Title:	*Dynamics of the nonlinear Klein-Gordon equation in the nonrelativistic limit*
Journal Information:	*Annali di Matematica Pura ed Applicata (1923 -)*
Author(s):	S. Pasquali.
Volume, pages:	*903–972*, DOI:[www.doi.org/10.1007/s10231-018-0805-1]

Dynamics of the nonlinear Klein-Gordon equation in the nonrelativistic limit, I

S. Pasquali*

October 15, 2018

Abstract

The nonlinear Klein-Gordon (NLKG) equation on a manifold M in the nonrelativistic limit, namely as the speed of light c tends to infinity, is considered. In particular, a higher-order normalized approximation of NLKG (which corresponds to the NLS at order $r = 1$) is constructed, and when M is a smooth compact manifold or $\mathbb{R}^d$ it is proved that the solution of the approximating equation approximates the solution of the NLKG locally uniformly in time. When $M = \mathbb{R}^d$, $d \geq 2$, it is proved that solutions of the linearized order r normalized equation approximate solutions of linear Klein-Gordon equation up to times of order $\mathcal{O}(c^{2(r-1)})$ for any $r > 1$.
Keywords: nonrelativistic limit, nonlinear Klein-Gordon
MSC2010: 37K55, 70H08, 70K45, 81Q05

1 Introduction

In this paper the nonlinear Klein-Gordon (NLKG) equation in the nonrelativistic limit, namely as the speed of light c tends to infinity, is studied. Formal computations going back to the the first half of the last century suggest that, up to corrections of order $\mathcal{O}(c^{-2})$, the system should be described by the nonlinear Schrödinger (NLS) equation. Subsequent mathematical results have shown that the NLS describes the dynamics over time scales of order $\mathcal{O}(1)$.

The nonrelativistic limit for the Klein-Gordon equation on $\mathbb{R}^d$ has been extensively studied over more then 30 years, and essentially all the known results only show convergence of the solutions of NLKG to the solutions of the approximate equation for times of order $\mathcal{O}(1)$. The typical statement ensures convergence locally uniformly in time. In a first series of results (see [Tsu84], [Naj90] and [Mac01]) it was shown that, if the initial data are in a certain smoothness class, then the solutions converge in a weaker topology to the solutions of the approximating equation. These are informally called "results with loss of smoothness". Although in this paper a longer time convergence is proved, these results also fill in this group.

Some other results, essentially due to Machihara, Masmoudi, Nakanishi and Ozawa, ensure convergence without loss of regularity in the energy space, again over time scales of order $\mathcal{O}(1)$ (see [MNO02], [MN02] and [N$^+$08]).

* *Email:* `stefano.pasquali@unimi.it`

Concerning radiation solutions there is a remarkable result (see [Nak02]) by Nakanishi, who considered the complex NLKG in the defocusing case, in which it is known that all solutions scatter (and thus the scattering operator exists), and proved that the scattering operator of the NLKG equation converges to the scattering operator of the NLS. It is important to remark that this result is not contained in the one proved here and does not contain it.

Recently Lu and Zhang in [LZ16] proved a result which concerns the NLKG with a quadratic nonlinearity. Here the problem is that the typical scale over which the standard approach allows to control the dynamics is $\mathcal{O}(c^{-1})$, while the dynamics of the approximating equation takes place over time scales of order $\mathcal{O}(1)$. In that work the authors are able to use a normal form transformation (in a spirit quite different from ours) in order to extend the time of validity of the approximation over the $\mathcal{O}(1)$ time scale. We did not try to reproduce or extend that result.

In this paper some results for the dynamics of NLKG are obtained. Actually two kinds of results are proved: a global existence result for NLKG (see Theorem 2.1), uniform as $c \to \infty$, and approximation results (see Theorem 2.3 and Theorem 2.4) showing that solutions of NLKG can be approximated by solutions of suitable higher order NLS equations. Approximation results are different in the case where the equation lives on $\mathbb{R}^3$ or in a compact manifold: when M is a smooth compact manifold or $\mathbb{R}^d$ the solution of NLS approximates the solution of the original equation locally uniformly in time; when $M = \mathbb{R}^d$, $d \geq 2$, it is possible to prove that solutions of the *linearized* approximating equation approximate solutions of the linear Klein-Gordon equation up to times of order $\mathcal{O}(c^{2(r-1)})$, for any $r > 1$.

The present paper can be thought as an example in which techniques from canonical perturbation theory are used together with results from the theory of dispersive equations in order to understand the singular limit of some Hamiltonian PDEs. In this context, the nonrelativistic limit of the NLKG is a relevant example.

The issue of nonrelativistic limit has been studied also in the more general Maxwell-Klein-Gordon system ([BMS04], [MN03]), in the Klein-Gordon-Zakharov system ([MN08], [MN10]), in the Hartree equation ([CO06]) and in the pseudo-relativistic NLS ([CS16]). However, all these results proved the convergence of the solutions of the limiting system in the energy space ([CO06] studied also the convergence in H^k), *locally uniformly in time*; no information could be obtained about the convergence of solutions for longer (in the case of NLKG, that means c-dependent) timescales.

Other examples of singular perturbation problems that have been studied either with canonical perturbation theory or with other techniques (typically multiscale analysis) are the problem of the continuous approximation of lattice dynamics (see e.g. [BP06], [Sch10]) and the semiclassical analysis of Schrödinger operators (see e.g. [RT87], [AC07]). In the framework of lattice dynamics, the time scale covered by all known results is that typical of averaging theorems, which corresponds to our $\mathcal{O}(1)$ time scale. Hopefully the methods developed in the present paper could allow to extend the time of validity of those results.

The paper is organized as follows. In sect. 2 we state the results of the paper, together with some examples and comments. In sect. 3 we show Strichartz estimates for the linear KG equation and for the KG equation with potential, as well as a global existence result uniform with respect to c for the cubic NLKG equation on $\mathbb{R}^3$. In sect. 4 we state the main abstract result of the paper. In the subsequent sect. 5 we present the proof of the abstract result, which is based on a Galerkin averaging technique, along with some remarks and variant of the result. Next, in sect. 6 we apply the abstract theorem to the NLKG equation, making some explicit computations of the normal form at the first and at the second step. In the following sect. 7 we deduce some results about the approximation of solutions locally uniformly in time, while in sect. 8 we discuss the approximation for longer timescales: in particular, to deduce the latter we will exploit some dispersive properties of the KG equation reported in sect. 3. Finally, in Appendix A we will report some Birkhoff Normal Form estimates (the approach is essentially the same as in [Bam99]), and in Appendix B we will prove some interpolation theory results for relativistic Sobolev spaces, and we exploit them to deduce Strichartz estimates for the KG equation with potential.

Acknowledgments. This work is based on author's PhD thesis. He would like to express his thanks to his supervisor Professor Dario Bambusi.

2 Statement of the Main Results

The NLKG equation describes the motion of a spinless particle with mass $m > 0$. Consider first the real NLKG

$$\frac{\hbar^2}{2mc^2}u_{tt} - \frac{\hbar^2}{2m}\Delta u + \frac{mc^2}{2}u + \lambda|u|^{2(l-1)}u = 0, \tag{1}$$

where $c > 0$ is the speed of light, $\hbar > 0$ is the Planck constant, $\lambda \in \mathbb{R}$, $l \geq 2$, $c > 0$.

In the following $m = 1$, $\hbar = 1$. As anticipated above, one is interested in the behaviour of solutions as $c \to \infty$.

First it is convenient to reduce equation (1) to a first order system, by making the following symplectic change variables

$$\psi := \frac{1}{\sqrt{2}}\left[\left(\frac{\langle\nabla\rangle_c}{c}\right)^{1/2} u - i\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2} v\right], \;\; v = u_t/c^2,$$

where

$$\langle\nabla\rangle_c := (c^2 - \Delta)^{1/2}, \tag{2}$$

which reduces (1) to the form

$$-i\psi_t = c\langle\nabla\rangle_c\psi + \frac{\lambda}{2^l}\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right]^{2l-1}, \tag{3}$$

which is hamiltonian with Hamiltonian function given by

$$H(\bar{\psi},\psi) = \langle\bar{\psi}, c\langle\nabla\rangle_c\psi\rangle + \frac{\lambda}{2l}\int\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\frac{\psi+\bar{\psi}}{\sqrt{2}}\right]^{2l}\mathrm{d}x. \tag{4}$$

To state our first result, introduce for any $k \in \mathbb{R}$ and for any $1 < p < \infty$ the following relativistic Sobolev spaces

$$\mathscr{W}_c^{k,p}(\mathbb{R}^3) := \left\{ u \in L^p : \|u\|_{\mathscr{W}_c^{k,p}} := \left\| c^{-k} \langle \nabla \rangle_c^k u \right\|_{L^p} < +\infty \right\}, \tag{5}$$

$$\mathscr{H}_c^{k}(\mathbb{R}^3) := \left\{ u \in L^2 : \|u\|_{\mathscr{H}_c^{k}} := \left\| c^{-k} \langle \nabla \rangle_c^k u \right\|_{L^2} < +\infty \right\}, \tag{6}$$

and remark that the energy space is $\mathscr{H}_c^{1/2}$. Remark that for finite $c > 0$ such spaces coincide with the standard Sobolev spaces, while for $c = \infty$ they are equivalent to the Lebesgue spaces L^p.

In the following the notation $a \preceq b$ is used to mean: there exists a positive constant K that does not depend on c such that $a \leq Kb$.

We begin with a global existence result for the NLKG (3) in the cubic case, $l = 2$, for small initial data.

Theorem 2.1. *Consider Eq.* (3) *with $l = 2$ on $\mathbb{R}^3$.*
There exists $\epsilon_* > 0$ such that, if the norm of the initial datum ψ_0 fulfills

$$\|\psi_0\|_{\mathscr{H}_c^{1/2}} \leq \epsilon_*, \tag{7}$$

then the corresponding solution $\psi(t)$ of (3) *exists globally in time:*

$$\|\psi(t)\|_{L_t^\infty \mathscr{H}_c^{1/2}} \preceq \|\psi_0\|_{\mathscr{H}_c^{1/2}}, \tag{8}$$

All the constants do not depend on c.

Remark 2.2. *For finite c this is the standard result for small amplitude solution, while for $c = \infty$ it becomes the standard result for the NLS: thus Theorem 2.1 interpolates between these apparently completely different situations. Remark that the lack of a priori estimates for the solutions of NLKG in the limit $c \to \infty$ was the main obstruction in order to obtain global existence results uniform in c in standard Sobolev spaces.*

One is now interested in discussing the approximation of the solutions of NLKG with NLS-type equations. Before giving the result we describe the general strategy we use to get them.

Remark that Eq. (1) is Hamiltonian with Hamiltonian function (4). If one divides the Hamiltonian by a factor c^2 (which corresponds to a rescaling of time) and expands in powers of c^{-2} it takes the form

$$\langle \psi, \bar{\psi} \rangle + \frac{1}{c^2} P_c(\psi, \bar{\psi}) \tag{9}$$

with a suitable funtion P_c. One can notice that this Hamiltonian is a perturbation of $h_0 := \langle \psi, \bar{\psi} \rangle$, which is the generator of the standard Gauge transform, and which in particular admits a flow that is periodic in time. Thus the idea is to exploit canonical perturbation theory in order to conjugate such a Hamiltonian system to a system in normal form, up to remainders of order $\mathcal{O}(c^{-2r})$, for any given $r \geq 1$.

The problem is that the perturbation P_c has a vector field which is small only as an operator extracting derivatives. One can Taylor expand P_c and its vector field, but the number of derivatives extracted at each order increases. This situation is typical in singular perturbation problems. Problems of this kind have already been studied with canonical perturbation theory, but the price to pay to get a normal form is that the remainder of the perturbation turns out to be an operator that extracts a large number of derivatives.

In Sect. 6 the normal form equation is explicitly computed in the case $r = 2$:

$$\begin{aligned} -i\psi_t &= c^2\psi - \frac{1}{2}\Delta\psi + \frac{3}{4}\lambda|\psi|^2\psi \\ &+ \frac{1}{c^2}\left[\frac{51}{8}\lambda^2|\psi|^4\psi + \frac{3}{16}\lambda\left(2|\psi|^2\,\Delta\psi + \psi^2\Delta\bar{\psi} + \Delta(|\psi|^2\bar{\psi})\right) - \frac{1}{8}\Delta^2\psi\right], \end{aligned} \tag{10}$$

namely a singular perturbation of a Gauge-transformed NLS equation. If one, after a gauge transformation, only considers the first order terms, one has the NLS, for which radiation solution exist (for example in the defocusing case all solutions are of radiation type). For higher order NLS there are very few results (see for example [MS11]).

The standard way to exploit such a "singular" normal form is to use it just to construct some approximate solution of the original system, and then to apply Gronwall Lemma in order to estimate the difference with a true solution with the same initial datum (see for example [BCP02]).

This strategy works also here, but it only leads to a control of the solutions over times of order $\mathcal{O}(c^2)$. When scaled back to the physical time, this allows to justify the approximation of the solutions of NLKG by solutions of the NLS over time scales of order $\mathcal{O}(1)$, *on any manifold* admitting a Littlewood-Paley decomposition (such as Riemannian smooth compact manifolds, or $\mathbb{R}^d$; see the introduction of [Bou10] for the construction of Littlewood-Paley decomposition on manifolds).

Theorem 2.3. *Let M be a manifold which admits a Littlewood-Paley decomposition, and consider Eq.* (3) *on M.*

Fix $r \geq 1$, $R > 0$, $k_1 \gg 1$, $1 < p < +\infty$. Then $\exists\ k_0 = k_0(r) > 0$ with the following properties: for any $k \geq k_1$ there exists $c_{l,r,k,p,R} \gg 1$ such that for any $c > c_{l,r,k,p,R}$, if

$$\|\psi_0\|_{k+k_0,p} \leq R$$

and there exists $T = T_{r,k,p} > 0$ such that the solution ψ_r of the equation in normal form up to order r (96) *with initial datum ψ_0 satisfies*

$$\|\psi_r(t)\|_{k+k_0,p} \leq 2R, \ \ for \ \ 0 \leq t \leq T,$$

then

$$\|\psi(t) - \psi_r(t)\|_{k,p} \preceq \frac{1}{c^2}, \ \ for \ \ 0 \leq t \leq T. \tag{11}$$

where $\psi(t)$ is the solution of (3) *with initial datum ψ_0.*

A similar result has been obtained for the case $M = \mathbb{T}^d$ by Faou and Schratz, who aimed to construct numerical schemes which are robust in the nonrelativistic limit (see [FS14]; see also [BD12], [BZ16] and to [BFS16] for some numerical analysis of the nonrelativistic limit of the NLKG).

The idea one uses here in order to improve the time scale of the result is that of substituting Gronwall Lemma with a more sophisticated tool, namely dispersive estimates and the retarded Strichartz estimate. This can be done each time one can prove a dispersive or a Strichartz estimate (in the spaces $\mathscr{W}_c^{k,p}$ or $W^{k,p}$) for the linearization of equation (3) on the approximate solution uniformly in c.

It turns out that this is a quite hard task, and we were able to accomplish it only for the linear KG equation on $\mathbb{R}^d$. In order to state our approximation result, we consider the approximate equation given by the Hamilton equations of the normal form truncated at order $\mathcal{O}(c^{-2r})$, and let ψ_r be a solution of such a linearized normal form equation.

Theorem 2.4. *Consider* (1) *on* $\mathbb{R}^d$*,* $d \geq 2$*. Fix* $r \geq 1$ *and* $k_1 \gg 1$*. Then* $\exists$ $k_0 = k_0(r) > 0$ *such that for any* $k \geq k_1$*, if we denote by* ψ_r *the solution of the linearized normal equation* (103) *with initial datum* $\psi_0 \in H^{k+k_0}$ *and by* ψ *the solution of the linear KG equation* (12) *with the same initial datum, then there exists* $c^* := c^*(r,k) > 0$ *such that for any* $c > c^*$

$$\sup_{t\in[0,T]} \|\psi(t) - \psi_r(t)\|_{H^k_x} \preceq \frac{1}{c^2}, \quad T \preceq c^{2(r-1)}.$$

This result has been proved in the case $r = 1$ in Appendix A of [CLM15].

In order to approximate small radiation solutions of the NLKG equation, we would need to use dispersive estimates for the normal form equation, which unfortunately are not present in the literature. We defer this problem to a future work.

There are other well known solutions of NLS which would be interesting to study; indeed, it is well known that in the case of mixed-type nonlinearity

$$i\psi_t = -\Delta\psi - (|\psi|^2 - |\psi|^4)\psi,$$

such an equation admits linearly stable solitary wave solutions; it can also be proved that the standing waves of NLS can be modified in order to obtain standing wave solutions of the normal form of order r, for any r. It would be of clear interest to prove that true solutions starting close to such standing wave remain close to them for long times (remark that the NLKG does not admit stable standing wave solutions, see [OT07]); in order to get such a result one should prove a Strichartz estimate for NLKG close to the approximate solution and uniformly in c.

Before closing the subsection, a few technical comments: the first one is that here we develop normal form in the framework of the spaces $W^{k,p}$, while known results in Galerkin averaging theory only allow to deal with the spaces H^k. This is due to the fact that the Fourier analysis is used in order to approximate the derivatives operators with bounded operators. Thus the first technical step needed in order to be able to exploit dispersion is to reformulate Galerkin averaging theory in terms of dyadic decompositions. This is done in Theorem 4.3.

3 Dispersive properties of the Klein-Gordon equation

We briefly recall some classical notion of Fourier analysis on $\mathbb{R}^d$. Recall the definition of the space of Schwartz (or rapidly decreasing) functions,

$$\mathcal{S} := \{f \in C^\infty(\mathbb{R}^d, \mathbb{R}) | \sup_{x \in \mathbb{R}^d} (1+|x|^2)^{\alpha/2} |\partial^\beta f(x)| < +\infty, \ \ \forall \alpha \in \mathbb{N}^d, \forall \beta \in \mathbb{N}^d\}.$$

In the following $\langle x \rangle := (1+|x|^2)^{1/2}$.
Now, for any $f \in \mathcal{S}$ the *Fourier transform* of f, $\hat{f} : \mathbb{R}^d \to \mathbb{R}$, is defined by the following formula

$$\hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-i\langle x, \xi \rangle} \mathrm{d}x, \ \ \forall \xi \in \mathbb{R}^d,$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^d$.

At the beginning we will obtain Strichartz estimates for the linear equation

$$-i\,\psi_t \ = \ c \langle \nabla \rangle_c \,\psi, \ \ x \in \mathbb{R}^3. \tag{12}$$

Proposition 3.1. *For any Schrödinger admissible couples (p,q) and (r,s), namely such that*

$$\begin{gathered} 2 \leq p, r \leq \infty, \\ 2 \leq q, s \leq 6, \\ \frac{2}{p} + \frac{3}{q} = \frac{3}{2}, \frac{2}{r} + \frac{3}{s} = \frac{3}{2}, \end{gathered}$$

one has

$$\| \langle \nabla \rangle_c^{\frac{1}{q}-\frac{1}{p}} \ e^{it\,c\langle \nabla \rangle_c} \ \psi_0 \|_{L^p_t L^q_x} \ \preceq \ c^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}} \ \| \langle \nabla \rangle_c^{1/2} \psi_0 \|_{L^2}, \tag{13}$$

$$\left\| \langle \nabla \rangle_c^{\frac{1}{q}-\frac{1}{p}} \ \int_0^t e^{i(t-s)\ c \langle \nabla \rangle_c} \ F(s)\ \mathrm{d}s \right\|_{L^p_t L^q_x} \ \preceq \ c^{\frac{1}{q}-\frac{1}{p}+\frac{1}{s}-\frac{1}{r}-1} \ \| \langle \nabla \rangle_c^{\frac{1}{r}-\frac{1}{s}+1} F \|_{L^{r'}_t L^{s'}_x}. \tag{14}$$

Remark 3.2. *The above result can be easily generalized to the d-dimensional case, $d \geq 2$, if we consider (p,q) and (r,s) such that*

$$\begin{gathered} 2 \leq p, r \leq \infty, \\ 2 \leq q, s \leq \frac{2d}{d-2}, \\ \frac{2}{p} + \frac{d}{q} = \frac{d}{2}, \frac{2}{r} + \frac{d}{s} = \frac{d}{2}, \\ (p,q,d), (r,s,d) \neq (2, +\infty, 2), \end{gathered}$$

Remark 3.3. *By choosing $p = +\infty$ and $q = 2$, we get the following a priori estimate for finite energy solutions of* (12),

$$\|c^{1/2}\langle\nabla\rangle_c^{1/2}\, e^{it\,c\langle\nabla\rangle_c}\, \psi_0\|_{L_t^\infty L_x^2} \;\preceq\; \|c^{1/2}\langle\nabla\rangle_c^{1/2}\psi_0\|_{L^2}.$$

We also point out that, since the operators $\langle\nabla\rangle$ and $\langle\nabla\rangle_c$ commute, the above estimates in the spaces $L_t^p L_x^q$ extend to estimates in $L_t^p W_x^{k,q}$ for any $k \geq 0$.

Proof. We recall a result reported by D'Ancona-Fanelli in [DF08] for the operator $\langle\nabla\rangle := \langle\nabla\rangle_1$.

Lemma 3.4. *For all (p, q) Schrödinger-admissible exponents (ie, s.t. $\frac{2}{p}+\frac{3}{q} = \frac{3}{2}$)*

$$\|e^{i\tau\,\langle\nabla\rangle}\, \phi_0\|_{L_\tau^p\, W_y^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2},q}} = \|\langle\nabla\rangle^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}}\, e^{it\,\langle\nabla\rangle}\, \phi_0\|_{L_\tau^p\, L_y^q} \;\leq\; \|\phi_0\|_{L_y^2}.$$

Now, the solution of equation (12) satifies $\hat{\psi}(t,\xi) = e^{ic\langle\xi\rangle_c t}\hat{\psi}_0(\xi)$. We then define $\eta := \xi/c$, in order to have that

$$\hat{\phi}(c^2 t, \eta) := \hat{\psi}(t, c\eta) = \hat{\psi}(t, \xi),$$

and in particular that $\hat{\phi}_0(\eta) = \hat{\psi}_0(\xi)$.
Since

$$\langle\xi\rangle_c = \sqrt{c^2 + |\xi|^2} = c\sqrt{1 + |\xi|^2/c^2}, \tag{15}$$

we get

$$\begin{aligned}\hat{\phi}(t,\eta) &= e^{it\,c^2\langle\xi/c\rangle}\,\hat{\phi}_0(\xi/c)\\ &= e^{i\,tc^2\,\langle\eta\rangle}\,\hat{\phi}_0(\eta)\\ &= e^{i\,\tau\,\langle\eta\rangle}\,\hat{\phi}_0(\eta)\end{aligned}$$

if we set $\tau := c^2 t$. Now, by setting $y := cx$ a simple scaling argument leads to

$$\|e^{i\tau\,\langle\nabla\rangle}\, \phi_0\|_{L_\tau^p\, L_y^q} \;\preceq\; \|\langle\nabla\rangle^{\frac{1}{p}-\frac{1}{q}+\frac{1}{2}}\, \phi_0\|_{L^2} \;=\; \|\,\langle\eta\rangle^{\frac{1}{p}-\frac{1}{q}+\frac{1}{2}}\, \hat{\phi}_0\|_{L^2}$$

and since

$$\begin{aligned}\|\,\langle\eta\rangle^k\, \hat{\phi}_0\|_{L^2}^2 \;&=\; \int_{\mathbb{R}^3} \langle\eta\rangle^{2k}\; |\hat{\phi}_0(\eta)|^2\, \mathrm{d}\eta\\ &=\; \int_{\mathbb{R}^3} \left\langle\frac{\xi}{c}\right\rangle^{2k}\; |\hat{\phi}_0(\eta/c)|^2\, \frac{\mathrm{d}\xi}{c^3} \;=\; \frac{1}{c^{2k+3}} \int_{\mathbb{R}^3} \langle\xi\rangle_c^{2k}\; |\hat{\psi}_0(\xi)|^2\, \mathrm{d}\xi,\end{aligned}$$

we get

$$\|\,\langle\eta\rangle^{\frac{1}{p}-\frac{1}{q}+\frac{1}{2}}\, \hat{\phi}_0\|_{L^2} \;=\; \frac{1}{c^{\frac{3}{2}-\frac{1}{q}+\frac{1}{p}+\frac{1}{2}}}\; \|\langle\nabla\rangle_c^{\frac{1}{p}-\frac{1}{q}+\frac{1}{2}}\; \psi_0\|_{L^2}, \tag{16}$$

while on the other hand

$$\begin{aligned}\psi(t,x) &= (2\pi)^{-d/2}\int_{\mathbb{R}^3} e^{i\langle\xi,x\rangle}\ \hat{\psi}(t,\xi)\,\mathrm{d}\xi\ = (2\pi)^{-d/2}\int_{\mathbb{R}^3} e^{i\langle\eta,cx\rangle}\ \hat{\psi}(t,c\eta)\ c^3\mathrm{d}\eta\\ &= (2\pi)^{-d/2}\ c^3\int_{\mathbb{R}^3} e^{i\langle\eta,cx\rangle}\ \hat{\phi}(c^2t,\eta)\,\mathrm{d}\eta\ = c^3\ \phi(c^2t,cx),\end{aligned}$$

yields

$$\|\psi\|_{L^p_tL^q_x}\ =\ c^{3-\ 3/q-\ 2/p}\ \|\phi\|_{L^p_\tau L^q_y}. \tag{17}$$

Hence we can deduce (13); via a scaling argument we can also deduce (14). □

One important application of the Strichartz estimates for the free Klein-Gordon equation is Theorem 2.1, namely a global existence result *uniform with respect to* c for the NLKG equation (3) with cubic nonlinearity (this means $l=2$), with small initial data.

Proof of Theorem 2.1. It just suffices to apply Duhamel formula,

$$\psi(t)=e^{itc\nabla_c}\psi_0+i\frac{\lambda}{2^l}\int_0^t e^{i(t-s)c\nabla_c}\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right]^{2l-1},$$

and Proposition 3.1 with $p=+\infty$, in order to get that

$$\|\psi(t)\|_{L^\infty_t\mathscr{H}^{1/2}_c}\preceq\|\psi_0\|_{\mathscr{H}^{1/2}_c}+c^{1/s-1/r}\left\|\nabla_c^{1/r-1/s}\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right]^3\right\|_{L^{r'}_tL^{s'}_x},$$

but by choosing $r=+\infty$ and by Hölder inequality we get

$$\begin{aligned}\|\psi(t)\|_{L^\infty_t\mathscr{H}^{1/2}_c}&\preceq\|\psi_0\|_{\mathscr{H}^{1/2}_c}+\left\|\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right]^3\right\|_{L^1_tL^2_x}\\ &\preceq\|\psi_0\|_{\mathscr{H}^{1/2}_c}+\left\|\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right]^2\right\|_{L^1_tL^3_x}\left\|\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right\|_{L^\infty_tL^6_x}\\ &\preceq\|\psi_0\|_{\mathscr{H}^{1/2}_c}+\left\|\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right\|^2_{L^2_tL^6_x}\left\|\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right\|_{L^\infty_tL^6_x}\\ &\preceq\|\psi_0\|_{\mathscr{H}^{1/2}_c}+\|\psi\|^2_{L^2_t\mathscr{W}_c^{-1/2,6}}\|\psi\|_{L^\infty_t\mathscr{W}_c^{-1/2,6}}\\ &\preceq\|\psi_0\|_{\mathscr{H}^{1/2}_c}+\|\psi\|^2_{L^2_t\mathscr{W}_c^{-1/3,6}}\|\psi\|_{L^\infty_t\mathscr{H}^{1/2}_c},\end{aligned}$$

and one can conclude by a standard continuation argument. □

We also give a formulation of the Kato-Ponce inequality for the relativistic Sobolev spaces.

Proposition 3.5. *Let $f, g \in \mathcal{S}(\mathbb{R}^3)$, and let $c > 0$, $1 < r < \infty$ and $k \geq 0$. Then*

$$\|f\, g\|_{\mathscr{W}_c^{k,r}} \preceq \|f\|_{\mathscr{W}_c^{k,r_1}} \|g\|_{L^{r_2}} + \|f\|_{L^{r_3}} \|g\|_{\mathscr{W}_c^{k,r_4}}, \tag{18}$$

with

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r_3} + \frac{1}{r_4}, \quad 1 < r_1, r_4 < +\infty.$$

Remark 3.6. *For $c = 1$ Eq.* (18) *reduces to the classical Kato-Ponce inequality.*

Proof. We follow an argument by Cordero and Zucco (see Theorem 2.3 in [CZ11]).
We introduce the dilation operator $S_c(f)(x) := f(x/c)$, for any $c > 0$.
Then we apply the classical Kato-Ponce inequality to the rescaled product $S_c(fg) = S_c(f)\, S_c(g)$,

$$\|S_c(fg)\|_{W^{k,r}} \preceq \|S_c(f)\|_{W^{k,r_1}} \|S_c(g)\|_{L^{r_2}} + \|S_c(f)\|_{L^{r_3}} \|S_c(g)\|_{W^{k,r_4}}, \tag{19}$$

where

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r_3} + \frac{1}{r_4}, \quad 1 < r_1, r_4 < +\infty.$$

Now, combining the commutativity property

$$\langle\nabla\rangle^k S_c(f)(x) = c^{-k} S_c(\langle\nabla\rangle_c^k\, f)(x),$$

with the equality $\|S_c(f)\|_{L^r} = c^{-3/r}\|f\|_{L^r}$, we can rewrite (19) as

$$\|\langle\nabla\rangle^k (f\, g)\|_{L^r} \preceq \|\langle\nabla\rangle^k f\|_{L^{r_1}} \|g\|_{L^{r_2}} + \|f\|_{L^{r_3}} \|\langle\nabla\rangle^k g\|_{L^{r_4}},$$

and this leads to the thesis. □

We conclude with another dispersive result, which could be interesting in itself: by exploiting the boundedness of the wave operators for the Schrödinger equation, we can deduce Strichartz estimates for the KG equation with potential.

Theorem 3.7. *Let $c \geq 1$, and consider the operator*

$$\mathcal{H}(x) := c(c^2 - \Delta + V(x))^{1/2} = \mathcal{H}_0(1 + \langle\nabla\rangle_c^{-2} V)^{1/2}, \tag{20}$$

where $V \in C(\mathbb{R}^3, \mathbb{R})$ is a potential such that

$$|V(x)| + |\nabla V(x)| \preceq \langle x\rangle^{-\beta}, \quad x \in \mathbb{R}^3,$$

for some $\beta > 5$, and that 0 is neither an eigenvalue nor a resonance for the operator $-\Delta + V(x)$. Let (p, q) be a Schrödinger admissible couple, and assume that $\psi_0 \in \langle\nabla\rangle_c^{-1/2} L^2$ is orthogonal to the bound states of $-\Delta + V(x)$. Then

$$\|\langle\nabla\rangle_c^{\frac{1}{q}-\frac{1}{p}}\, e^{it\mathcal{H}(x)} \psi_0\|_{L_t^p L_x^q} \preceq c^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}} \|\langle\nabla\rangle_c^{1/2}\, \psi_0\|_{L^2}. \tag{21}$$

In order to prove Theorem 3.7 we recall Yajima's result on wave operators [Yaj95] (where we denote by $P_c(-\Delta+V)$ the projection onto the continuous spectrum of the operator $-\Delta+V$).

Theorem 3.8. *Assume that*

- *0 is neither an eigenvalue nor a resonance for* $-\Delta+V$*;*

- $|\partial^\alpha V(x)| \preceq \langle x\rangle^{-\beta}$ *for* $|\alpha|\le k$*, for some* $\beta>5$*.*

Consider the strong limits

$$\mathcal{W}_\pm := \lim_{t\to\pm\infty} e^{it(-\Delta+V)}e^{it\Delta}, \;\; \mathcal{Z}_\pm := \lim_{t\to\pm\infty} e^{-it\Delta}e^{it(\Delta-V)}P_c(-\Delta+V).$$

Then $\mathcal{W}_\pm : L^2 \to P_c(-\Delta+V)L^2$ *are isomorphic isometries which extend into isomorphisms* $\mathcal{W}_\pm : W^{k,p} \to P_c(-\Delta+V)W^{k,p}$ *for all* $p\in[1,+\infty]$*, with inverses* $\mathcal{Z}_\pm$*. Furthermore, for any Borel function* $f(\cdot)$ *we have*

$$f(-\Delta+V)P_c(-\Delta+V) = \mathcal{W}_\pm f(-\Delta)\mathcal{Z}_\pm, \;\; f(-\Delta) = \mathcal{Z}_\pm f(-\Delta+V)P_c(-\Delta+V)\mathcal{W}_\pm. \tag{22}$$

Now, in the case $c=1$ one can derive Strichartz estimates for $\mathcal{H}(x)$ from the Strichartz estimates for the free KG equation, just by applying the aforementioned Theorem by Yajima in the case $k=1$ (since $1/p-1/q+1/2\in[0,5/6]$ for all Schrödinger admissible couples (p,q)). This was already proved in [BC11] (see Lemma 6.3). In the general case, this will follow from an interpolation theory argument, and we defer it to Appendix B.

4 Galerkin Averaging Method

Consider the scale of Banach spaces $W^{k,p}(M,\mathbb{C}^n\times\mathbb{C}^n)\ni(\psi,\bar\psi)$ ($k\ge1$, $1<p<+\infty$, $n\in\mathbb{N}_0$) endowed by the standard symplectic form. Having fixed k and p, and $U_{k,p}\subset W^{k,p}$ open, we define the gradient of $H\in C^\infty(U_{k,p},\mathbb{R})$ w.r.t. $\bar\psi$ as the unique function s.t.

$$\left\langle \nabla_{\bar\psi}H, \bar h\right\rangle = \mathrm{d}_{\bar\psi}H\bar h, \;\; \forall h\in W^{k,p},$$

so that the Hamiltonian vector field of a Hamiltonian function H is given by

$$X_H(\psi,\bar\psi) = (i\nabla_{\bar\psi}H, \; -i\nabla_\psi H).$$

The open ball of radius R and center 0 in $W^{k,p}$ will be denoted by $B_{k,p}(R)$.

Now, we call an *admissible family of cut-off (pseudo-differential) operators* a sequence $(\pi_j(D))_{j\ge0}$, where $\pi_j(D): W^{k,p}\to W^{k,p}$ for any $j\ge0$, such that

- for any $j\ge0$ and for any $f\in W^{k,p}$

$$f = \sum_{j\ge0}\pi_j(D)f;$$

- for any $j \geq 0$ $\pi_j(D)$ can be extended to a self-adjoint operator on L^2, and there exist constants K_1, $K_2 > 0$ such that

$$K_1 \left(\sum_{j\geq 0} \|\pi_j(D)f\|^2_{L^2} \right)^{1/2} \leq \|f\|_{L^2} \leq K_2 \left(\sum_{j\geq 0} \|\pi_j(D)f\|^2_{L^2} \right)^{1/2} ;$$

- for any $j \geq 0$, if we denote by $\Pi_j(D) := \sum_{l=0}^{j} \pi_l(D)$, there exist positive constants K', (possibly depending on k and p) such that

$$\|\Pi_j f\|_{k,p} \leq K' \, \|f\|_{k,p} \;\; \forall f \in W^{k,p};$$

- there exist positive constants K''_1, K''_2 (possibly depending on k and p) and an increasing and unbounded sequence $(K_j)_{j\in\mathbb{N}} \subset \mathbb{R}_+$ such that

$$K''_1 \|f\|_{W^{k,p}} \leq \left\| \left[\sum_{j\in\mathbb{N}} K_j^{2k} |\pi_j(D)f|^2 \right]^{1/2} \right\|_{L^p} \leq K''_2 \|f\|_{W^{k,p}}. \tag{23}$$

Remark 4.1. *Let $k \geq 0$, M be either $\mathbb{R}^d$ or the d-dimensional torus $\mathbb{T}^d$, and consider the Sobolev space $H^k = H^k(M)$. One can readily check that Fourier projection operators on H^k*

$$\pi_j \psi(x) := (2\pi)^{-d/2} \int_{j-1\leq |k| \leq j} \hat{\psi}(k) e^{ik\cdot x} \mathrm{d}k, \;\; j \geq 1$$

form an admissible family of cut-off operators. In this case we have

$$\Pi_N \psi(x) := (2\pi)^{-d/2} \int_{|k| \leq N} \hat{\psi}(k) e^{ik\cdot x} \mathrm{d}k, \;\; N \geq 0,$$

and the constants $(K_j)_{j\in\mathbb{N}}$ in (23) *are given by $K_j := j$.*

Remark 4.2. *Let $k \geq 0$, $1 < p < +\infty$, we now introduce the Littlewood-Paley decomposition on the Sobolev space $W^{k,p} = W^{k,p}(\mathbb{R}^d)$ (see [Tay11], Ch. 13.5).*

In order to do this, define the cutoff operators in $W^{k,p}$ in the following way: start with a smooth, radial nonnegative function $\phi_0 : \mathbb{R}^d \to \mathbb{R}$ such that $\phi_0(\xi) = 1$ for $|\xi| \leq 1/2$, and $\phi_0(\xi) = 0$ for $|\xi| \geq 1$; then define $\phi_1(\xi) := \phi_0(\xi/2) - \phi_0(\xi)$, and set

$$\phi_j(\xi) := \phi_1(2^{1-j}\xi), \;\; j \geq 2. \tag{24}$$

Then $(\phi_j)_{j\geq 0}$ is a partition of unity,

$$\sum_{j\geq 0} \phi_j(\xi) = 1.$$

Now, for each $j \in \mathbb{N}$ and each $f \in W^{k,2}$, we can define $\phi_j(D)f$ by

$$\mathcal{F}(\phi_j(D)f)(\xi) := \phi_j(\xi)\hat{f}(\xi).$$

It is well known that for $p \in (1,+\infty)$ the map $\Phi : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d, l^2)$,

$$\Phi(f) := (\phi_j(D)f)_{j\in\mathbb{N}},$$

maps $L^p(\mathbb{R}^d)$ isomorphically onto a closed subspace of $L^p(\mathbb{R}^d, l^2)$, and we have compatibility of norms ([Tay11], Ch. 13.5, (5.45)-(5.46)),

$$K'_p\|f\|_{L^p} \leq \|\Phi(f)\|_{L^p(\mathbb{R}^d,l^2)} := \left\| \left[\sum_{j\in\mathbb{N}} |\phi_j(D)f|^2 \right]^{1/2} \right\|_{L^p} \leq K_p\|f\|_{L^p},$$

and similarly for the $W^{k,p}$-norm, i.e. for any $k > 0$ and $p \in (1,+\infty)$

$$K'_{k,p}\|f\|_{W^{k,p}} \leq \left\| \left[\sum_{j\in\mathbb{N}} 2^{2jk}|\phi_j(D)f|^2 \right]^{1/2} \right\|_{L^p} \leq K_{k,p}\|f\|_{W^{k,p}}. \tag{25}$$

We then define the cutoff operator Π_N by

$$\Pi_N\psi := \sum_{j\leq N} \phi_j(D)\psi. \tag{26}$$

Hence, according to the above definition, the sequence $(\phi_j(D))_{j\geq 0}$ is an admissible family of cut-off operators.
We point out that the Littlewood-Paley decomposition, along with equality (25), can be extended to compact manifolds (see [BGT04]), as well as to some particular non-compact manifolds (see [Bou10]).

Now we consider a Hamiltonian system of the form

$$H = h_0 + \epsilon\, h + \epsilon\, F, \tag{27}$$

where $\epsilon > 0$ is a parameter. We fix an admissible family of cut-off operators $(\pi_j(D))_{j\geq 0}$ on $W^{k,p}(\mathbb{R}^d)$. We assume that

PER h_0 generates a linear periodic flow Φ^t with period 2π,

$$\Phi^{t+2\pi} = \Phi^t \quad \forall t.$$

We also assume that Φ^t is analytic from $W^{k,p}$ to itself for any $k \geq 1$, and for any $p \in (1,+\infty)$;

INV for any $k \geq 1$, for any $p \in (1,+\infty)$, Φ^t leaves invariant the space $\Pi_j W^{k,p}$ for any $j \geq 0$. Furthermore, for any $j \geq 0$

$$\pi_j(D) \circ \Phi^t = \Phi^t \circ \pi_j(D);$$

NF h is in normal form, namely

$$h \circ \Phi^t = h.$$

Next we assume that both the Hamiltonian and the vector field of both h and F admit an asymptotic expansion in ϵ of the form

$$h \sim \sum_{j\geq 1} \epsilon^{j-1} h_j, \;\; F \sim \sum_{j\geq 1} \epsilon^{j-1} F_j, \tag{28}$$

$$X_h \sim \sum_{j\geq 1} \epsilon^{j-1} X_{h_j}, \;\; X_F \sim \sum_{j\geq 1} \epsilon^{j-1} X_{F_j}, \tag{29}$$

and that the following properties are satisfied

HVF There exists $R^* > 0$ such that for any $j \geq 1$

- X_{h_j} is analytic from $B_{k+2j,p}(R^*)$ to $W^{k,p}$;
- X_{F_j} is analytic from $B_{k+2(j-1),p}(R^*)$ to $W^{k,p}$.

Moreover, for any $r \geq 1$ we have that

- $X_{h-\sum_{j=1}^{r} \epsilon^{j-1} h_j}$ is analytic from $B_{k+2(r+1),p}(R^*)$ to $W^{k,p}$;
- $X_{F-\sum_{j=1}^{r} \epsilon^{j-1} F_j}$ is analytic from $B_{k+2r,p}(R^*)$ to $W^{k,p}$.

The main result of this section is the following theorem.

Theorem 4.3. *Fix $r \geq 1$, $R > 0$, $k_1 \gg 1$, $1 < p < +\infty$. Consider (27), and assume PER, INV (with respect to the Littlewood-Paley decomposition), NF and HVF. Then $\exists\, k_0 = k_0(r) > 0$ with the following properties: for any $k \geq k_1$ there exists $\epsilon_{r,k,p} \ll 1$ such that for any $\epsilon < \epsilon_{r,k,p}$ there exists $\mathcal{T}_\epsilon^{(r)} : B_{k,p}(R) \to B_{k,p}(2R)$ analytic canonical transformation such that*

$$H_r := H \circ \mathcal{T}_\epsilon^{(r)} = h_0 + \sum_{j=1}^{r} \epsilon^j \mathcal{Z}_j + \epsilon^{r+1} \mathcal{R}^{(r)},$$

where $\mathcal{Z}_j$ are in normal form, namely

$$\{\mathcal{Z}_j, h_0\} = 0, \tag{30}$$

and

$$\sup_{B_{k+k_0,p}(R)} \|X_{\mathcal{Z}_j}\|_{W^{k,p}} \leq C_{k,p},$$

$$\sup_{B_{k+k_0,p}(R)} \|X_{\mathcal{R}^{(r)}}\|_{W^{k,p}} \leq C_{k,p}, \tag{31}$$

$$\sup_{B_{k,p}(R)} \|\mathcal{T}_\epsilon^{(r)} - id\|_{W^{k,p}} \leq C_{k,p}\, \epsilon. \tag{32}$$

In particular, we have that

$$\mathcal{Z}_1(\psi,\bar\psi) = h_1(\psi,\bar\psi) + \langle F_1\rangle(\psi,\bar\psi),$$

where $\langle F_1\rangle(\psi,\bar\psi) := \int_0^{2\pi} F_1 \circ \Phi^t(\psi,\bar\psi) \frac{\mathrm{d}t}{2\pi}$.

5 Proof of Theorem 4.3

We first make a Galerkin cutoff through the Littlewood-Paley decomposition (see [Tay11], Ch. 13.5).

In order to do this, fix $N \in \mathbb{N}$, $N \gg 1$, and introduce the cutoff operators Π_N in $W^{k,p}$ by

$$\Pi_N \psi := \sum_{j \leq N} \phi_j(D)\psi,$$

where $\phi_j(D)$ are the operators we introduced in Remark 4.2.

We notice that by assumption INV the Hamiltonian vector field of h_0 generates a continuous flow Φ^t which leaves $\Pi_N W^{k,p}$ invariant.
Now we set $H = H_{N,r} + \mathcal{R}_{N,r} + \mathcal{R}_r$, where

$$H_{N,r} := h_0 + \epsilon\, h_{N,r} + \epsilon\, F_{N,r}, \tag{33}$$

$$h_{N,r} := \sum_{j=1}^{r} \epsilon^{j-1} h_{j,N}, \quad h_{j,N} := h_j \circ \Pi_N, \tag{34}$$

$$F_{N,r} := \sum_{j=1}^{r} \epsilon^{j-1} F_{j,N}, \quad F_{j,N} := F_j \circ \Pi_N, \tag{35}$$

and

$$\mathcal{R}_{N,r} := h_0 + \sum_{j=1}^{r} \epsilon^j h_j + \sum_{j=1}^{r} \epsilon^j F_j - H_{N,r}, \tag{36}$$

$$\mathcal{R}_r := \epsilon \left(h - \sum_{j=1}^{r} \epsilon^{j-1} h_j \right) + \epsilon \left(F - \sum_{j=1}^{r} \epsilon^{j-1} F_j \right). \tag{37}$$

The system described by the Hamiltonian (33) is the one that we will put in normal form.

In the following we will use the notation $a \preceq b$ to mean: there exists a positive constant K independent of N and R (but dependent on r, k and p), such that $a \leq Kb$.

We exploit the following intermediate results:

Lemma 5.1. *For any $k \geq k_1$ and $p \in (1, +\infty)$ there exists $B_{k,p}(R) \subset W^{k,p}$ s.t. $\forall\, \sigma > 0$, $N > 0$*

$$\sup_{B_{k+\sigma+2(r+1),p}(R)} \|X_{\mathcal{R}_{N,r}}(\psi, \bar{\psi})\|_{W^{k,p}} \preceq \frac{\epsilon}{2^{\sigma(N+1)}}, \tag{38}$$

$$\sup_{B_{k+2(r+1),p}(R)} \|X_{\mathcal{R}_r}(\psi, \bar{\psi})\|_{W^{k,p}} \preceq \epsilon^{r+1}. \tag{39}$$

Proof. We recall that $\mathcal{R}_{N,r} = h_0 + \sum_{j=1}^r \epsilon^j h_j + \sum_{j=1}^r \epsilon^j F_j - H_{N,r}$.
Now, $\|id - \Pi_N\|_{W^{k+\sigma,p} \to W^{k,p}} \preceq 2^{-\sigma(N+1)}$, since

$$\begin{aligned}
\left\| \sum_{j \geq N+1} \phi_j(D) f \right\|_{W^{k,p}} &\preceq \left\| \left[\sum_{j \geq N+1} |2^{jk} \phi_j(D) f|^2 \right]^{1/2} \right\|_{L^p} \\
&\preceq 2^{-\sigma(N+1)} \left\| \left[\sum_{j \geq N+1} |2^{j(k+\sigma)} \phi_j(D) f|^2 \right]^{1/2} \right\|_{L^p} \\
&\preceq 2^{-\sigma(N+1)} \|f\|_{W^{k+\sigma,p}},
\end{aligned}$$

hence

$$\begin{aligned}
&\sup_{\psi \in B_{k+2(r+1)+\sigma,p}(R)} \|X_{\mathcal{R}_{N,r}}(\psi, \bar{\psi})\|_{W^{k,p}} \\
&\preceq \|dX_{\sum_{j=1}^r \epsilon^j (h_j + F_j)}\|_{L^\infty(B_{k+2(r+1),p}(R), W^{k,p})} \|id - \Pi_N\|_{L^\infty(B_{k+2(r+1)+\sigma,p}(R), B_{k+2(r+1),p})} \\
&\preceq \epsilon\, 2^{-\sigma(N+1)}.
\end{aligned}$$

The estimate of $X_{\mathcal{R}_r}$ follow from the hypothesis HVF.

$\square$

Lemma 5.2. *Let $j \geq 1$. Then for any $k \geq k_1 + 2(j-1)$ and $p \in (1, +\infty)$ there exists $B_{k,p}(R) \subset W^{k,p}$ such that*

$$\begin{aligned}
\sup_{B_{k,p}(R)} \|X_{h_{j,N}}(\psi, \bar{\psi})\|_{k,p} &\leq K_{j,k,p}^{(h)} 2^{2jN}, \\
\sup_{B_{k,p}(R)} \|X_{F_{j,N}}(\psi, \bar{\psi})\|_{k,p} &\leq K_{j,k,p}^{(F)} 2^{2(j-1)N},
\end{aligned}$$

where

$$\begin{aligned}
K_{j,k,p}^{(h)} &:= \sup_{B_{k,p}(R)} \|X_{h_j}(\psi, \bar{\psi})\|_{k-2j,p}, \\
K_{j,k,p}^{(F)} &:= \sup_{B_{k,p}(R)} \|X_{F_j}(\psi, \bar{\psi})\|_{k-2(j-1),p}.
\end{aligned}$$

Proof. It follows from

$$\sup_{\psi \in B_{k,p}(R)} \left\| \sum_{h \leq N} \phi_h(D) X_{F_{j,N}}(\psi, \bar{\psi}) \right\|_{W^{k,p}} \preceq \sup_{\psi \in B_{k,p}(R)} \left\| \left[\sum_{h \leq N} |2^{hk} \phi_h(D) X_{F_{j,N}}(\psi, \bar{\psi})|^2 \right]^{1/2} \right\|_{L^p} \tag{40}$$

$$\leq 2^{2(j-1)N} \sup_{\psi \in B_{k,p}(R)} \left\| \left[\sum_{h \leq N} |2^{h[k-2(j-1)]} \phi_h(D) X_{F_{j,N}}(\psi, \bar{\psi})|^2 \right]^{1/2} \right\|_{L^p} \tag{41}$$

$$\preceq 2^{2(j-1)N} \sup_{\psi \in B_{k,p}(R)} \|X_{F_{j,N}}(\psi, \bar{\psi})\|_{k-2(j-1),p} \tag{42}$$

$$= K_{j,k,p}^{(F)}\, 2^{2(j-1)N}, \tag{43}$$

and similarly for $X_{h_{j,N}}$. $\square$

Next we have to normalize the system (33). In order to do this we need a slight reformulation of Theorem 4.4 in [Bam99]. Here we report a statement of the result adapted to our context.

Lemma 5.3. *Let $k \geq k_1 + 2r$, $p \in (1, +\infty)$, $R > 0$, and consider the system* (33). *Assume that $\epsilon < 2^{-4Nr}$, and that*

$$(K^{(F,r)}_{k,p} + K^{(h,r)}_{k,p}) r 2^{2Nr} \epsilon < 2^{-9} e^{-1} \pi^{-1} R, \tag{44}$$

where

$$K^{(F,r)}_{k,p} := \sup_{1 \leq j \leq r} \sup_{\psi \in B_{k,p}(R)} \| X_{F_j}(\psi, \bar{\psi}) \|_{k-2(j-1),p},$$

$$K^{(h,r)}_{k,p} := \sup_{1 \leq j \leq r} \sup_{\psi \in B_{k,p}(R)} \| X_{h_j}(\psi, \bar{\psi}) \|_{k-2j,p}.$$

Then there exists an analytic canonical transformation $\mathcal{T}^{(r)}_{\epsilon,N} : B_{k,p}(R) \to B_{k,p}(2R)$ such that

$$\sup_{B_{k,p}(R/2)} \| \mathcal{T}^{(r)}_{\epsilon,N}(\psi, \bar{\psi}) - (\psi, \bar{\psi}) \|_{W^{k,p}} \leq 4\pi r K^{(F,r)}_{k,p} 2^{2Nr} \epsilon,$$

and that puts (33) *in normal form up to a small remainder,*

$$H_{N,r} \circ \mathcal{T}^{(r)}_{\epsilon,N} = h_0 + \epsilon h_{N,r} + \epsilon Z^{(r)}_N + \epsilon^{r+1} \mathcal{R}^{(r)}_N, \tag{45}$$

with $Z^{(r)}_N$ is in normal form, namely $\{h_{0,N}, Z^{(r)}_N\} = 0$, and

$$\begin{aligned} \sup_{B_{k,p}(R/2)} \| X_{Z^{(r)}_N}(\psi, \bar{\psi}) \|_{k,p} &\leq 4\, 2^{2Nr} \epsilon \left(r K^{(F,r)}_{k,p} + r K^{(h,r)}_{k,p} \right) r 2^{2Nr} K^{(F,r)}_{k,p} \\ &= 4 r^2 K^{(F,r)}_{k,p} (K^{(F,r)}_{k,p} + K^{(h,r)}_{k,p}) 2^{4Nr} \epsilon, \end{aligned} \tag{46}$$

$$\sup_{B_{k,p}(R/2)} \| X_{\mathcal{R}^{(r)}_N}(\psi, \bar{\psi}) \|_{k,p} \tag{47}$$

$$\leq 2^8 e \frac{T}{R} (K^{(F,r)}_{k,p} + K^{(F,r)}_{k,p}) r 2^{2Nr} \tag{48}$$

$$\times \left[\frac{4T}{R} \left(2^9 3^2 e \frac{T}{R} (K^{(F,r)}_{k,p} + K^{(F,r)}_{k,p}) K^{(F,r)}_{k,p} r^2 2^{4Nr} \epsilon + 5 K^{(h,r)}_{k,p}\, r 2^{2Nr} + 5 K^{(F,r)}_{k,p}\, r 2^{2Nr} \right) r \right]^r \tag{49}$$

The proof of Lemma 5.3 is postponed to Appendix A.

Remark 5.4. *In the original notation of Theorem 4.4 in [Bam99] we set*

$$\begin{aligned} \mathcal{P} &= W^{k,p}, \\ h_\omega &= h_0, \\ \hat{h} &= \epsilon h_{N,r}, \\ f &= \epsilon F_{N,r}, \\ f_1 &= r = g \equiv 0, \\ F &= K^{(F,r)}_{k,p}\, r 2^{2Nr}\, \epsilon, \\ F_0 &= K^{(h,r)}_{k,p}\, r 2^{2Nr}\, \epsilon. \end{aligned}$$

Remark 5.5. *Actually, Lemma 5.3 would also hold under a weaker smallness assumption on ϵ: it would be enough that $\epsilon < 2^{-2N}$, and that*

$$\epsilon \left[K_{k,p}^{(F,r)} \frac{1 - 2^{2Nr}\epsilon^r}{1 - 2^{2N}\epsilon} + K_{k,p}^{(h,r)} \frac{2^{2N}(1 - 2^{2Nr}\epsilon^r)}{1 - 2^{2N}\epsilon} \right] < 2^{-9}e^{-1}\pi^{-1}R \tag{50}$$

is satified. However, condition (50) *is less explicit than* (44)*, that allows us to apply directly the scheme of [Bam99]. The disadvantage of the stronger smallness assumption* (44) *is that it holds for a smaller range of ϵ, and that at the end of the proof it will force us to choose a larger parameter $\sigma = 4r^2$. By using* (50) *and by making a more careful analysis, it may be possible to prove Theorem 4.3 also by choosing $\sigma = 2r$.*

Now we conclude with the proof of the Theorem 4.3.

Proof. Now consider the transformation $\mathcal{T}_{\epsilon,N}^{(r)}$ defined by Lemma 5.3, then

$$(\mathcal{T}_{\epsilon,N}^{(r)})^* H = h_0 \ + \sum_{j=1}^{r} \epsilon^j h_{j,N} \ + \epsilon Z_N^{(r)} + \epsilon^{r+1}\mathcal{R}_N^{(r)} + \epsilon^r \mathcal{R}_{Gal}$$

where we recall that

$$\epsilon^r \mathcal{R}_{Gal} := (\mathcal{T}_{\epsilon,N}^{(r)})^* (\mathcal{R}_{N,r} + \mathcal{R}_r).$$

By exploiting the Lemma 5.3 we can estimate the vector field of $\mathcal{R}_N^{(r)}$, while by using Lemma 5.1 and (123) we get

$$\sup_{B_{k+\sigma+2(r+1),p}(R/2)} \|X_{\mathcal{R}_{Gal}}(\psi, \bar{\psi})\|_{W^{k,p}} \preceq \left(\frac{\epsilon}{2^{\sigma(N+1)}} + \frac{\epsilon^{r+1}}{\sigma + 2(r+1)} \right). \tag{51}$$

To get the result choose

$$\begin{aligned} k_0 &= \sigma + 2(r+1), \\ N &= r\sigma^{-1}\log_2(1/\epsilon) - 1, \\ \sigma &= 4r^2. \end{aligned}$$

□

Remark 5.6. *The compatibility condition $N \geq 1$ and* (44) *lead to*

$$\epsilon \leq \left[2^{-9}e^{-1}\pi^{-1}R(K_{k,p}^{(F,r)} + K_{k,p}^{(h,r)})^{-1} r^{-1} 2^{-2r} \right]^{\frac{\sigma}{2r}} =: \epsilon_{r,k,p} \leq 2^{-2\sigma/r} \leq 2^{-8r}.$$

Remark 5.7. *We point out the fact that Theorem 4.3 holds for the scale of Banach spaces $W^{k,p}(M, \mathbb{C}^n \times \mathbb{C}^n)$, where $k \geq 1$, $1 < p < +\infty$, $n \in \mathbb{N}_0$, and where M is a smooth manifold on which the Littlewood-Paley decomposition can be constructed, for example a compact manifold (see sect. 2.1 in [BGT04]), $\mathbb{R}^d$, or a noncompact manifold satisfying some technical assumptions (see [Bou10]).*

If we restrict to the case $p = 2$, and we consider M as either $\mathbb{R}^d$ or the d-dimensional torus $\mathbb{T}^d$, we can prove an analogous result for Hamiltonians $H(\psi, \bar{\psi})$ with $(\psi, \bar{\psi}) \in H^k := W^{k,2}(M, \mathbb{C} \times \mathbb{C})$. In the following we denote by $B_k(R)$ the open ball of radius R and center 0 in H^k. We recall that the Fourier projection operator on H^k is given by

$$\pi_j \psi(x) := (2\pi)^{-d/2} \int_{j-1 \leq |k| \leq j} \hat{\psi}(k) e^{ik\cdot x} \mathrm{d}k, \quad j \geq 1.$$

Theorem 5.8. *Fix $r \geq 1$, $R > 0$, $k_1 \gg 1$. Consider (27), and assume PER, INV (with respect to Fourier projection operators), NF and HVF. Then $\exists\ k_0 = k_0(r) > 0$ with the following properties: for any $k \geq k_1$ there exists $\epsilon_{r,k} \ll 1$ such that for any $\epsilon < \epsilon_{r,k}$ there exists $\mathcal{T}_\epsilon^{(r)} : B_k(R) \to B_k(2R)$ transformation s.t.*

$$H_r := H \circ \mathcal{T}_\epsilon^{(r)} = h_0 + \sum_{j=1}^{r} \epsilon^j \mathcal{Z}_j + \epsilon^{r+1} \mathcal{R}^{(r)},$$

where $\mathcal{Z}_j$ are in normal form, namely

$$\{\mathcal{Z}_j, h_0\} = 0, \tag{52}$$

and

$$\sup_{B_{k+k_0}(R)} \|X_{\mathcal{R}^{(r)}}\|_{H^k} \leq C_k, \tag{53}$$

$$\sup_{B_k(R)} \|\mathcal{T}_\epsilon^{(r)} - id\|_{H^k} \leq C_k\, \epsilon. \tag{54}$$

In particular, we have that

$$\mathcal{Z}_1(\psi, \bar{\psi}) = h_1(\psi, \bar{\psi}) + \langle F_1 \rangle (\psi, \bar{\psi}),$$

where $\langle F_1 \rangle (\psi, \bar{\psi}) := \int_0^{2\pi} F_1 \circ \Phi^t(\psi, \bar{\psi}) \frac{\mathrm{d}t}{2\pi}$.

The only technical difference between the proofs of Theorem 4.3 and the proof of Theorem 5.8 is that we exploit the Fourier cut-off operator

$$\Pi_N \psi(x) := \int_{|k| \leq N} \hat{\psi}(k) e^{ik\cdot x} \mathrm{d}k,$$

as in [Bam05]. This in turn affects (38), which in this case reads

$$\sup_{B_{k+\sigma+2(r+1)}(R)} \|X_{\mathcal{R}_{N,r}}(\psi, \bar{\psi})\|_{H^k} \preceq \frac{\epsilon}{N^\sigma}, \tag{55}$$

and (51), for which we have to choose a bigger cut-off, $N = \epsilon^{-r\sigma}$.

6 Application to the nonlinear Klein-Gordon equation

6.1 The real nonlinear Klein-Gordon equation

We first consider the Hamiltonian of the real non-linear Klein-Gordon equation with power-type nonlinearity on a smooth manifold M (M is such the Littlewood-Paley decomposition is well-defined; take, for example, a smooth compact manifold, or $\mathbb{R}^d$). The Hamiltonian is of the form

$$H(u,v) = \frac{c^2}{2}\langle v,v\rangle + \frac{1}{2}\langle u, \langle\nabla\rangle_c^2 u\rangle \; + \; \lambda \int \frac{u^{2l}}{2l}, \tag{56}$$

where $\langle\nabla\rangle_c := (c^2-\Delta)^{1/2}$, $\lambda \in \mathbb{R}$, $l \geq 2$.
If we introduce the complex-valued variable

$$\psi := \frac{1}{\sqrt{2}}\left[\left(\frac{\langle\nabla\rangle_c}{c}\right)^{1/2} u - i\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2} v\right], \tag{57}$$

(the corresponding symplectic 2-form becomes $i\mathrm{d}\psi \wedge \mathrm{d}\bar{\psi}$), the Hamiltonian (56) in the coordinates $(\psi,\bar{\psi})$ is

$$H(\bar{\psi},\psi) = \langle\bar{\psi}, c\langle\nabla\rangle_c\psi\rangle + \frac{\lambda}{2l}\int\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\frac{\psi+\bar{\psi}}{\sqrt{2}}\right]^{2l}\mathrm{d}x. \tag{58}$$

If we rescale the time by a factor c^2, the Hamiltonian takes the form (27), with $\epsilon = \frac{1}{c^2}$, and

$$H(\psi,\bar{\psi}) = h_0(\psi,\bar{\psi}) + \epsilon\, h(\psi,\bar{\psi}) + \epsilon\, F(\psi,\bar{\psi}), \tag{59}$$

where

$$h_0(\psi,\bar{\psi}) = \langle\bar{\psi},\psi\rangle, \tag{60}$$

$$h(\psi,\bar{\psi}) = \langle\bar{\psi}, \left(c\langle\nabla\rangle_c - c^2\right)\psi\rangle \sim \sum_{j\geq 1}\epsilon^{j-1}\,\langle\bar{\psi}, a_j\Delta^j\psi\rangle =: \sum_{j\geq 1}\epsilon^{j-1} h_j(\psi,\bar{\psi}), \tag{61}$$

$$\begin{aligned} F(\psi,\bar{\psi}) &= \frac{\lambda}{2^{l+1}l}\int\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}(\psi+\bar{\psi})\right]^{2l}\mathrm{d}x \qquad (62)\\ &\sim \frac{\lambda}{2^{l+1}l}\int(\psi+\bar{\psi})^{2l}\mathrm{d}x\\ &- \epsilon b_2\int\left[(\psi+\bar{\psi})^{2l-1}\Delta(\psi+\bar{\psi}) + \ldots + (\psi+\bar{\psi})\Delta((\psi+\bar{\psi})^{2l-1})\right]\mathrm{d}x\\ &+ \mathcal{O}(\epsilon^2)\\ &=: \sum_{j\geq 1}\epsilon^{j-1}\,F_j(\psi,\bar{\psi}), \qquad (63)\end{aligned}$$

where $(a_j)_{j\geq 1}$ and $(b_j)_{j\geq 1}$ are real coefficients, and $F_j(\psi,\bar{\psi})$ is a polynomial function of the variables ψ and $\bar{\psi}$ (along with their derivatives) and which admits a bounded vector field from a neighborhood of the origin in $W^{k+2(j-1),p}$ to $W^{k,p}$ for any $1 < p < +\infty$.

This description clearly fits the scheme treated in the previous section, and one can easily check that assumptions PER, NF and HVF are satisfied. Therefore we can apply Theorem 4.3 to the Hamiltonian (59).

Remark 6.1. *About the normal forms obtained by applying Theorem 4.3, we remark that in the first step (case $r = 1$ in the statement of the Theorem) the homological equation we get is of the form*

$$\{\chi_1, h_0\} + F_1 = \langle F_1 \rangle, \tag{64}$$

where $F_1(\psi, \bar{\psi}) = \frac{\lambda}{2^{l+1} l} \int (\psi + \bar{\psi})^{2l} \mathrm{d}x$. Hence the transformed Hamiltonian is of the form

$$H_1(\psi, \bar{\psi}) = h_0(\psi, \bar{\psi}) + \frac{1}{c^2} \left[-\frac{1}{2} \left\langle \bar{\psi}, \Delta\psi \right\rangle + \langle F_1 \rangle (\psi, \bar{\psi}) \right] + \frac{1}{c^4} \mathcal{R}^{(1)}(\psi, \bar{\psi}). \tag{65}$$

If we neglect the remainder and we derive the corresponding equation of motion for the system, we get

$$-i\psi_t = \psi + \frac{1}{c^2} \left[-\frac{1}{2} \Delta\psi + \frac{\lambda}{2^{l+1}} \binom{2l}{l} |\psi|^{2(l-1)} \psi \right], \tag{66}$$

which is the NLS, and the Hamiltonian which generates the canonical transformation is given by

$$\chi_1(\psi, \bar{\psi}) = \frac{\lambda}{2^{l+1} l} \sum_{\substack{j=0,\dots,2l \\ j \neq l}} \frac{1}{i\, 2(l-j)} \binom{2l}{j} \int \psi^{2l-j} \bar{\psi}^j \mathrm{d}x. \tag{67}$$

Remark 6.2. *Now we iterate the construction by passing to the case $r = 2$, and for simplicity we consider only the case $l = 2$, which at the first step yields the cubic NLS. In this case one has that*

$$\begin{aligned} \chi_1(\psi, \bar{\psi}) &= \int_0^T \tau \left[F_1(\Phi^\tau(\psi, \bar{\psi})) \ - \ \langle F_1 \rangle (\Phi^\tau(\psi, \bar{\psi})) \right] \frac{\mathrm{d}\tau}{T} \\ &= \frac{\lambda}{16} \int_0^{2\pi} \tau \int \left[|e^{i\tau}\psi + e^{-i\tau}\bar{\psi}|^4 - 6|\psi|^4 \right] \mathrm{d}x \frac{\mathrm{d}\tau}{2\pi}. \end{aligned}$$

Since

$$|e^{i\tau}\psi + e^{-i\tau}\bar{\psi}|^4 = e^{4i\tau}\psi^4 + 4e^{2i\tau}\psi^3\bar{\psi} + 6\psi^2\bar{\psi}^2 + 4e^{-2i\tau}\psi\bar{\psi}^3 + e^{-4i\tau}\bar{\psi}^4$$

and since $\int_0^{2\pi} \tau e^{in\tau} \mathrm{d}\tau = \frac{2\pi}{i\,n}$ for any non-zero integer n, we finally get

$$\chi_1(\psi, \bar{\psi}) = \frac{\lambda}{16} \int \frac{\psi^4 - \bar{\psi}^4}{4i} + \frac{2}{i}(\psi^3\bar{\psi} - \psi\bar{\psi}^3) \, \mathrm{d}x.$$

If we neglect the remainder of order c^{-6}, we have that

$$\begin{aligned} H\circ\mathcal{T}^{(1)} &= h_0+\frac{1}{c^2}h_1+\frac{1}{c^4}\{\chi_1,h_1\}+\frac{1}{c^4}h_2+ \\ &+\frac{1}{c^2}\langle F_1\rangle+\frac{1}{c^4}\{\chi_1,F_1\}+\frac{1}{2c^4}\{\chi_1,\{\chi_1,h_0\}\}+\frac{1}{c^4}F_2 \qquad (68)\\ &= h_0+\frac{1}{c^2}\left[h_1+\langle F_1\rangle\right]+\frac{1}{c^4}\left[\{\chi_1,h_1\}+h_2+\{\chi_1,F_1\}+\frac{1}{2}\{\chi_1,\langle F_1\rangle-F_1\}+F_2\right], \qquad (69)\end{aligned}$$

where $h_1(\psi,\bar\psi)=-\frac{1}{2}\left\langle\bar\psi,\Delta\psi\right\rangle$.

Now we compute the terms of order $\frac{1}{c^4}$.

$$\begin{aligned}\{\chi_1,h_1\} &= \mathrm{d}\chi_1 X_{h_1}=\frac{\partial\chi_1}{\partial\psi}\cdot i\frac{\partial h_1}{\partial\bar\psi}-i\frac{\partial\chi_1}{\bar\psi}\frac{\partial h_1}{\partial\psi} \qquad (70)\\ &= -\frac{\lambda}{32}\int\left[\Delta\psi\left(\psi^3+6\psi^2\bar\psi-2\bar\psi^3\right)-\Delta\bar\psi(2\psi^3-6\psi\bar\psi^2-\bar\psi^3)\right], \qquad (71)\end{aligned}$$

$$h_2=-\frac{1}{8}\left\langle\bar\psi,\Delta^2\psi\right\rangle, \qquad (72)$$

$$\begin{aligned}\{\chi_1,F_1\} &= \frac{\lambda^2}{32}\int(4\psi^3+12\psi^2\bar\psi+12\psi\bar\psi^2+4\bar\psi^3)(\psi^3+6\psi^2\bar\psi-2\bar\psi^3)+ \qquad (73)\\ &-(4\psi^3+12\psi^2\bar\psi+12\psi\bar\psi^2+4\bar\psi^3)(2\psi^3-6\psi\bar\psi^2-\bar\psi^3)\,\mathrm{d}x, \qquad (74)\end{aligned}$$

$$\{\chi_1,\langle F_1\rangle\}=\frac{\lambda^2}{2}\int\left[|\psi|^2\psi\,\left(\psi^3+6\psi^2\bar\psi-2\bar\psi^3\right)-|\psi|^2\bar\psi\,\left(2\psi^3-6\psi\bar\psi^2-\bar\psi^3\right)\right]\,\mathrm{d}x, \qquad (75)$$

$$F_2=\frac{\lambda}{16}\int\left[(\psi^3+3\psi^2\bar\psi+3\psi\bar\psi^2+\psi^3)\,\Delta\psi+(\bar\psi^3+3\bar\psi^2\psi+3\bar\psi\psi^2+\psi^3)\,\Delta\bar\psi\right]\,\mathrm{d}x. \qquad (76)$$

Now, one can easily verify that $\langle\{\chi_1,h_1\}\rangle=\langle\{\chi_1,\langle F_1\rangle\}\rangle=0$, *and that*

$$\begin{aligned}\langle\{\chi_1,F_1\}\rangle &= \frac{\lambda^2}{32}\int(-8|\psi|^6+72|\psi|^6+4|\psi|^6)+(4|\psi|^6+72|\psi|^6-8|\psi|^6)\,\mathrm{d}x \qquad (77)\\ &= \frac{17}{4}\lambda^2\int|\psi|^6\,\mathrm{d}x, \qquad (78)\end{aligned}$$

$$\begin{aligned}\langle F_2\rangle &= \frac{\lambda}{16}\int 3\psi\bar\psi^2\,\Delta\psi+3\bar\psi\psi^2\,\Delta\bar\psi\,\mathrm{d}x \qquad (79)\\ &= \frac{\lambda}{16}\int 3|\psi|^2(\psi\,\Delta\psi+\psi\Delta\bar\psi)\,\mathrm{d}x. \qquad (80)\end{aligned}$$

Hence, up to a remainder of order $O\left(\frac{1}{c^6}\right)$, we have that

$$
\begin{aligned}
H_2 = h_0 &+ \frac{1}{c^2}\int\left[-\frac{1}{2}\left\langle\bar{\psi},\Delta\psi\right\rangle + \frac{3}{8}\lambda|\psi|^4\right]\,\mathrm{d}x \\
&+ \frac{1}{c^4}\int\left[\frac{17}{8}\lambda^2|\psi|^6 + \frac{3}{16}\lambda|\psi|^2(\bar{\psi}\,\Delta\psi + \psi\,\Delta\bar{\psi}) - \frac{1}{8}\left\langle\bar{\psi},\Delta^2\psi\right\rangle\right]\,\mathrm{d}x,
\end{aligned} \tag{81}
$$

which, by neglecting h_0 (that yields only a gauge factor) and by rescaling the time, leads to the following equations of motion

$$
\begin{aligned}
-i\psi_t \;=\; & -\frac{1}{2}\Delta\psi + \frac{3}{4}\lambda|\psi|^2\psi \\
& + \frac{1}{c^2}\left[\frac{51}{8}\lambda^2|\psi|^4\psi + \frac{3}{16}\lambda\left(2|\psi|^2\,\Delta\psi + \psi^2\Delta\bar{\psi} + \Delta(|\psi|^2\bar{\psi})\right) - \frac{1}{8}\Delta^2\psi\right].
\end{aligned} \tag{82}
$$

To the author's knowledge, Eq. (82) *has never been studied before. It is the nonlinear analogue of a linear higher-order Schrödinger equation that appears in [CM12] and [CLM15] in the context of semi-relativistic equations. Indeed, the linearization of Eq.* (82) *is studied within the framework of relativistic quantum field theory, as an approximation of nonlocal kinetic terms; Carles, Lucha and Moulay studied the well-posedness of these approximations, as well as the convergence of the equations as the order of truncation goes to infinity, in the linear case, also when one takes into account the effects of some time-independent potentials (e.g. bounded potentials, the harmonic-oscillator potential and the Coulomb potential).*

Apparently, little is known for the nonlinear equation (82)*: we just mention [CG07], in which the well-posedness of a higher-order Schrodinger equation has been studied, and [PX13], in which the scattering theory for a fourth-order Schrödinger equation in dimensions $1 \leq d \leq 4$ is studied.*

6.2 The complex nonlinear Klein-Gordon equation

Now we consider the Hamiltonian of the complex non-linear Klein-Gordon equation with power-type nonlinearity on a smooth manifold M (take, for example, a smooth compact manifold, or $\mathbb{R}^d$)

$$
H(w,p_w) = \frac{c^2}{2}\left\langle p_w,p_w\right\rangle + \frac{1}{2}\left\langle w,\langle\nabla\rangle_c^2 w\right\rangle \;+\; \lambda\int\frac{|w|^{2l}}{2l}, \tag{83}
$$

where $w:\mathbb{R}\times M\to\mathbb{C}$, $\langle\nabla\rangle_c := (c^2-\Delta)^{1/2}$, $\lambda\in\mathbb{R}$, $l\geq 2$.
If we rewrite the Hamiltonian in terms of $u := Re(w)$ and $v := Im(w)$, we have

$$
H(u,v,p_u,p_v) = \frac{c^2}{2}(\langle p_u,p_u\rangle + \langle p_v,p_v\rangle) + \frac{1}{2}(|\nabla u|^2 + |\nabla v|^2) + \frac{c^2}{2}(u^2+v^2) + \lambda\int\frac{(u^2+v^2)^l}{2l}. \tag{84}
$$

We will consider by simplicity only the cubic case ($l=2$), but the argument may be readily generalized to the other power-type nonlinearities.

If we introduce the variables

$$\psi := \frac{1}{\sqrt{2}}\left[\left(\frac{\langle\nabla\rangle_c}{c}\right)^{1/2} u - i\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2} p_u\right], \tag{85}$$

$$\phi := \frac{1}{\sqrt{2}}\left[\left(\frac{\langle\nabla\rangle_c}{c}\right)^{1/2} v + i\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2} p_v\right], \tag{86}$$

(the corresponding symplectic 2-form becomes $i\mathrm{d}\psi\wedge\mathrm{d}\bar{\psi} - i\mathrm{d}\phi\wedge\mathrm{d}\bar{\phi}$), the Hamiltonian (83) in the coordinates $(\psi,\phi,\bar{\psi},\bar{\phi})$ reads

$$\begin{aligned} H(\psi,\phi,\bar{\psi},\bar{\phi}) &= \left\langle\bar{\psi}, c\langle\nabla\rangle_c\psi\right\rangle + \left\langle\bar{\phi}, c\langle\nabla\rangle_c\phi\right\rangle && (87)\\ &+ \frac{\lambda}{16}\int_M \left[\left\langle\psi+\bar{\psi}, \frac{c}{\langle\nabla\rangle_c}(\psi+\bar{\psi})\right\rangle + \left\langle\phi+\bar{\phi}, \frac{c}{\langle\nabla\rangle_c}(\phi+\bar{\phi})\right\rangle\right]^2 \mathrm{d}x, && (88)\end{aligned}$$

with corresponding equations of motion

$$\begin{cases} -i\psi_t &= c\langle\nabla\rangle_c\psi + \frac{1}{4}\left[\left\langle\psi+\bar{\psi}, \frac{c}{\langle\nabla\rangle_c}(\psi+\bar{\psi})\right\rangle + \left\langle\phi+\bar{\phi}, \frac{c}{\langle\nabla\rangle_c}(\phi+\bar{\phi})\right\rangle\right]\frac{c}{\langle\nabla\rangle_c}(\psi+\bar{\psi}),\\ i\phi_t &= c\langle\nabla\rangle_c\phi + \frac{1}{4}\left[\left\langle\psi+\bar{\psi}, \frac{c}{\langle\nabla\rangle_c}(\psi+\bar{\psi})\right\rangle + \left\langle\phi+\bar{\phi}, \frac{c}{\langle\nabla\rangle_c}(\phi+\bar{\phi})\right\rangle\right]\frac{c}{\langle\nabla\rangle_c}(\phi+\bar{\phi}). \end{cases}$$

If we rescale the time by a factor c^2, the Hamiltonian takes the form (27), with $\epsilon = \frac{1}{c^2}$, and

$$H(\psi,\phi,\bar{\psi},\bar{\phi}) = H_0(\psi,\phi,\bar{\psi},\bar{\phi}) + \epsilon\, h(\psi,\phi,\bar{\psi},\bar{\phi}) + \epsilon\, F(\psi,\phi,\bar{\psi},\bar{\phi}), \tag{89}$$

where

$$\begin{aligned} H_0(\psi,\phi,\bar{\psi},\bar{\phi}) &= \left\langle\bar{\psi},\psi\right\rangle + \left\langle\bar{\phi},\phi\right\rangle, && (90)\\ h(\psi,\phi,\bar{\psi},\bar{\phi}) &= \left\langle\bar{\psi}, \left(c\langle\nabla\rangle_c - c^2\right)\psi\right\rangle - \left\langle\bar{\phi}, \left(c\langle\nabla\rangle_c - c^2\right)\phi\right\rangle \\ &\sim \sum_{j\geq 1}\epsilon^{j-1}\left(\left\langle\bar{\psi}, a_j\Delta^j\psi\right\rangle + \left\langle\bar{\phi}, a_j\Delta^j\phi\right\rangle\right) \\ &=: \sum_{j\geq 1}\epsilon^{j-1}(h_j(\psi,\phi,\bar{\psi},\bar{\phi})), && (91)\\ F(\psi,\phi,\bar{\psi},\bar{\phi}) &= \frac{\lambda}{16}\int_{\mathbb{T}} \left[\left\langle\psi+\bar{\psi}, \frac{c}{\langle\nabla\rangle_c}(\psi+\bar{\psi})\right\rangle + \left\langle\phi+\bar{\phi}, \frac{c}{\langle\nabla\rangle_c}(\phi+\bar{\phi})\right\rangle\right]^2 \mathrm{d}x, \\ &\sim \frac{\lambda}{16}\int \left[|\psi+\bar{\psi}|^2 + |\phi+\bar{\phi}|^2\right]^2 \mathrm{d}x \\ &+ \mathcal{O}(\epsilon) \\ &=: \sum_{j\geq 1}\epsilon^{j-1}\, F_j(\psi,\phi,\bar{\psi},\bar{\phi}), && (92)\end{aligned}$$

where $(a_j)_{j\geq 1}$ are real coefficients, and $F_j(\psi,\phi,\bar{\psi},\bar{\phi})$ is a polynomial function of the variables ψ, ϕ, $\bar{\psi}$, $\bar{\phi}$ (along with their derivatives) and which admits a bounded vector field from a neighborhood of the origin in $W^{k+2(j-1),p}(\mathbb{R}^d, \mathbb{C}^2\times\mathbb{C}^2)$ to $W^{k,p}(\mathbb{R}^d, \mathbb{C}^2\times\mathbb{C}^2)$ for any $1 < p < +\infty$.

This description clearly fits the scheme treated in sect. 4 with $n = 2$, and one can easily check that assumptions PER, NF and HVF are satisfied. Therefore we can apply Theorem 4.3 to the Hamiltonian (89).

Remark 6.3. *About the normal forms obtained by applying Theorem 4.3, we remark that in the first step (case $r = 1$ in the statement of the Theorem) the homological equation we get is of the form*

$$\{\chi_1, h_0\} + F_1 = \langle F_1 \rangle , \tag{93}$$

where $F_1(\psi,\bar\psi) = \frac{\lambda}{16}\int \left[|\psi+\bar\psi|^2 + |\phi+\bar\phi|^2\right]^2 \mathrm{d}x$. *Hence the transformed Hamiltonian is of the form*

$$\begin{aligned} H_1(\psi,\phi,\bar\psi,\bar\phi) &= h_0(\psi,\phi,\bar\psi,\bar\phi) + \frac{1}{c^2}\left[-\frac{1}{2}\left(\langle\bar\psi,\Delta\psi\rangle + \langle\bar\phi,\Delta\phi\rangle\right) + \langle F_1\rangle(\psi,\phi,\bar\psi,\bar\phi)\right] \\ &\quad + \frac{1}{c^4}\mathcal{R}^{(1)}(\psi,\phi,\bar\psi,\bar\phi), \end{aligned} \tag{94}$$

where

$$\begin{aligned} \langle F_1\rangle &= \frac{\lambda}{16}\left[6\psi^2\bar\psi^2 + 6\phi^2\bar\phi^2 + 8\psi\bar\psi\phi\bar\phi + 2\psi^2\phi^2 + 2\bar\psi^2\bar\phi^2\right] \\ &= \frac{\lambda}{8}\left[3(|\psi|^2+|\phi|^2)^2 + 2(\psi\phi - \bar\psi\bar\phi)^2\right]. \end{aligned}$$

If we neglect the remainder and we derive the corresponding equations of motion for the system, we get

$$\begin{cases} -i\psi_t &= \psi + \frac{1}{c^2}\left\{-\frac{1}{2}\Delta\psi + \frac{\lambda}{4}\left[3(|\psi|^2+|\phi|^2)\psi + 2(\psi\phi+\bar\psi\bar\phi)\bar\phi\right]\right\}, \\ i\phi_t &= \phi + \frac{1}{c^2}\left\{-\frac{1}{2}\Delta\phi + \frac{\lambda}{4}\left[3(|\psi|^2+|\phi|^2)\phi + 2(\psi\phi+\bar\psi\bar\phi)\bar\psi\right]\right\}, \end{cases} \tag{95}$$

which is a system of two coupled NLS equations.

7 Dynamics

Now we want to exploit the result of the previous section in order to deduce some consequences about the dynamics of the NLKG equation (3) in the non-relativistic limit. Consider the *simplified system*, that is the Hamiltonian H_r in the notations of Theorem 4.3, where we neglect the remainder:

$$H_{simp} := h_0 + \epsilon(h_1 + \langle F_1\rangle) + \sum_{j=2}^{r}\epsilon^j(h_j + Z_j).$$

We recall that in the case of the NLKG the simplified system is actually the NLS (given by $h_0+\epsilon(h_1+\langle F_1\rangle)$), plus higher-order normalized corrections. Now let ψ_r be a solution of

$$-i\,\dot\psi_r \;=\; X_{H_{simp}}(\psi_r), \tag{96}$$

then $\psi_a(t,x) := \mathcal{T}^{(r)}(\psi_r(c^2t,x))$ solves

$$\dot{\psi}_a = ic\langle\nabla\rangle_c\psi_a + \frac{\lambda}{2l}\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\frac{\psi_a+\bar{\psi}_a}{\sqrt{2}}\right]^{2l-1} - \frac{1}{c^{2r}}X_{\mathcal{T}^{(r)*}\mathcal{R}^{(r)}}(\psi_a,\bar{\psi}_a), \tag{97}$$

that is, the NLKG plus a remainder of order c^{-2r} (in the following we will refer to equation (97) as *approximate equation*, and to ψ_a as the *approximate solution* of the original NLKG). We point out that the original NLKG and the approximate equation differ only by a remainder of order c^{-2r}, which is evaluated on the approximate solution. This fact is extremely important: indeed, if one can prove the smoothness of the approximate solution (which often is easier to check than the smoothness of the solution of the original equation), then the contribution of the remainder may be considered small in the nonrelativistic limit. This property is rather general, and has been already applied in the framework of normal form theory (see for example [BCP02]).

Now let ψ be a solution of the NLKG equation (3) with initial datum ψ_0, and let $\delta := \psi - \psi_a$ be the error between the solution of the approximate equation and the original one. One can check that δ fulfills

$$\dot{\delta} = ic\langle\nabla\rangle_c\delta + [P(\psi_a+\delta,\bar{\psi}_a+\bar{\delta}) - P(\psi_a,\bar{\psi}_a)] + \frac{1}{c^{2r}}X_{\mathcal{T}^{(r)*}\mathcal{R}^{(r)}}(\psi_a(t),\bar{\psi}_a(t)),$$

where

$$P(\psi,\bar{\psi}) = \frac{\lambda}{2l}\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\frac{\psi+\bar{\psi}}{\sqrt{2}}\right]^{2l-1}. \tag{98}$$

Thus we get

$$\dot{\delta} = i\,c\langle\nabla\rangle_c\delta + dP(\psi_a(t))\delta + \mathcal{O}(\delta^2) + \mathcal{O}\left(\frac{1}{c^{2r}}\right);$$

$$\delta(t) = e^{itc\langle\nabla\rangle_c}\delta_0 + \int_0^t e^{i(t-s)c\langle\nabla\rangle_c}dP(\psi_a(s))\delta(s)\mathrm{d}s + \mathcal{O}(\delta^2) + \mathcal{O}\left(\frac{1}{c^{2r}}\right). \tag{99}$$

By applying Gronwall inequality to (99) we obtain

Proposition 7.1. *Fix $r \geq 1$, $R > 0$, $k_1 \gg 1$, $1 < p < +\infty$. Then $\exists\ k_0 = k_0(r) > 0$ with the following properties: for any $k \geq k_1$ there exists $c_{l,r,k,p,R} \gg 1$ such that for any $c > c_{l,r,k,p,R}$, if we assume that*

$$\|\psi_0\|_{k+k_0,p} \leq R$$

and that there exists $T = T_{r,k,p} > 0$ such that the solution of (96) satisfies

$$\|\psi_r(t)\|_{k+k_0,p} \leq 2R, \ \ for\ \ 0 \leq t \leq T,$$

then

$$\|\delta(t)\|_{k,p} \leq C_{k,p}\,c^{-2r}, \ \ for\ \ 0 \leq t \leq T. \tag{100}$$

Remark 7.2. *If we restrict to $p = 2$, and to $M = \mathbb{T}^d$, the above result is actually a reformulation of Theorem 3.2 in [FS14]. We also remark that the time interval $[0,T]$ in which estimate (100) is valid is* independent *of c.*

Remark 7.3. *By exploiting estimate* (32) *about the canonical transformation, Proposition 7.1 leads immediately to a proof of Theorem 2.3.*

In order to study the evolution of the error between the approximate solution and the solution of the NLKG over longer (namely, c-dependent) time scales, we observe that the error is described by

$$\dot{\delta}(t) = i\, c\langle\nabla\rangle_c \delta(t) + dP(\psi_a(t))\delta(t); \tag{101}$$

$$\delta(t) = e^{itc\langle\nabla\rangle_c}\delta_0 + \int_0^t e^{i(t-s)c\langle\nabla\rangle_c} dP(\psi_a(s))\delta(s)\mathrm{d}s, \tag{102}$$

up to a remainder which is small, if we assume the smoothness of ψ_a.

Equation (101) in the context of dispersive PDEs is known as *semirelativistic spinless Salpeter equation* with a time-dependent potential. This system was introduced as a first order in time analogue of the KG equation for the Lorentz-covariant description of bound states within the framework of relativistic quantum field theory, and, despite the nonlocality of its Hamiltonian, some of its properties have already been studied (see [Suc63] for a study from a physical point of view; for a more mathematical approach see [Läm93] and the more recent works [CM12] and [CLM15], which are closer to the spirit of our approximation).

It seems reasonable to estimate the solution of Equation (101) by studying and by exploiting its dispersive properties, and this will be the aim of the following sections. From now on we will consider by simplicity only the three-dimensional case, $d = 3$, but the argument may also be applied to $M = \mathbb{R}^d$ for $d \geq 2$.

8 Long time approximation

Now we study the evolution of the the error between the approximate solution ψ_a, namely the solution of (97), and the original solution ψ of (3) for long (that means, c-dependent) time intervals. As pointed out in Sect. 2, we will prove a result only for the linear case; we will also begin to discuss the long time approximation of the NLKG, but we defer more precise results to a future work.

8.1 Linear case

Fix $r \geq 1$, and take $\psi_0 \in H^{k+k_0}$, where $k_0 > 0$ and $k \gg 1$ are the ones in Theorem 4.3. In [CM12] and [CLM15] the authors proved that the linearized normal form system, namely the one that corresponds (up to a rescaling of time by a factor c^2) to

$$\begin{aligned} -i\dot{\psi}_r &= X_{h_0+\sum_{j=1}^{r}\epsilon^j h_j}(\psi_r), \\ \psi_r(0) &= \psi_0, \end{aligned} \tag{103}$$

admits a unique solution in $L^\infty(\mathbb{R})H^{k+k_0}(\mathbb{R}^3)$ (this is a simple application of the properties of the Fourier transform), and by a perturbative argument they also proved the global existence also for the higher oder Schrödinger equation with a bounded time-independent potential.

Moreover, by following the arguments of Theorem 4.1 in [KAY12] and Lemma 4.3 in [CLM15] one obtains the following dispersive estimates and local-in-time Strichartz estimates for solutions of the linearized normal form equation (103).

Proposition 8.1. *Let $r \geq 1$, and denote by $\mathcal{U}_r(t)$ the evolution operator of* (103)*. Then one has the following local-in-time dispersive estimate*

$$\|\mathcal{U}_r(t)\|_{L^1(\mathbb{R}^3)\to L^\infty(\mathbb{R}^3)} \preceq c^{3\left(1-\frac{1}{r}\right)}|t|^{-3/(2r)}, \quad 0<|t|\leq c^{2(r-1)}. \tag{104}$$

On the other hand, $\mathcal{U}_r(t)$ is unitary on $L^2(\mathbb{R}^3)$.
Now introduce the following set of admissible exponent pairs:

$$\Delta_r := \{(p,q) : (1/p, 1/q) \text{ lies in the closed quadrilateral } ABCD\}, \tag{105}$$

where

$$A=\left(\frac{1}{2},\frac{1}{2}\right), \quad B=\left(1,\frac{1}{\tau_r}\right), \quad C=(1,0), \quad D=\left(\frac{1}{\tau_r'},0\right), \quad \tau_r=\frac{2r-1}{r-1}, \quad \frac{1}{\tau_r}+\frac{1}{\tau_r'}=1.$$

Then for any $(p,q)\in\Delta_r \setminus \{(2,2),(1,\tau_r),(\tau_r',\infty)\}$

$$\|\mathcal{U}_r(t)\|_{L^p(\mathbb{R}^3)\to L^q(\mathbb{R}^3)} \preceq c^{3\left(1-\frac{1}{r}\right)\left(\frac{1}{p}-\frac{1}{q}\right)}|t|^{-\frac{3}{2r}\left(\frac{1}{q}-\frac{1}{p}\right)}, \quad 0<|t|\leq c^{2(r-1)}. \tag{106}$$

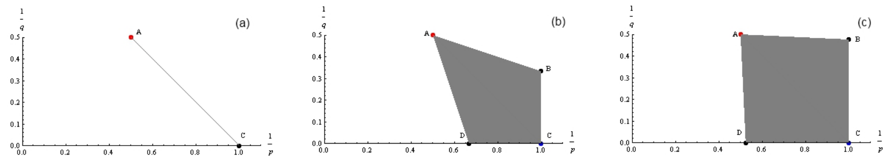

Figure 1: Set of admissible exponents Δ_r for different values of r: (a) r=1 (this is the Schrödinger case); (b) r=2; (c) r=11.

Let $r\geq 1$: in the following lemma (p,q) is called an order-r admissible pair when $2\leq q\leq +\infty$ for $r\geq 2$ ($2\leq q\leq 6$ for $r=1$), and

$$\frac{2}{p}+\frac{3}{rq}=\frac{3}{2r}. \tag{107}$$

Proposition 8.2. *Let $r \geq 1$, and denote by $\mathcal{U}_r(t)$ the evolution operator of* (103)*. Let (p,q) and (r,s) be order-r admissible pairs, then for any $T\preceq c^{2(r-1)}$*

$$\|\mathcal{U}_r(t)\phi_0\|_{L^p([0,T])L^q(\mathbb{R}^3)} \preceq c^{3\left(1-\frac{1}{r}\right)\left(\frac{1}{2}-\frac{1}{q}\right)}\|\phi_0\|_{L^2(\mathbb{R}^3)} = c^{\left(1-\frac{1}{r}\right)\frac{2r}{p}}\|\phi_0\|_{L^2(\mathbb{R}^3)}. \tag{108}$$

Now, we want to estimate the space-time norm of the error $\delta=\psi-\psi_a$. In the linear case we can observe that δ satisfies

$$\dot{\delta}=ic\langle\nabla\rangle_c\delta+\frac{1}{c^{2r}}X_{\mathcal{T}^{(r)*}\mathcal{R}^{(r)}}(\psi_a(t),\bar{\psi}_a(t)). \tag{109}$$

Remark 8.3. *By applying the Strichartz estimate* (14) *(choose $p=+\infty$, $q=2$, $r=+\infty$, $s=2$), together with estimate* (31) *for the vector field of the remainder $\mathcal{R}^{(r)}$, estimate* (32) *for the canonical transformation $\mathcal{T}^{(r)}$, and estimate* (106) *(choose $p=q=2$), we can deduce Theorem 2.4.*

8.2 The nonlinear case: radiation solutions

Now, assume that we want to recover the approach of Sect. 8.1 to approximate radiation solutions of the NLKG equation for long (c-dependent) timescales.

To pursue such a program, even by a perturbative argument, we would need to consider a (small) radiation solution $\psi_r = \eta_{rad}$ of the normalized system (96) that exists for all times, and such that it satisfies the dispersive estimates (106), in order to ensure the approximation up to times of order $\mathcal{O}(c^{2(r-1)})$. However, for $r > 1$ the issues of global existence and dispersive estimates for (96) are still open problems, as we point out in the following remarks.

Remark 8.4. *The assumption of global existence for ψ_r is actually a delicate matter. Equation* (96) *is a nonlinear perturbation of a higher-order Schrödinger equation.*

We recall that in [CM12] and [CLM15] the authors proved that the linearized system admits a unique solution in $L^\infty(\mathbb{R})H^k(\mathbb{R}^3)$, and by a perturbative argument they also proved the global existence also for the higher oder Schrödinger equation with a bounded time-independent potential.

In the nonlinear case little is known: see for example [CG07] for the well-posedness for a higher-order nonlinear Schrödinger equation.

Even if we restrict to the case $r = 2$, the issues of global existence and scattering for Eq. (82) *have not been solved. Even though some results for the linearization of Eq.* (82) *have already been established (see [BAKS00] and [KAY12] for dispersive estimates, and [CLM15] for Strichartz estimates), the study of the fourth-order NLS-type (4NLS) equation is still open: while there are some papers dealing with the local well-posedness of 4NLS (see for example [HJ07] for the one-dimensional case, [HJ11] for the multidimensional case), global well-posedness and scattering results are much less known. The recent [RWZ16] gives the first global well-posedness and scattering result for small radiation solutions of 4NLS in any dimension $d \geq 1$, but unfortunately does not cover Eq.* (82)*, due to technical reasons.*

We defer a more detailed study of Eq. (82) *(and in general of the normal form equation* (96)*), together with the approximation of small radiation solutions of for the NLKG on $\mathbb{R}^d$, $d \geq 3$, up to times of order $\mathcal{O}(c^2)$ (or longer), to a future work.*

Remark 8.5. *We point out that the case of the one-dimensional defocusing NLKG is also interesting, since for $\lambda = 1$ the normalized equation at first step is the defocusing NLS, which is integrable. It would be interesting also to understand whether globally well-posedness and scattering hold also the normalized order 2 equation* (82)*, which we later exploit to approximate solutions of the NLKG up to times of order $\mathcal{O}(c^2)$.*

Even though there is a one-dimensional integrable 4NLS equation related to the dynamics of a vortex filament (see [S+ 03] and references therein),

$$i\psi_t + \psi_{xx} + \frac{1}{2}|\psi|^2\psi - \nu\left[\psi_{xxxx} + \frac{3}{2}|\psi|^2\psi_{xx} + \frac{3}{2}\psi_x^2\bar{\psi} + \frac{3}{8}|\psi|^4\psi + \frac{1}{2}(|\psi|^2)_{xx}\psi\right] = 0,\ \nu \in \mathbb{R} \tag{110}$$

apparently there is no obvious relation between the above equation and Eq. (82)*. Furthermore, while the issue of local well-posedness for one-dimensional fourth-order Nonlinear Schrödinger has been quite studied (see for example [HJ07]),* *there is only a recent result (see [RWZ16]) about global well-posedness and scattering for small radiation solutions of 4NLS, which unfortunately does not cover Eq.* (82)*, due to technical reasons.*

Therefore it seems difficult to give an explicit condition for global well-posedness and scattering for the normal form equation also in the one-dimensional case.

8.3 The nonlinear case: standing waves solutions

Now we consider the approximation of another important type of solutions, the so-called standing waves solutions.

The issue of (in)stability of standing waves and solitons has a long history: for the NLS equation and the NLKG the orbital stability of standing waves has been discussed first in [SS85]; for the NLS the orbital stability of one soliton solutions has been treated in [GSS87], while the asymptotic stability has been discussed in [Cuc01] for one soliton solutions, and in [RSS05] and [RSS03] for N-solitons. For the higher-order Schrödinger equation we mention [MS11], which deals with orbital stability of standing waves for fourth-order NLS-type equations. For the NLKG equation, the instability of solitons and standing waves has been studied in [SW99], [IKV06] and [OT07].

As for the case of radiation solution, we should fix $r \geq 1$, and consider a standing wave solution ψ_r of (96), namely of the form

$$\psi_r(t,x) = e^{it\omega}\eta_\omega(x), \tag{111}$$

where $\omega \in \mathbb{R}$, and $\eta_\omega \in \mathcal{S}(\mathbb{R}^3)$ solves

$$-\omega\eta_\omega = X_{H_{simp}}(\eta_\omega).$$

Remark 8.6. *Of course the existence of a standing wave for the simplified equation* (96) *is a far from trivial question (see [GSS87] for the NLS equation, and [MS11] for the fourth-order NLS-type equation).*

For $r = 1$ and $\lambda = 1$ (namely, the defocusing case), we can exploit the criteria in [GSS87] for existence and stability of standing waves for the NLS: we recall that if we fix $\omega > 0$ and we consider η_ω to be the ground state of the corresponding equation, we have that the standing wave solution is orbitally stable for $\frac{1}{2} < l < \frac{7}{6}$, and unstable for $\frac{7}{6} < l < \frac{5}{2}$.

We also point out that in the case of a standing wave solution, if $\delta(t)$ satisfies (101), then by Duhamel formula

$$\dot{\delta} = ic\langle\nabla\rangle_c\delta(t) + \mathrm{d}P(\psi_a(t), \bar{\psi}_a(t))\delta(t).$$

Since

$$P(e^{it\omega}\eta_\omega, e^{-it\omega}\bar{\eta}_\omega) = 2^{l-1/2}\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2} Re(e^{it\omega}\eta_\omega)\right]^{2l-1},$$

we have that

$$\mathrm{d}P(\eta_\omega, \bar{\eta}_\omega)e^{it\omega}h = 2^{l-1/2}\left(\frac{c}{\langle\nabla\rangle_c}\right)\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\cos(\omega t)\eta_\omega\right]^{2(l-1)}(e^{it\omega}h + e^{-it\omega}\bar{h}),$$

and by setting $\delta = e^{-it\omega}h$, one gets

$$-i\dot{h} = (c\langle\nabla\rangle_c + \omega)h + 2^{l-1/2}\cos^{2(l-1)}(\omega t)\left(\frac{c}{\langle\nabla\rangle_c}\right)\left[\left(\frac{c}{\langle\nabla\rangle_c}\right)^{1/2}\eta_\omega\right]^{2(l-1)}(h + e^{-2it\omega}\bar{h}) \tag{112}$$

$$+\left[\mathrm{d}P(\psi_a(s),\bar{\psi}_a(s)) - \mathrm{d}P(\eta_\omega,\bar{\eta}_\omega)\right]h. \tag{113}$$

Eq. (112) is a Salpeter spinless equation with a periodic time-dependent potential; therefore, in order to get some information about the error, one would need the corresponding Strichartz estimates for Eq. (112). Unfortunately, in the literature of dispersive estimates there are only few results for PDEs with time-dependent potentials, and the majority of them is of perturbative nature; for the Schrödinger equation we mention [DPV05] and [Gol09], in which Strichartz estimates are proved in a non-perturbative framework.

Remark 8.7. *By using Proposition 7.1 one can show that the NLKG can be approximated by the simplified equation* (7.1) *locally uniformly in time, up to an error of order* $\mathcal{O}(c^{-2r})$.

Remark 8.8. *One could ask whether one could get a similar result for more general (in particular, moving) soliton solution of* (96). *Apart from the issue of existence and stability for such solutions, one can check that, provided that a moving soliton solution for* (96) *exists, then the error* $\delta(t)$ *must solve a* (112)*-type equation, namely a spinless Salpeter equation with a time-dependent moving potential. Unfortunately, since Eq.* (112)*, unlike KG, is not manifestly covariant, one cannot apparently reduce to an analogue equation, and once again one cannot justify the approximation over the* $\mathcal{O}(1)$*-timescale.*

A Proof of Lemma 5.3

In order to normalize system (33), we used an adaptation of Theorem 4.4 in [Bam99]. The result is based on the method of Lie transform, that we will recall in the following.

Let $k \geq k_1$ and $p \in (1, +\infty)$ be fixed.
Given an auxiliary function χ analytic on $W^{k,p}$, we consider the auxiliary differential equation

$$\dot{\psi} = i\nabla_{\bar{\psi}}\chi(\psi,\bar{\psi}) =: X_\chi(\psi,\bar{\psi}) \tag{114}$$

and denote by Φ_χ^t its time-t flow. A simple application of Cauchy inequality gives

Lemma A.1. *Let* χ *and its symplectic gradient be analytic in* $B_{k,p}(\rho)$*. Fix* $\delta < \rho$*, and assume that*

$$\sup_{B_{k,p}(R-\delta)} \|X_\chi(\psi,\bar{\psi})\|_{k,p} \leq \delta.$$

Then, if we consider the time-t flow Φ_χ^t *of* X_χ *we have that for* $|t| \leq 1$

$$\sup_{B_{k,p}(R-\delta)} \|\Phi_\chi^t(\psi,\bar{\psi}) - (\psi,\bar{\psi})\|_{k,p} \leq \sup_{B_{k,p}(R-\delta)} \|X_\chi(\psi,\bar{\psi})\|_{k,p}.$$

Definition A.2. *The map* $\Phi := \Phi_\chi^1$ *will be called the* Lie transform *generated by* χ.

Remark A.3. *Given* G *analytic on* $W^{k,p}$, *consider the differential equation*

$$\dot{\psi} = X_G(\psi, \bar{\psi}), \tag{115}$$

where by X_G *we denote the vector field of* G. *Now define*

$$\Phi^* G(\phi, \bar{\phi}) := G \circ \Phi(\psi, \bar{\psi}).$$

In the new variables $(\phi, \bar{\phi})$ *defined by* $(\psi, \bar{\psi}) = \Phi(\phi, \bar{\phi})$ *equation* (115) *is equivalent to*

$$\dot{\phi} = X_{\Phi^* G}(\phi, \bar{\phi}). \tag{116}$$

Using the relation

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Phi_\chi^t)^* G = (\Phi_\chi^t)^* \{\chi, G\},$$

we formally get

$$\Phi^* G = \sum_{l=0}^{\infty} G_l, \tag{117}$$

$$G_0 := G, \tag{118}$$

$$G_l := \frac{1}{l}\{\chi, G_{l-1}\}, \quad l \geq 1. \tag{119}$$

In order to estimate the terms appearing in (117) we exploit the following results

Lemma A.4. *Let* $R > 0$, *and assume that* χ, G *are analytic on* $B_{k,p}(R)$. *Then, for any* $d \in (0, R)$ *we have that* $\{\chi, G\}$ *is analytic on* $B_{k,p}(R-d)$, *and*

$$\sup_{B_{k,p}(R-d)} \|X_{\{\chi,G\}}(\psi, \bar{\psi})\|_{k,p} \preceq \frac{2}{d}. \tag{120}$$

Lemma A.5. *Let* $R > 0$, *and assume that* χ, G *are analytic on* $B_{k,p}(R)$. *Let* $l \geq 1$, *and consider* G_l *as defined in* (117); *for any* $d \in (0, R)$ *we have that* G_l *is analytic on* $B_{k,p}(R-d)$, *and*

$$\sup_{B_{k,p}(R-d)} \|X_{G_l}(\psi, \bar{\psi})\|_{k,p} \preceq \left(\frac{2e}{d}\right)^l. \tag{121}$$

Proof. Fix l, and denote $\delta := d/l$. We look for a sequence $C_m^{(l)}$ such that

$$\sup_{B_{k,p}(R-m\delta)} \|X_{G_m}(\psi, \bar{\psi})\|_{k,p} \preceq C_m^{(l)}, \quad \forall m \leq l.$$

By (120) we can define the sequence

$$\begin{aligned} C_0^{(l)} &:= \sup_{B_{k,p}(R)} \|X_G(\psi, \bar{\psi})\|_{k,p}, \\ C_m^{(l)} &= \frac{2}{\delta m} C_{m-1}^{(l)} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p} \\ &= \frac{2l}{dm} C_{m-1}^{(l)} \sup_{B_{k,p}(R)} \|X_\chi(\psi, \bar{\psi})\|_{k,p}. \end{aligned}$$

One has

$$C_l^{(l)} = \frac{1}{l!}\left(\frac{2l}{d}\sup_{B_{k,p}(R)}\|X_\chi(\psi,\bar\psi)\|_{k,p}\right)^l \sup_{B_{k,p}(R)}\|X_G(\psi,\bar\psi)\|_{k,p},$$

and by using the inequality $l^l < l!e^l$ we can conclude. □

Remark A.6. *Let $k \geq k_1$, $p \in (1,+\infty)$, and assume that χ, F are analytic on $B_{k,p}(R)$. Fix $d \in (0,R)$, and assume also that*

$$\sup_{B_{k,p}(R)}\|X_\chi(\psi,\bar\psi)\|_{k,p} \leq d/3,$$

Then for $|t| \leq 1$

$$\sup_{B_{k,p}(R-d)}\|X_{(\Phi^t_\chi)^*F-F}(\psi,\bar\psi)\|_{k,p} = \sup_{B_{k,p}(R-d)}\|X_{F\circ\Phi^t_\chi-F}(\psi,\bar\psi)\|_{k,p} \tag{122}$$

$$\overset{(120)}{\leq} \frac{5}{d}\sup_{B_{k,p}(R)}\|X_\chi(\psi,\bar\psi)\|_{k,p}\sup_{B_{k,p}(R)}\|X_F(\psi,\bar\psi)\|_{k,p}. \tag{123}$$

Lemma A.7. *Let $k \geq k_1$, $p \in (1,+\infty)$, and assume that G is analytic on $B_{k,p}(R)$, and that h_0 satisfies PER. Then there exists χ analytic on $B_{k,p}(R)$ and Z analytic on $B_{k,p}(R)$ with Z in normal form, namely $\{h_0,Z\}=0$, such that*

$$\{h_0,\chi\} \; + \; G \; = \; Z. \tag{124}$$

Furthermore, we have the following estimates on the vector fields

$$\sup_{B_{k,p}(R)}\|X_Z(\psi,\bar\psi)\|_{k,p} \leq \sup_{B_{k,p}(R)}\|X_G(\psi,\bar\psi)\|_{k,p}, \tag{125}$$

$$\sup_{B_{k,p}(R)}\|X_\chi(\psi,\bar\psi)\|_{k,p} \preceq \sup_{B_{k,p}(R)}\|X_G(\psi,\bar\psi)\|_{k,p}. \tag{126}$$

Proof. One can check that the solution of (124) is

$$\chi(\psi,\bar\psi) = \frac{1}{T}\int_0^T t\,\left[G(\Phi^t(\psi,\bar\psi)) - Z(\Phi^t(\psi,\bar\psi))\right]\mathrm{d}t,$$

with $T = 2\pi$. Indeed,

$$\begin{aligned}
\{h_0,\chi\}(\psi,\bar\psi) &= \frac{\mathrm{d}}{\mathrm{d}s}_{|s=0}\chi(\Phi^s(\psi,\bar\psi)) \\
&= \frac{1}{2\pi}\int_0^{2\pi} t\frac{\mathrm{d}}{\mathrm{d}s}_{|s=0}\left[G(\Phi^{t+s}(\psi,\bar\psi)) - Z(\Phi^{t+s}(\psi,\bar\psi))\right]\mathrm{d}t \\
&= \frac{1}{2\pi}\int_0^{2\pi} t\frac{\mathrm{d}}{\mathrm{d}t}\left[G(\Phi^t(\psi,\bar\psi)) - Z(\Phi^t(\psi,\bar\psi))\right]\mathrm{d}t \\
&= \frac{1}{2\pi}\left[tG(\Phi^t(\psi,\bar\psi)) - tZ(\Phi^t(\psi,\bar\psi))\right]_{t=0}^{2\pi} - \frac{1}{2\pi}\int_0^{2\pi}\left[G(\Phi^t(\psi,\bar\psi)) - Z(\Phi^t(\psi,\bar\psi))\right]\mathrm{d}t \\
&= G(\psi,\bar\psi) - Z(\psi,\bar\psi).
\end{aligned}$$

Finally, (125) follows from the fact that

$$X_\chi(\psi,\bar{\psi}) = \frac{1}{T}\int_0^T t\,\Phi^{-t}\circ X_{G-Z}(\Phi^t(\psi,\bar{\psi})\mathrm{d}t$$

by applying property (25). $\square$

Lemma A.8. *Let $k \geq k_1$, $p \in (1,+\infty)$, and assume that G is analytic on $B_{k,p}(R)$, and that h_0 satisfies PER. Let χ be analytic on $B_{k,p}(R)$, and assume that it solves* (124)*. For any $l \geq 1$ denote by $h_{0,l}$ the functions defined recursively as in* (117) *from h_0. Then for any $d \in (0,R)$ one has that $h_{0,l}$ is analytic on $B_{k,p}(R-d)$, and*

$$\sup_{B_{k,p}(R-d)} \|X_{h_{0,l}}(\psi,\bar{\psi})\|_{k,p} \leq 2\sup_{B_{k,p}(R)} \|X_G(\psi,\bar{\psi})\|_{k,p}\left(\frac{5}{d}\sup_{B_{k,p}(R)} \|X_\chi(\psi,\bar{\psi})\|_{k,p}\right)^l. \tag{127}$$

Proof. By using (124) one gets that $h_{0,1} = Z - G$ is analytic on $B_{k,p}(R)$. Then by exploiting (123) one gets the result. $\square$

Lemma A.9. *Let $k_1 \gg 1$, $p \in (1,+\infty)$, $R > 0$, $m \geq 0$, and consider the Hamiltonian*

$$H^{(m)}(\psi,\bar{\psi}) = h_0(\psi,\bar{\psi}) + \epsilon\hat{h}(\psi,\bar{\psi}) + \epsilon Z^{(m)}(\psi,\bar{\psi}) + \epsilon^{m+1}F^{(m)}(\psi,\bar{\psi}). \tag{128}$$

Assume that h_0 satisfies PER and INV, that $\hat{h}$ satisfies NF, and that

$$\sup_{B_{k,p}(R)} \|X_{\hat{h}}(\psi,\bar{\psi})\|_{k,p} \leq F_0,$$
$$\sup_{B_{k,p}(R)} \|X_{F^{(0)}}(\psi,\bar{\psi})\|_{k,p} \leq F.$$

Fix $\delta < R/(m+1)$, and assume also that $Z^{(m)}$ are analytic on $B_{k,p}(R-m\delta)$, and that

$$\begin{aligned}
\sup_{B_{k,p}(R-m\delta)} \|X_{Z^{(0)}}(\psi,\bar{\psi})\|_{k,p} &= 0,\\
\sup_{B_{k,p}(R-m\delta)} \|X_{Z^{(m)}}(\psi,\bar{\psi})\|_{k,p} &\leq F\sum_{i=0}^{m-1}\epsilon^i K_s^i, \quad m\geq 1,\\
\sup_{B_{k,p}(R-m\delta)} \|X_{F^{(m)}}(\psi,\bar{\psi})\|_{k,p} &\leq F\,K_s^m, \quad m\geq 1,
\end{aligned} \tag{129}$$

with $K_s := \frac{2\pi}{\delta}(18F+5F_0)$.
Then, if $\epsilon K_s < 1/2$ there exists a canonical transformation $\mathcal{T}_\epsilon^{(m)}$ analytic on $B_{k,p}(R-(m+1)\delta)$ such that

$$\sup_{B_{k,p}(R-m\delta)} \|\mathcal{T}_\epsilon^{(m)}(\psi,\bar{\psi}) - (\psi,\bar{\psi})\|_{k,p} \leq 2\pi\epsilon^{m+1}F, \tag{130}$$

$H^{(m+1)} := H^{(m)}\circ\mathcal{T}^{(m)}$ has the form (128) *and satisfies* (129) *with m replaced by $m+1$.*

Proof. The key point of the lemma is to look for $\mathcal{T}_\epsilon^{(m)}$ as the time-one map of the Hamiltonian vector field of an analytic function $\epsilon^{m+1}\chi_m$. Hence, consider the differential equation

$$(\dot{\psi}, \dot{\bar{\psi}}) = X_{\epsilon^{m+1}\chi_m}(\psi, \bar{\psi}); \tag{131}$$

by standard theory we have that, if $\|X_{\epsilon^{m+1}\chi_m}\|_{B_{k,p}(R-m\delta)}$ is sufficiently small and $(\psi_0, \bar{\psi}_0) \in B_{k,p}(R-(m+1)\delta)$, then the solution of (131) exists for $|t| \leq 1$. Therefore we can define $\mathcal{T}^t_{m,\epsilon} : B_{k,p}(R-(m+1)\delta) \to B_{k,p}(R-m\delta)$, and in particular the corresponding time-one map $\mathcal{T}_\epsilon^{(m)} := \mathcal{T}^1_{m,\epsilon}$, which is an analytic canonical transformation, ϵ^{m+1}-close to the identity. We have

$$\begin{aligned}
&(\mathcal{T}_\epsilon^{(m+1)})^* \left(h_0 + \epsilon\hat{h} + \epsilon Z^{(m)} + \epsilon^{m+1}F^{(m)}\right) = h_0 + \epsilon\hat{h} + \epsilon Z^{(m)} \\
&\quad + \epsilon^{m+1}\left[\{\chi_m, h_0\} + F^{(m)}\right] + \\
&\quad + \left(h_0 \circ \mathcal{T}^{(m+1)} - h_0 - \epsilon^{m+1}\{\chi_m, h_0\}\right) + \epsilon(\hat{h} \circ \mathcal{T}^{(m+1)} - \hat{h}) + \epsilon\left(Z^{(m)} \circ \mathcal{T}^{(m+1)} - Z^{(m)}\right) \qquad (132)\\
&\quad + \epsilon^{m+1}\left(F^{(m)} \circ \mathcal{T}^{(m+1)} - F^{(m)}\right). \qquad (133)
\end{aligned}$$

It is easy to see that the first three terms are already normalized, that the term in the second line is the non-normalized part of order m+1 that will vanish through the choice of a suitable χ_m, and that the last lines contains all the terms of order higher than m+1.

Now we want to determine χ_m in order to solve the so-called "homological equation"

$$\{\chi_m, h_0\} + F^{(m)} = Z_{m+1},$$

with Z_{m+1} in normal form. The existence of χ_m and Z_{m+1} is ensured by Lemma A.7, and by applying (125) and the inductive hypothesis we get

$$\sup_{B_{k,p}(R-m\delta)} \|X_{\chi_m}(\psi, \bar{\psi})\|_{k,p} \leq 2\pi F, \tag{134}$$

$$\sup_{B_{k,p}(R-m\delta)} \|X_{Z_{m+1}}(\psi, \bar{\psi})\|_{k,p} \leq 2\pi F. \tag{135}$$

Now define $Z^{(m+1)} := Z^{(m)} + \epsilon^m Z_{m+1}$, and notice that by Lemma A.1 we can deduce the estimate of $X_{Z^{(m+1)}}$ on $B_{k,p}(R-(m+1)\delta)$ and (130) at level $m+1$. Next, set $\epsilon^{m+2}F^{(m+1)} :=$ (132) + (133). Then we can use (123) and (127), in order to get

$$\begin{aligned}
&\sup_{B_{k,p}(R-(m+1)\delta)} \|X_{\epsilon^{m+2}F^{(m+1)}}(\psi, \bar{\psi})\|_{k,p} \qquad (136)\\
&\leq \left(\frac{10}{\delta}\epsilon^m K_s^m\,\epsilon F + \frac{5}{\delta}\epsilon F_0 + \frac{5}{\delta}\epsilon F\sum_{i=0}^{m-1}\epsilon^i K_s^i + \frac{5}{\delta}\epsilon F\,\epsilon^m K_s^m\right)\epsilon^{m+1}\sup_{B_{k,p}(R-m\delta)}\|X_{\chi_m}(\psi, \bar{\psi})\|_{k,p} \\
&= \epsilon^{m+2}\left(\frac{10}{\delta}\epsilon^m K_s^m\,F + \frac{5}{\delta}F_0 + \frac{5}{\delta}F\sum_{i=0}^{m-1}\epsilon^i K_s^i + \frac{5}{\delta}F\,\epsilon^m K_s^m\right)\sup_{B_{k,p}(R-m\delta)}\|X_{\chi_m}(\psi, \bar{\psi})\|_{k,p}. \qquad (137)
\end{aligned}$$

If $m=0$, then the third term is not present, and (137) reads

$$\sup_{B_{k,p}(R-\delta)} \|X_{\epsilon^2 F^{(1)}}(\psi,\bar\psi)\|_{k,p} \le \epsilon^2 \left(\frac{15}{\delta}F + \frac{5}{\delta}F\right) 2\pi F < \epsilon^2 K_s F.$$

If $m \ge 1$, we exploit the smallness condition $\epsilon K_s < 1/2$, and (137) reads

$$\sup_{B_{k,p}(R-(m+1)\delta)} \|X_{\epsilon^{m+2} F^{(m+1)}}(\psi,\bar\psi)\|_{k,p} < \left(\frac{18}{\delta}\epsilon F + \frac{5}{\delta}\epsilon F_0\right) 2\pi\, \epsilon F\, \epsilon^m K_s^m = \epsilon^{m+2}\, F K_s^{m+1}.$$

□

Now fix $R>0$.

Proof. (of Lemma 5.3) The Hamiltonian (33) satisfies the assumptions of Lemma A.9 with $m=0$, $F_{N,r}$ in place of $F^{(0)}$ and $h_{N,r}$ in place of $\hat h$, $F = K_{k,p}^{(F,r)}\, r 2^{2Nr}$, $F_0 = K_{k,p}^{(h,r)}\, r 2^{2Nr}$ (for simplicity we will continue to denote by F and F_0 the last two quantities). So we apply Lemma A.9 with $\delta = R/4$, provided that

$$\frac{8\pi}{R}(18F + 5F_0)\epsilon < \frac{1}{2},$$

which is true due to (44). Hence there exists an analytic canonical transformation $\mathcal{T}_{\epsilon,N}^{(1)} : B_{k,p}(3R/4) \to B_{k,p}(R)$ with

$$\sup_{B_{k,p}(3R/4)} \|\mathcal{T}_{\epsilon,N}^{(1)}(\psi,\bar\psi) - (\psi,\bar\psi)\|_{k,p} \le 2\pi F\,\epsilon,$$

such that

$$H_{N,r} \circ \mathcal{T}_{\epsilon,N}^{(1)} = h_0 + \epsilon h_{N,r} + \epsilon Z_N^{(1)} + \epsilon^2 \mathcal{R}_N^{(1)}, \tag{138}$$

$$Z_N^{(1)} := \langle F_{N,r}\rangle, \tag{139}$$

$$\begin{aligned}\epsilon^2 \mathcal{R}_N^{(1)} &:= \epsilon^2 F^{(1)} \\ &= \left(h_0 \circ \mathcal{T}_{\epsilon,N}^{(1)} - h_0 - \epsilon\{\chi_1, h_0\}\right) + \epsilon(\hat h_{N,r} \circ \mathcal{T}_{\epsilon,N}^{(1)} - \hat h_{N,r}) + \epsilon\left(Z_N^{(1)} \circ \mathcal{T}_{\epsilon,N}^{(1)} - Z_N^{(1)}\right) \\ &\quad + \epsilon^2\left(F_{N,r} \circ \mathcal{T}_{\epsilon,N}^{(1)} - F_{N,r}\right),\end{aligned} \tag{140}$$

$$\sup_{B_{k,p}(3R/4)} \|X_{h_{N,r}+Z_N^{(1)}}(\psi,\bar\psi)\|_{k,p} \le F_0 + F =: \tilde F_0, \tag{141}$$

$$\sup_{B_{k,p}(3R/4)} \|X_{\mathcal{R}_N^{(1)}}(\psi,\bar\psi)\|_{k,p} \le \frac{8\pi}{R}(18F + 5F_0)F =: \tilde F. \tag{142}$$

Again (138) satisfies the assumptions of Lemma A.9 with $m=0$, and $h_{N,r}+Z_N^{(1)}$ and $\mathcal{R}_N^{(1)}$ in place of $F^{(0)}$ and $\hat h$.
Now fix $\delta := \delta(R) = \frac{R}{4r}$, and apply r times Lemma A.9; we get an Hamiltonian of the form (45), such that

$$\sup_{B_{k,p}(R/2)} \|X_{Z_N^{(r)}}(\psi,\bar\psi)\|_{k,p} \le 2\tilde F, \tag{143}$$

$$\sup_{B_{k,p}(R/2)} \|X_{\mathcal{R}_N^{(r)}}(\psi,\bar\psi)\|_{k,p} \le \tilde F. \tag{144}$$

□

B Interpolation theory for relativistic Sobolev spaces

In this section we show an analogue of Theorem 6.4.5 (7) in [BL76] for the relativistic Sobolev spaces $\mathscr{W}_c^{k,p}$, $k \in \mathbb{R}$, $1 < p < +\infty$. We recall that

$$\mathscr{W}_c^{k,p}(\mathbb{R}^3) := \left\{ u \in L^p : \|u\|_{\mathscr{W}_c^{k,p}} := \|c^{-k} \langle \nabla \rangle_c^k u\|_{L^p} < +\infty \right\}, \quad k \in \mathbb{R}, \quad 1 < p < +\infty.$$

In order to state the main result of this section, we exploit notations and well known results coming from complex interpolation theory (see [BL76] for a detailed introduction to this topic).

In order to study the relativistic Sobolev spaces, we have to recall the notion of Fourier multipliers.

Definition B.1. *Let $1 < p < +\infty$, and $\rho \in \mathcal{S}'$. We call ρ a* Fourier multiplier *on $L^p(\mathbb{R}^d)$ if the convolution $(\mathcal{F}^{-1}\rho) * f \in L^p(\mathbb{R}^d)$ for all $f \in L^p(\mathbb{R}^d)$, and if*

$$\sup_{\|f\|_{L^p}=1} \|(\mathcal{F}^{-1}\rho) * f\|_{L^p} < +\infty. \tag{145}$$

The linear space of all such ρ is denoted by M_p, and is endowed with the above norm (145)*.*

One can check that for any $p \in (1, +\infty)$ one has $M_p = M_{p'}$ (where $1/p + 1/p' = 1$), and that by Parseval's formula $M_2 = L^\infty$. Furthermore, by Riesz-Thorin theorem one gets that for any $\rho \in M_{p_0} \cap M_{p_1}$ and for any $\theta \in (0,1)$

$$\|\rho\|_{M_p} \leq \|\rho\|_{M_{p_0}}^{1-\theta} \|\rho\|_{M_{p_1}}^{\theta}, \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}. \tag{146}$$

In particular, one can deduce that $\|\cdot\|_{M_p}$ decreases with $p \in (1,2]$, and that $M_p \subset M_q$ for any $1 < p < q \leq 2$.

More generally, if H_0 and H_1 are Hilbert spaces, one can introduce a similar definition of Fourier multiplier. We use the notation $\mathcal{S}'(H_0, H_1)$ in order to denote the space of all linear continous maps from $\mathcal{S}(\mathbb{R}^d, H_0)$ to H_1.

Definition B.2. *Let $1 < p < +\infty$, let H_0 and H_1 be two Hilbert spaces, and consider $\rho \in \mathcal{S}'(H_0, H_1)$. We call ρ a* Fourier multiplier *if the convolution $(\mathcal{F}^{-1}\rho) * f \in L^p(H_1)$ for all $f \in L^p(H_0)$, and if*

$$\sup_{\|f\|_{L^p(H_0)}=1} \|(\mathcal{F}^{-1}\rho) * f\|_{L^p(H_1)} < +\infty. \tag{147}$$

The linear space of all such ρ is denoted by $M_p(H_0, H_1)$, and is endowed with the above norm (147)*.*

Next we recall *Mihlin multipier theorem* (Theorem 6.1.6 in [BL76]).

Theorem B.3. *Let H_0 and H_1 be Hilbert spaces, and assume that $\rho : \mathbb{R}^d \to L(H_0, H_1)$ be such that*

$$|\xi|^\alpha \|D^\alpha \rho(\xi)\|_{L(H_0,H_1)} \leq K, \quad \forall \xi \in R^d, |\alpha| \leq L$$

for some integer $L > d/2$. Then $\rho \in M_p(H_0, H_1)$ for any $1 < p < +\infty$, and

$$\|\rho\|_{M_p} \leq C_p K, \quad 1 < p < +\infty.$$

Now, recall the Littlewood-Paley functions $(\phi_j)_{j\geq 0}$ defined in (24), and introduce the maps $\mathcal{J} : \mathcal{S}' \to \mathcal{S}'$ and $\mathcal{P} : \mathcal{S}' \to \mathcal{S}'$ via formulas

$$(\mathcal{J}f)_j := \phi_j * f, \;\; j \geq 0, \tag{148}$$

$$\mathcal{P}g := \sum_{j\geq 0} \tilde{\phi}_j * g_j, \;\; j \geq 0, \tag{149}$$

where $g = (g_j)_{j\geq 0}$ with $g_j \in \mathcal{S}'$ for all j, and

$$\tilde{\phi}_0 := \phi_0 + \phi_1,$$
$$\tilde{\phi}_j := \phi_{j-1} + \phi_j + \phi_{j+1}, \;\; j \geq 1.$$

One can check that $\mathcal{P} \circ \mathcal{J}f = f \; \forall f \in \mathcal{S}'$, since $\tilde{\phi}_j * \phi_j = \phi_j$ for all j. We then introduce for $c \geq 1$ and $k \geq 0$ the space

$$l_c^{2,k} := \{(z_j)_{j\in\mathbb{Z}} \; : \; c^{-k} \sum_{j\in\mathbb{Z}} (c^2 + |j|^2)^k |z_j|^2 < +\infty\}.$$

Theorem B.4. *Let $c \geq 1$, $k \geq 0$, $1 < p < +\infty$. Then $\langle\nabla\rangle_c^k L^p$ is a retract of $L^p(l_c^{2,k})$, namely that the operators*

$$\mathcal{J} : \mathscr{W}_c^{k,p} \to L^p(l_c^{2,k})$$
$$\mathcal{P} : L^p(l_c^{2,k}) \to \mathscr{W}_c^{k,p}$$

satisfy $\mathcal{P} \circ \mathcal{J} = id$ on $\mathscr{W}_c^{k,p}$.

Proof. First we show that $\mathcal{J} : \mathscr{W}_c^{k,p} \to L^p(l_c^{2,k})$ is bounded.
Since $\mathcal{J}f = (\mathcal{F}^{-1}\chi_c) * \mathcal{J}_c^k f$, where

$$(\chi_c(\xi))_j := (c^2 + |\xi|^2)^{-k/2} \hat{\phi}_j(\xi), \;\; j \geq 0$$
$$\mathcal{J}_c^k f := \mathcal{F}^{-1}((c^2 + |\xi|^2)^{k/2} \hat{f}),$$

we have that for any $\alpha \in \mathbb{N}^d$

$$|\xi|^\alpha \|D^\alpha \chi_c(\xi)\|_{L(\mathbb{C}, l_c^{2,k})} \leq |\xi|^\alpha \sum_{j\geq 0} (2^{jk} c^k |D^\alpha (\chi_c(\xi))_j|) \leq K_\alpha$$

because the sum consists of at most two non-zero terms for each ξ. Thus $\mathcal{J} \in M_p(\mathscr{W}_c^{k,p}, L^p(l_c^{2,k}))$ by Mihlin multiplier Theorem.

On the other hand, consider $\mathcal{P} : L^p(l_c^{2,k}) \to \mathscr{W}_c^{k,p}$.
Since $\mathcal{J}_c^k \circ \mathcal{P}g = (\mathcal{F}^{-1}\delta_c) * g_{(k)}$, where

$$g = (g_j)_{j\geq 0},$$
$$g_{(k)} := (2^{jk} g_j)_{j\geq 0},$$
$$\delta_c(\xi) g := \sum_{j\geq 0} 2^{-jk} (c^2 + |\xi|^2)^{k/2} \tilde{\phi}_j(\xi) g_j,$$

we have that for any $\alpha \in \mathbb{N}^d$

$$|\xi|^\alpha \|D^\alpha \delta_c(\xi)\|_{L(l_c^{2,k}, \mathbb{C})} \leq |\xi|^\alpha \left[\sum_{j\geq 0} (2^{-jk} c^{-k} |D^\alpha (c^2 + |\xi|^2)^{k/2} \tilde{\phi}_j(\xi)|)^2 \right]^{1/2} \leq K_\alpha,$$

because the sum consists of at most four non-zero terms for each ξ. Thus $\mathcal{P} \in M_p(L^p(l_c^{2,k}), \mathscr{W}_c^{k,p})$ by Mihlin multiplier Theorem, and we can conclude. □ □

Corollary B.5. *Let $\theta \in (0,1)$, and assume that k_0, $k_1 \geq 0$ ($k_0 \neq k_1$) and p_0, $p_1 \in (1,+\infty)$ satisfy*

$$k = (1-\theta)k_0 + \theta k_1,$$
$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

Then $(\mathscr{W}_c^{k_0,p}, \mathscr{W}_c^{k_1,p})_\theta = \mathscr{W}_c^{k,p}$.

The previous corollary, combined with the classical *3 lines theorem* (Lemma 1.1.2 in [BL76]), immediately leads us to the following Proposition.

Proposition B.6. *Let $k_0 \neq k_1$, $1 < p < +\infty$, and assume that $T : \mathscr{W}_c^{k_0,p} \to \mathscr{W}_c^{k_0,p}$ has norm M_0, and that $T : \mathscr{W}_c^{k_1,p} \to \mathscr{W}_c^{k_1,p}$ has norm M_1. Then*

$$T : \mathscr{W}_c^{k,p} \to \mathscr{W}_c^{k,p},\ k = (1-\theta)k_0 + \theta k_1,$$

with norm $M \leq M_0^{1-\theta} M_1^\theta$.

Now we conclude with the proof of Theorem 3.7.

Theorem 3.7. Estimates (21) clearly follow from Proposition 3.1 if we can prove that for any α and for any $q \in [2,6]$

$$\|\langle\nabla\rangle_c^\alpha \mathcal{W}_\pm \langle\nabla\rangle_c^{-\alpha}\|_{L^q \to L^q} \preceq 1, \tag{150}$$
$$\|\langle\nabla\rangle_c^\alpha \mathcal{Z}_\pm \langle\nabla\rangle_c^{-\alpha}\|_{L^q \to L^q} \preceq 1. \tag{151}$$

Indeed in this case one would have

$$\|\langle\nabla\rangle_c^{1/q-1/p}\, e^{it\mathcal{H}(x)} P_c(-\Delta+V)\psi_0\|_{L_t^p L_x^q} = \|\langle\nabla\rangle_c^{1/q-1/p}\, \mathcal{W}_\pm e^{it\langle\nabla\rangle_c} \mathcal{Z}_\pm \psi_0\|_{L_t^p L_x^q},$$

but

$$\|\langle\nabla\rangle_c^{1/q-1/p}\, \mathcal{W}_\pm e^{it\langle\nabla\rangle_c} \mathcal{Z}_\pm \psi_0\|_{L_x^q} \preceq \|\langle\nabla\rangle_c^{1/q-1/p}\, e^{it\langle\nabla\rangle_c} \mathcal{Z}_\pm \psi_0\|_{L_x^q},$$

hence

$$\|\langle\nabla\rangle_c^{1/q-1/p}\, e^{it\mathcal{H}(x)} P_c(-\Delta+V)\psi_0\|_{L_t^p L_x^q} \preceq c^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}} \|\langle\nabla\rangle_c^{1/2} \mathcal{Z}_\pm \psi_0\|_{L^2} \preceq c^{\frac{1}{q}-\frac{1}{p}-\frac{1}{2}} \|\langle\nabla\rangle_c^{1/2} \psi_0\|_{L^2}.$$

To prove (151) we first show that it holds for $\alpha = 2k$, $k \in \mathbb{N}$. We argue by induction. The case $k = 0$ is true by Theorem 3.8. Now, suppose that (151) holds for $\alpha = 2(k-1)$, then

$$\begin{aligned}
&\|(c^2-\Delta)^k \mathcal{Z}_\pm (c^2-\Delta)^{-k}\|_{L^q\to L^q} = \|(c^2-\Delta)(c^2-\Delta)^{k-1} \mathcal{Z}_\pm (c^2-\Delta)^{-(k-1)} (c^2-\Delta)^{-1}\|_{L^q\to L^q} \\
&\leq c^2 \|(c^2-\Delta)^{k-1} \mathcal{Z}_\pm (c^2-\Delta)^{-(k-1)} (c^2-\Delta)^{-1}\|_{L^q\to L^q} \\
&\quad + \|-\Delta (c^2-\Delta)^{k-1} \mathcal{Z}_\pm (c^2-\Delta)^{-(k-1)} (c^2-\Delta)^{-1}\|_{L^q\to L^q} \\
&\leq c^2 \|(c^2-\Delta)^{k-1} \mathcal{Z}_\pm (c^2-\Delta)^{-(k-1)} (c^2-\Delta)^{-1}\|_{L^q\to L^q} \\
&\quad + \|-\Delta (c^2-\Delta)^{-1}\, (c^2-\Delta)^{k-1} \mathcal{Z}_\pm (c^2-\Delta)^{-(k-1)}\|_{L^q\to L^q} \\
&+ \|-\Delta (c^2-\Delta)^{k-1} [\mathcal{Z}_\pm, (c^2-\Delta)^{-1}] (c^2-\Delta)^{-(k-1)}\|_{L^q\to L^q} \\
&\preceq c^2 \|(c^2-\Delta)^{-1}\|_{L^q\to L^q} + \|-\Delta (c^2-\Delta)^{-1}\|_{L^q\to L^q} \preceq 1,
\end{aligned}$$

since

$$\|[\mathcal{Z}_\pm, (c^2 - \Delta)^{-1}]\|_{L^2 \to L^2} \preceq \frac{|\xi|}{(c^2 + |\xi|^2)^2} \leq (c^2 + |\xi|^2)^{-3/2}.$$

Similarly we can show (151) for $\alpha = -2k$, $k \in \mathbb{N}$. By Proposition B.6 one can extend the result to any $\alpha \in \mathbb{R}$ via interpolation theory. □ □

References

[AC07] Thomas Alazard and Rémi Carles. Semi-classical limit of Schrödinger–Poisson equations in space dimension n ≥ 3. *Journal of Differential Equations*, 233(1):241–275, 2007.

[BAKS00] Matania Ben-Artzi, Herbert Koch, and Jean-Claude Saut. Dispersion estimates for fourth order Schrödinger equations. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 330(2):87–92, 2000.

[Bam99] Dario Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Mathematische Zeitschrift*, 230(2):345–387, 1999.

[Bam05] Dario Bambusi. Galerkin averaging method and Poincaré normal form for some quasilinear PDEs. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 4(4):669–702, 2005.

[BC11] Dario Bambusi and Scipio Cuccagna. On dispersion of small energy solutions of the nonlinear Klein Gordon equation with a potential. *American journal of mathematics*, 133(5):1421–1468, 2011.

[BCP02] Dario Bambusi, Andrea Carati, and Antonio Ponno. The nonlinear Schrödinger equation as a resonant normal form. *DISCRETE AND CONTINUOUS DYNAMICAL SYSTEMS SERIES B*, 2(1):109–128, 2002.

[BD12] Weizhu Bao and Xuanchun Dong. Analysis and comparison of numerical methods for the Klein–Gordon equation in the nonrelativistic limit regime. *Numerische Mathematik*, 120(2):189–229, 2012.

[BFS16] Simon Baumstark, Erwan Faou, and Katharina Schratz. Uniformly accurate exponential-type integrators for klein-gordon equations with asymptotic convergence to classical splitting schemes in the nonlinear schroedinger limit. *arXiv preprint arXiv:1606.04652*, 2016.

[BGT04] Nicolas Burq, Pierre Gérard, and Nikolay Tzvetkov. Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds. *American Journal of Mathematics*, 126(3):569–605, 2004.

[BL76] Jöran Bergh and Jorgen Lofstrom. Interpolation spaces: an introduction. *Springer Verlag, Newe*, 1976.

[BMS04] Philippe Bechouche, Norbert J Mauser, and Sigmund Selberg. Nonrelativistic limit of Klein-Gordon-Maxwell to Schrödinger-Poisson. *American journal of mathematics*, 126(1):31–64, 2004.

[Bou10] Jean-Marc Bouclet. Littlewood-Paley decompositions on manifolds with ends. *Bull. Soc. Math. France*, 138(1):1–37, 2010.

[BP06] Dario Bambusi and Antonio Ponno. On metastability in FPU. *Communications in mathematical physics*, 264(2):539–561, 2006.

[BZ16] Weizhu Bao and Xiaofei Zhao. A uniformly accurate (UA) multiscale time integrator Fourier pseudospectral method for the Klein–Gordon–Schrödinger equations in the nonrelativistic limit regime. *Numerische Mathematik*, pages 1–41, 2016.

[CG07] Shangbin Cui and Cuihua Guo. Well-posedness of higher-order nonlinear Schrödinger equations in Sobolev spaces Hs (Rn) and applications. *Nonlinear Analysis: Theory, Methods & Applications*, 67(3):687–707, 2007.

[CLM15] Rémi Carles, Wolfgang Lucha, and Emmanuel Moulay. Higher-order Schrödinger and Hartree–Fock equations. *Journal of Mathematical Physics*, 56(12):122301, 2015.

[CM12] Rémi Carles and Emmanuel Moulay. Higher order Schrödinger equations. *Journal of Physics A: Mathematical and Theoretical*, 45(39):395304, 2012.

[CO06] Yonggeun Cho and Tohru Ozawa. On the semirelativistic Hartree-type equation. *SIAM journal on mathematical analysis*, 38(4):1060–1074, 2006.

[CS16] Woocheol Choi and Jinmyoung Seok. Nonrelativistic limit of standing waves for pseudo-relativistic nonlinear Schrödinger equations. *Journal of Mathematical Physics*, 57(2):021510, 2016.

[Cuc01] Scipio Cuccagna. Stabilization of solutions to nonlinear Schrödinger equations. *Communications on Pure and Applied Mathematics*, 54(9):1110–1145, 2001.

[CZ11] Elena Cordero and Davide Zucco. Strichartz estimates for the vibrating plate equation. *Journal of Evolution Equations*, 11(4):827–845, 2011.

[DF08] Piero D'Ancona and Luca Fanelli. Strichartz and smoothing estimates for dispersive equations with magnetic potentials. *Communications in Partial Differential Equations*, 33(6):1082–1112, 2008.

[DPV05] Piero D'Ancona, Vittoria Pierfelice, and Nicola Visciglia. Some remarks on the Schrödinger equation with a potential in $L^r_t L^s_x$. *Mathematische Annalen*, 333(2):271–290, 2005.

[FS14] Erwan Faou and Katharina Schratz. Asymptotic preserving schemes for the Klein–Gordon equation in the non-relativistic limit regime. *Numerische Mathematik*, 126(3):441–469, 2014.

[Gol09] Michael Goldberg. Strichartz estimates for the Schrödinger equation with time-periodic $L^{n/2}$ potentials. *Journal of Functional Analysis*, 256(3):718–746, 2009.

[GSS87] Manoussos Grillakis, Jalal Shatah, and Walter Strauss. Stability theory of solitary waves in the presence of symmetry, I. *Journal of Functional Analysis*, 74(1):160–197, 1987.

[HJ07] Zhaohui Huo and Yueling Jia. A refined well-posedness for the fourth-order nonlinear Schrödinger equation related to the vortex filament. *Communications in Partial Differential Equations*, 32(10):1493–1510, 2007.

[HJ11] Zhaohui Huo and Yueling Jia. Well-posedness for the fourth-order nonlinear derivative Schrödinger equation in higher dimension. *Journal de mathématiques pures et appliquées*, 96(2):190–206, 2011.

[IKV06] Valery Imaikin, Alexander Komech, and Boris Vainberg. On scattering of solitons for the Klein–Gordon equation coupled to a particle. *Communications in mathematical physics*, 268(2):321–367, 2006.

[KAY12] JinMyong Kim, Anton Arnold, and Xiaohua Yao. Global estimates of fundamental solutions for higher-order Schrödinger equations. *Monatshefte für Mathematik*, 168(2):253–266, 2012.

[Läm93] Claus Lämmerzahl. The pseudodifferential operator square root of the Klein–Gordon equation. *Journal of mathematical physics*, 34(9):3918–3932, 1993.

[LZ16] Yong Lu and Zhifei Zhang. Partially strong transparency conditions and a singular localization method in geometric optics. *Archive for Rational Mechanics and Analysis*, 222(1):245–283, 2016.

[Mac01] Shuji Machihara. The nonrelativistic limit of the nonlinear Klein-Gordon equation. *FUNKCIALAJ EKVACIOJ SERIO INTERNACIA*, 44(2):243–252, 2001.

[MN02] Nader Masmoudi and Kenji Nakanishi. From nonlinear Klein-Gordon equation to a system of coupled nonlinear Schrödinger equations. *Mathematische Annalen*, 324(2):359–389, 2002.

[MN03] Nader Masmoudi and Kenji Nakanishi. Nonrelativistic limit from Maxwell-Klein-Gordon and Maxwell-Dirac to Poisson-Schrödinger. *International Mathematics Research Notices*, 2003(13):697–734, 2003.

[MN08] Nader Masmoudi and Kenji Nakanishi. Energy convergence for singular limits of Zakharov type systems. *Inventiones mathematicae*, 172(3):535–583, 2008.

[MN10] Nader Masmoudi and Kenji Nakanishi. From the Klein–Gordon–Zakharov system to a singular nonlinear Schrödinger system. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 27(4):1073–1096, 2010.

[MNO02] Shuji Machihara, Kenji Nakanishi, and Tohru Ozawa. Nonrelativistic limit in the energy space for nonlinear Klein-Gordon equations. *Mathematische Annalen*, 322(3):603–621, 2002.

[MS11] Masaya Maeda and Jun-ichi Segata. Existence and stability of standing waves of fourth order nonlinear Schrödinger type equation related to vortex filament. *Funkcialaj Ekvacioj*, 54(1):1–14, 2011.

[N$^+$08] Kenji Nakanishi et al. Transfer of global wellposedness from nonlinear Klein-Gordon equation to nonlinear Schrödinger equation. *Hokkaido Mathematical Journal*, 37(4):749–771, 2008.

[Naj90] Branko Najman. The nonrelativistic limit of the nonlinear Klein-Gordon equation. *Nonlinear Analysis: Theory, Methods & Applications*, 15(3):217–228, 1990.

[Nak02] Kenji Nakanishi. Nonrelativistic limit of scattering theory for nonlinear Klein–Gordon equations. *Journal of Differential Equations*, 180(2):453–470, 2002.

[OT07] Masahito Ohta and Grozdena Todorova. Strong instability of standing waves for the nonlinear Klein–Gordon equation and the Klein–Gordon–Zakharov system. *SIAM Journal on Mathematical Analysis*, 38(6):1912–1931, 2007.

[PX13] Benoit Pausader and Suxia Xia. Scattering theory for the fourth-order Schrödinger equation in low dimensions. *Nonlinearity*, 26(8):2175–2191, 2013.

[RSS03] I Rodnianski, W Schlag, and A Soffer. Asymptotic stability of N-soliton states of NLS. *arXiv preprint math/0309114*, 2003.

[RSS05] Igor Rodnianski, Wilhelm Schlag, and Avraham Soffer. Dispersive analysis of charge transfer models. *Communications on pure and applied mathematics*, 58(2):149–216, 2005.

[RT87] Didier Robert and Hideo Tamura. Semi-classical estimates for resolvents and asymptotics for total scattering cross-sections. *Annales de l'IHP Physique théorique*, 46(4):415–442, 1987.

[RWZ16] Michael Ruzhansky, Baoxiang Wang, and Hua Zhang. Global well-posedness and scattering for the fourth order nonlinear Schrödinger equations with small data in modulation and Sobolev spaces. *Journal de Mathématiques Pures et Appliquées*, 105(1):31–65, 2016.

[S$^+$03] J Segata et al. Well-posedness for the fourth-order nonlinear Schrödinger-type equation related to the vortex filament. *Differential and Integral Equations*, 16(7):841–864, 2003.

[Sch10] Guido Schneider. Bounds for the nonlinear Schrödinger approximation of the Fermi–Pasta–Ulam system. *Applicable Analysis*, 89(9):1523–1539, 2010.

[SS85] Jalal Shatah and Walter Strauss. Instability of nonlinear bound states. *Communications in Mathematical Physics*, 100(2):173–190, 1985.

[SS03] Elias M Stein and Rami Shakarchi. Complex analysis. Princeton Lectures in Analysis, II, 2003.

[Suc63] J Sucher. Relativistic Invariance and the Square-Root Klein-Gordon Equation. *Journal of Mathematical Physics*, 4(1):17–23, 1963.

[SW99] A Soffer and Michael I Weinstein. Resonances, radiation damping and instabilitym in Hamiltonian nonlinear wave equations. *Inventiones mathematicae*, 136(1):9–74, 1999.

[Tay11] ME Taylor. Partial differential equations III. Nonlinear equations. Second. vol. 117. *Applied Mathematical Sciences. Springer, New York, pp. xxii*, 715, 2011.

[Tsu84] Masayoshi Tsutsumi. Nonrelativistic approximation of nonlinear Klein-Gordon equations in two space dimensions. *Nonlinear Analysis: Theory, Methods & Applications*, 8(6):637–643, 1984.

[Yaj95] Kenji Yajima. The $W^{k,p}$-continuity of wave operators for Schrödinger operators. *Journal of the Mathematical Society of Japan*, 47(3):551–581, 1995.