex filament of any shape. *Rend. Circ. Mat. Palermo* **22** (1906), 117–135 (in Italian).

[6] Doi, S., On the Cauchy problem for Schrödinger type equations and the regularity of solutions. *J. Math. Kyoto Univ.* **34** (1994), 319–328.

[7] Doi, S., Remarks on the Cauchy problem for Schrödinger–type equations. *Comm. Partial Differential Equations* **21** (1996), 163–178.

[8] Doi, S. , Smoothing effects for Schrödinger evolution equation and global behavior of geodesic flow. *Math. Ann.* **318** (2000), 355–389.

[9] Hasimoto, H. A soliton on a vortex filament. *J. Fluid Mech.* **51** (1972), 477–485.

[10] Hayashi, N., and Ozawa, T., Remarks on nonlinear Schrödinger equations in one space dimension. *Differential Integral Equations* **7** (1994), 453–461.

[11] Ichinose, W., On L^2 well-posedness of the Cauchy problem for Schrödinger type equations on a Riemannian manifold and Maslov theory. *Duke Math. J.* **56** (1988), 549–588.

[12] Kato, T., On the Cauchy problem for the (generalized) Korteweg-de Vries equation. In *Studies in applied mathematics* (ed. by V. Guillemin), Adv. Math. Suppl. Stud. 8, Academic Press, New York 1983, 93–128.

[13] Kenig, C. E., Ponce, G., and Vega, L., Oscillatory integrals and regularity of dispersive equations. *Indiana Univ. Math. J.* **40** (1991), 33–69.

[14] Kenig, C. E., Ponce, G., and Vega, L., Small solutions to nonlinear Schrödinger equations. *Ann. Inst. Henri Poincaré* **10** (1993), 255–288.

[15] Kenig, C. E., Ponce, G., and Vega, L., On the Zakharov and Zakharov-Schulman Systems. *J. Funct. Anal.* **127** (1995), 202–234.

[16] Kenig, C. E., Ponce, G., and Vega, L., Smoothing effects and local existence theory for the generalized nonlinear Schrödinger equations. *Invent. Math.* **134** (1998), 489–545.

[17] Kenig, C. E., Ponce, G, and Vega, L., The Cauchy problem for quasi-linear Schrödinger equations. *Invent. Math.* **158** (2004), 343–388.

[18] Kenig, C. E., Ponce, G, and Vega, L., On the ill-posedness of some canonical dispersive equations. *Duke Math. J* **106** (2001), 617–633.

[19] Kenig, C. E., Ponce, G, and Vega, L., On the Cauchy problem for linear Schrödinger systems with variable coefficient lower order terms. In *Harmonic analysis and number theory* (ed. by S. W. Drury and M. Ram Murty), CMS Conf. Proc. 21, Amer. Math. Soc., Providence, RI, 1997, 205–227.

[20] Kenig, C. E., Ponce, G, and Vega, L., On the smoothing properties of some dispersive hyperbolic systems. In *Nonlinear Waves* (Sapporo, 1995), GAKUTO Internat. Ser. Math. Sci. Appl. 10, Gakkōtosho, Tokyo 1997, 221–229.

[21] Kenig, C. E., Ponce, G, and Vega, L., The Initial Value Problem for the General Quasi-linear Schrödinger Equation. Preprint.

[22] Kenig, C. E., Ponce, G, and Vega, L., Variable coefficient Schrödinger flows for ultrahyperbolic operators. *Adv. Math.* **196** (2005), 373–486.

[23] Kenig, C. E., Ponce, G, and Vega, L., The general quasilinear ultrahyperbolic Schrödinger equation. *Adv. Math.*, to appear.

[24] Kruzhkov, D. J., and Faminskii, A. V., Generalized solutions for the Cauchy problem for the Korteweg-de Vries equation. *Math. USSR-Sb.* **48** (1990), 391–421.

[25] Lee, S., Bilinear restriction estimates for surfaces with curvatures of different signs. *Trans. Amer. Math. Soc.* **358** (2006) 3511–3533.

[26] Lim, W-K., and Ponce, G. On the initial value problem for the one dimensional quasilinear Schrödinger equation. *SIAM J. Math. Anal.* **34** (2003), 435–459.

[27] Mizohata, S., *On the Cauchy Problem*. Notes Rep. Math. Sci. Engrg. 3, Academic Press, Orlando, FL, 1985.

[28] Poppenberg, M. Smooth solutions for a class of fully nonlinear Schrödinger type equations. *Nonlinear Anal.* **45** (2001), 723–741.

[29] Rolvung, C., Non-isotropic Schrödinger equations. PhD. Dissertation, University of Chicago, 1998.

[30] Rogers, K., Vargas, A., Vega L., Pointwise convergence of solutions to the nonelliptic Schrödinger equation. *Indiana Univ. Math. J.*, to appear.

[31] Sjölin, P., Regularity of solutions to the Schrödinger equations. *Duke Math. J.* **55** (1987), 699–715.

[32] Sulem, C., Sulem P.L., *The Nonlinear Schrödinger Equation, Self-focusing and Wave Collapse*. Appl. Math. Sci. 139, Springer-Verlag, New York 1999.

[33] Tao, T., A Sharp bilinear restriction estimate for paraboloids. *Geom. Funct. Anal.* **13** (2003), 1359–1384.

[34] Vargas, A., Restriction theorems for a surface with negative curvature. *Math. Z.* 249 (2005), 97–111.

[35] Vega, L., El multiplicador de Schrödinger: la función maximal y los operadores de restricción. PhD Thesis, Universidad Autónoma de Madrid, 1988.

[36] Vega, L., The Schrödinger equation: pointwise convergence to the initial data. *Proc. Amer. Math. Soc.* **102** (1988), 874–878.

[37] Vega, L., Small perturbations of the free Schrödinger equation. In *Recent Advances in Partial Differential Equations* (ed. by M. A. Herrero, E. Zuazua), RAM Res. Appl. Math. 30, John Wiley & Sons, Chichester 1994, 115–130.

Departamento de Matemáticas, Universidad del País Vasco, Apdo. 644, 48080 Bilbao, Spain
E-mail: luis.vega@ehu.es