Communications in Partial Differential Equations, 39: 1479–1522, 2014

ISSN 0360-5302 print/1532-4133 online
DOI: 10.1080/03605302.2013.860988

Multi-Solitary Waves for the Nonlinear Klein-Gordon Equation

JACOPO BELLAZZINI[1], MARCO GHIMENTI[2], AND STEFAN LE COZ[3]

[1]Università degli Studi di Sassari, Sassari, Italy
[2]Dipartimento di Matematica, Università di Pisa, Pisa, Italy
[3]Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse, France

We consider the nonlinear Klein-Gordon equation in $\mathbb{R}^d$. We call multi-solitary waves a solution behaving at large time as a sum of boosted standing waves. Our main result is the existence of such multi-solitary waves, provided the composing boosted standing waves are stable. It is obtained by solving the equation backward in time around a sequence of approximate multi-solitary waves and showing convergence to a solution with the desired property. The main ingredients of the proof are finite speed of propagation, variational characterizations of the profiles, modulation theory and energy estimates.

Keywords Asymptotic behavior; Klein-Gordon equation; Multi-soliton.

2010 Mathematics Subject Classification 35Q51; 35C08; 35Q40; 35L71; 37K40.

1. Introduction

We consider the following nonlinear Klein-Gordon equation

$$u_{tt} - \Delta u + mu - |u|^{p-1}u = 0 \qquad \text{(NLKG)}$$

where $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{C}$, $m \in (0, +\infty)$, and the nonlinearity is H^1-subcritical, i.e., $1 < p < 1 + \frac{4}{d-2}$ if $d \geq 3$ or $1 < p < +\infty$ if $d = 1, 2$.

This equation arises in particular in quantum physics where it has been proposed as a simple model describing a self-interacting scalar field. Mathematically speaking, the Klein-Gordon equation is one of the model dispersive equations. It is a Hamiltonian equation which is invariant under gauge and Lorentz transform and in particular it conserves energy, charge and momentum. Due to the sign

Received May 7, 2013; Accepted October 21, 2013
Address correspondence to Stefan Le Coz, Institut de Mathématiques de Toulouse, Université Paul Sabatier 118 route de Narbonne, 31062 Toulouse Cedex 9, France; E-mail: slecoz@math.univ-toulouse.fr

of the nonlinearity, the equation is focusing. At the balance between dispersion and focusing, we find "truly" nonlinear solutions: the stationary/standing/solitary waves.

A *standing wave* with frequency $\omega \in \mathbb{R}$ is a solution of (NLKG) of the form $u(t, x) = e^{i\omega t}\varphi_\omega(x)$. Such solution has the particularity to exist globally and to remain localized at any time. In the physics literature this kind of solutions are sometimes referred to as *Q-balls*. A *soliton* (or *solitary wave*) with speed $v \in \mathbb{R}^d$, frequency $\omega \in \mathbb{R}$ and initial phase and position $\theta \in \mathbb{R}$, $x_0 \in \mathbb{R}^d$ is a boosted standing wave solution of (NLKG). More precisely a soliton is a solution of (NLKG) of the form

$$e^{i\frac{\omega}{\gamma}t+i\theta}\varphi_{\omega,v}(x - vt - x_0),$$

where $\varphi_{\omega,v}$ is a profile depending on ω and v, and $\gamma := \frac{1}{\sqrt{1-|v|^2}}$ is the Lorentz factor.

We shall consider ground states solitons, i.e., boosted standing waves with ground states profiles (profiles minimizing a certain action functional, see Section 2 for a more precise definition). The orbital stability properties of such solitons have been widely studied. It started with the work of Shatah [47] where it was shown that there exists a critical frequency $\omega_c > 0$ such that if $p < 1 + 4/d$ and $\omega_c < |\omega| < m$ then standing waves are stable under radial perturbation. Later on, Shatah [48] proved that the stationary solution (i.e., the standing wave with $\omega = 0$) is strongly unstable and the picture for standing waves was completed by Shatah and Strauss [49] when they showed that if either $p \geq 1 + 4/d$ or if $|\omega| < \omega_c$ then standing waves are unstable. These results were generalized and consolidated by Grillakis, Shatah and Strauss in their celebrated works [14, 15]. The stability theory of solitons was revisited by Stuart [51] via the modulational approach introduced by Weinstein [52] for nonlinear Schrödinger equations. Compare to prior results, Stuart [51] provided two improvements: first, he treated the whole range of possible speeds $|v| < 1$ without the radiality assumptions, second he gave the laws of the modulations parameters. In particular, it was shown in [51] that the ground state solitons are stable if the parameters are within the following open set

$$\mathcal{O}_{\text{stab}} := \left\{ (\omega, \theta, v, x) \in \mathbb{R}^{2+2d};\ |\omega| < \sqrt{m},\ |v| < 1,\ \frac{1}{1 + \frac{4}{p-1} - d} < \frac{\omega^2}{m} \right\}. \quad (1)$$

Note that $\mathcal{O}_{\text{stab}}$ is nonempty only if $p < 1 + \frac{4}{d}$, i.e., the nonlinearity is L^2-subcritical. Instability was further investigated by Liu, Ohta and Todorova [26, 42, 43] (see also [16] for a companion result), who proved that when standing waves are unstable, then the instability is either strong (i.e., by blow up in possibly infinite time) when $p < 1 + 4/d$ or very strong (i.e., by blow up in finite time) when $p \geq 1 + 4/d$.

Recently, further informations on the dynamics of (NLKG) around solitons have been obtained by Nakanishi and Schlag. In [41], using a method referred to as Hadamard approach in dynamical systems, they show the existence of a center-stable manifold which contains all solutions of (NLKG) staying close to the solitons manifold and describe precisely this manifold. Furthermore, in [38], they adopt a Lyapunov-Perron approach for the study of the dynamics around ground state stationary solitons of (NLKG) for the 3-d cubic case and in a radial setting.

In particular, they show the existence of a center stable manifold such that the following trichotomy occurs for a solution with initial data close to the ground state stationary solution. On one side of the center stable manifold, the solution scatters to 0, on the other side it blows up in finite time and on the center stable manifold itself the solution scatters to the ground state. The same authors [40] extended later their results in the non-radial setting. One can also refer to their monograph [39] for a complete introduction to the mathematical study of equations similar to (NLKG), in particular the study of the dynamics of the equation around stationary/standing waves.

In this paper we address the question whether it is possible to construct a multi-soliton solution for (NLKG), i.e., a solution behaving at large time like a sum of solitons. Multi-solitons are long time known to exist for integrable equations such as the Korteweg-de Vries equation or the 1-d cubic nonlinear Schrödinger equation. Indeed, existence of multi-solitons follows from the inverse scattering transform, see e.g., the survey of Miura [36] for the Korteweg-de Vries equation and the work of Zakharov and Shabat [55] for the 1-d cubic nonlinear Schrödinger equation. In the recent years, there has been a series of works around the existence and dynamical properties of multi-solitons for various dispersive equations.

One of the first existence result of multi-solitons for non-integrable equations was obtained by Merle [35] as a by-product of the construction of a multiple blow-up points solution to the L^2-critical nonlinear Schrödinger equation (indeed a pseudo-conformal transform of this solution gives the multi-soliton). Later on, Perelman [44, 45] (see also [46]) studied asymptotic stability of a sum of solitons of nonlinear Schrödinger equation under spectral hypotheses and in weighted spaces. In the energy space, Martel, Merle and Tsai [29, 34] showed the existence and orbital stability of multi-solitons made of stable solitons. The existence of multi-solitons made of unstable solitons was obtained by Côte, Martel and Merle [8] for ground state and by Côte and Le Coz [7] for excited states under a high speed assumption. Further results on the existence of exotic solutions like a train of infinitely many solitons were obtained by Le Coz, Li and Tsai [22, 23].

For the non-integrable generalized Korteweg-de Vries equation, Martel [27] showed the existence and uniqueness of multi-solitons for L^2-subcritical nonlinearities. These multi-solitons were shown to be stable and asymptotically stable by Martel et al. [33]. Combet [6] investigated further the existence of multi-solitons in the supercritical case and showed the existence and uniqueness of a N-parameter family of multi-solitons. Outstanding results on the description of the interaction between two solitons were recently obtained by Martel and Merle [30–32].

Despite the many works on multi-solitons previously cited, to our knowledge the present paper and the recent preprint [9] are the first works dealing with existence of multi-soliton type solutions for nonlinear Klein-Gordon equations (see nevertheless [10] for related results on the nonlinear wave equation).

Our goal is to prove the following existence result for multi-solitons.

Theorem 1. *Assume that $1 < p < 1 + \frac{4}{d}$. For any $N \in \mathbb{N}$, take $(\omega_j, \theta_j, v_j, x_j)_{j=1,\dots,N} \subset \mathcal{O}_{\mathrm{stab}}$ and let (φ_j) be the associated ground state profiles $\varphi_j := \varphi_{\omega_j, v_j}$, and (γ_j) the Lorentz factors $\gamma_j := (1 - |v_j|^2)^{-\frac{1}{2}}$. Denote the corresponding solitons by*

$$\mathcal{R}_j(t, x) := e^{i\frac{\omega_j}{\gamma_j}t + i\theta_j}\varphi_j(x - v_j t - x_j).$$

Define

$$v_\star := \min\{|v_j - v_k|, j, k = 1, \ldots, N, j \neq k\} \qquad (\textit{minimal relative speed}) \qquad (2)$$

$$\omega_\star := \max\{|\omega_j|, j = 1, \ldots, N\} \qquad (\textit{maximal frequency}). \qquad (3)$$

There exists $\alpha = \alpha(d, N) > 0$, *such that if* $v_j \neq v_k$ *for any* $j \neq k$, *then there exist* $T_0 \in \mathbb{R}$ *and a solution* u *to* (NLKG) *existing on* $[T_0, +\infty)$ *and such that for all* $t \in [T_0, +\infty)$ *the following estimate holds*

$$\left\| u(t) - \sum_{j=1}^{N} \mathcal{R}_j(t) \right\|_{H^1} + \left\| \partial_t u(t) - \sum_{j=1}^{N} \partial_t \mathcal{R}_j(t) \right\|_{L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}.$$

Remark 1. During the preparation of this paper we have been aware of the work [9] by Côte and Muñoz. Our two results are companions in the following sense. In [9] the authors used spectral theory and a topological argument to prove the existence of multi-solitons made of unstable solitons. To the contrary, we use finite speed of propagation, classical modulation theory and energy estimates to obtain the existence of multi-solitary waves based on stable solitons. Merging our results together would give the existence of multi-solitons made with any kind (stable or unstable) of solitons.

Our strategy for the proof of Theorem 1 is to solve (NLKG) backwards around suitable approximate solutions. It is inspired by the works of Martel, Merle and Tsai on multi-solitons of Schrödinger equations [29, 34] (see also [7, 8] where similar strategies were enforced). The main new ingredients on which we rely are the variational characterizations of the profile and a coercivity property of the total linearized action.

We start by introducing the mathematical framework in which we are going to work in Section 2. After transforming (NLKG) into its Hamiltonian form (4), we list the tools which are going to be useful for our purposes: Cauchy Theory in $H^1 \times L^2$ and $H^s \times H^{s-1}$, Conservation laws, Finite Propagation Speed, standing waves, Lorentz transform and finally definitions of solitons and their profiles.

Then we go on with the core of the proof of Theorem 1. We consider a sequence of times $T^n \to \infty$, a set of final data $u_n = \sum \mathcal{R}_j(T^n)$ and the associated solutions (u_n) of (NLKG) backward in time. The sequence (u_n) provides us with a sequence of approximate multi-solitons, and we need to prove its convergence to a solution of (NLKG) satisfying to the conclusion of Theorem 1. For this purpose, we show that each u_n exists backwards in time up to some time T_0 independent of n and decay uniformly in n to the sum of solitons (Proposition 1). Eventually a compactness argument (Lemma 2) permits to show that (u_n) converges to a multi-soliton of (NLKG) on $[T_0, \infty)$. Most steps are performed in Section 3, apart from uniform estimates whose proof needs more preparation.

The proof of the uniform estimates relies on several ingredients: coercivity of the Hessian of the action around each component of the multi-soliton, modulation theory and slow variation of localized conservation laws, energy, charge, momenta.

We study the profiles of the solitons in Section 4. We characterize the profiles variationally using the conserved quantities of (NLKG) (Proposition 3), and show that the ground state profiles are at the mountain pass level, the least energy level and the Nehari level. Our proofs are self-contained and do not rely on the (NLS) case.

After obtaining the variational characterizations, we prove a coercivity property (Lemma 8) for the second variation of the action functional around a soliton (linearized action). To this aim, we study the spectrum of the linearized action and prove in particular the non-degeneracy of the kernel (i.e., the kernel contains only the eigenvectors generated by the invariances of the equation, see Lemma 7). It is usually a crucial point in these matters. We underline that the coercivity properties are related to the fact that our standing waves are stable.

Coercivity of the linearized action is obtained provided orthogonality conditions hold. This prompt the question of obtaining orthogonality conditions around a sum of solitons, which is resolved in Section 5 via modulation theory. The modulation result is twofold. First, it shows that close to a sum of solitons one can recover orthogonality conditions (see (35)) by adjusting the modulation parameters phases, translation and scaling. Second, it gives the dynamical laws followed by the parameters (see (36)).

Finally, we define cutoff functions around each soliton and use them to localize the action around each soliton. We use these localized actions to build a global action adapted to the sum of solitons. Several properties are transported from the local actions to the global one. In particular, the global action inherits from the coercivity (see Lemma 12). Due to errors generated by the cutoff it is not a conserved quantity, but we can however prove that it is almost conserved (i.e., it varies slowly). We use these properties combined with the modulation result to prove the uniform estimates.

This work is organized as follows. In Section 2, we set the mathematical work context. Section 3 contains the proof of the main result, assuming uniform estimates. In Section 4, we establish variational characterizations of the profiles and use them to prove a coercivity statement for the hessian of the action functional related to a soliton. In Section 5, we explain the modulation theory in the neighborhood of a sum of solitons. Finally, we put all pieces together in Section 6 to prove the uniform estimates. The Appendix contains the proof of a compact injection used in Section 3 and interactions estimates used in Section 6.

2. Mathematical Context

In this section we introduce rigorously all the necessary material for our study and restate our result in the Hamiltonian formulation for (NLKG), which is a more suitable formulation for our needs. But before let us precise some notations. We denote by $L^q(\mathbb{R}^d)$ the standard Lebesgue space and its norm by $\|\cdot\|_q$. The space $L^2(\mathbb{R}^d)$ is viewed as a real Hilbert space endowed with the scalar product

$$(u, v)_2 = \mathrm{Re} \int_{\mathbb{R}^d} u\bar{v}dx.$$

The Sobolev spaces $H^s(\mathbb{R}^d)$ are endowed with their usual norms $\|\cdot\|_{H^s}$. For the product space $L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ we use the norm

$$\left\| \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\|_{L^2 \times L^2} = \sqrt{\|u_1\|_2^2 + \|u_2\|_2^2},$$

with similar convention for $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$, $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$, etc. We shall sometimes us the following notational shortcut:

$$(u, \nabla v)_2 := \begin{pmatrix} \left(u, \frac{\partial v}{\partial x_1}\right)_2 \\ \vdots \\ \left(u, \frac{\partial v}{\partial x_d}\right)_2 \end{pmatrix}.$$

Finally, unless otherwise specified the components of a vector $W \in L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ will be denoted by W_1 and W_2.

The *Hamiltonian Formalism* for the nonlinear Klein-Gordon equation (NLKG) is formulated as follows. For $(u_1, u_2) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ we define the following Hamiltonian (which we will call *energy* in the sequel)

$$E\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} := \frac{1}{2}\|u_2\|_2^2 + \frac{1}{2}\|\nabla u_1\|_2^2 + \frac{m}{2}\|u_1\|_2^2 - \frac{1}{p+1}\|u_1\|_{p+1}^{p+1}.$$

Define the matrix $J := \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Then u is a solution of (NLKG) if and only if $(u_1, u_2) := (u, u_t)$ solves the following equation

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = JE'\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{4}$$

From now on we shall work only with the Hamiltonian equation (4).

Due to this Hamiltonian formulation the energy is (at least formally) conserved. In addition, the invariance of (4) under phase shifts and space translations generates two other *conservations laws*, the *charge* Q and the (vectorial) *momentum* P, defined in the following way:

$$Q\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \operatorname{Im}\int_{\mathbb{R}^d} u_1 \bar{u}_2 dx, \qquad P\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \operatorname{Re}\int_{\mathbb{R}^d} \nabla u_1 \bar{u}_2 dx.$$

With our restrictions on the growth of the nonlinearity in Theorem 1 (L^2-subcritical), it is well-known that the Cauchy problem for (4) is globally well-posedness in the energy space $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. More precisely, the following well-posedness theory holds.

Cauchy Theory in $H^1 \times L^2$. *Assume* $1 < p < 1 + \frac{4}{d}$. *For any initial data* $U_0 = (u_1^0, u_2^0) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ *there exists a unique maximal solution of* (4)

$$U \in \mathscr{C}(\mathbb{R}, H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)) \cap \mathscr{C}^1(\mathbb{R}, L^2(\mathbb{R}^d) \times H^{-1}(\mathbb{R}^d)).$$

Furthermore, we have the following properties.

Conservation of energy, charge and momentum: *for all* $t \in \mathbb{R}$, *we have*

$$E(U(t)) = E(U_0), \quad Q(U(t)) = Q(U_0), \quad P(U(t)) = P(U_0),$$

Global estimate: *there exist* $C_0 > 0$ *such that* $\|U\|_{\mathscr{C}(\mathbb{R}, H^1 \times L^2)} \leq C_0 \|U_0\|_{H^1 \times L^2}$.

Uniqueness in light cones: *If $\widetilde{U}$ is another solution to* (4) *on* $(0, T)$ *for* $T > 0$ *with* $\widetilde{U}(0) = U_0$ *in* $\{x : |x - x_0| < T\}$ *for some* $x_0 \in \mathbb{R}^d$, *then* $\widetilde{U} \equiv U$ *on the backward light cone* $\{(t, x) \in [0, T) \times \mathbb{R}^d : |x - x_0| < T - t\}$.

Continuous dependency upon the initial data: *if* $(U_0^n) \subset H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ *converges to* U_0 *in* $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$, *then the associated solutions* (U_n) *of* (4) *converge in* $\mathscr{C}\left(I, H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)\right)$ *for any compact time interval* $I \subset \mathbb{R}$ *to the solution* U *of* (4) *with initial data* $U(0) = U_0$.

For this set of results, we refer to the classical papers by Ginibre and Velo [12, 13], or the recent review in the paper [19] by Killip et al. For our purposes, we will also need a more refined result on local well-posedness in the slightly larger space $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$ for some $s < 1$ (see Lindblad and Sogge [24] or Nakamura and Ozawa [37]).

Cauchy Theory in $H^s \times H^{s-1}$. *Let* $s > 0$ *be such that either* $s > d/2$ *or* $1/2 \leq s < d/2$ *and* $p < 1 + \frac{4}{d-2s}$. *For any initial data* $U_0 = (u_1^0, u_2^0) \in H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$ *there exists a unique maximal solution of* (4)

$$U \in \mathscr{C}((-T_\star, T^\star), H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)).$$

Furthermore, we have the continuous dependent upon the initial data: *if* $(U_0^n) \subset H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$ *converges to* U_0 *in* $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$, *then the associated solutions* (U_n) *of* (4) *converge to* U *in* $\mathscr{C}(I, H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d))$ *for any compact time interval* $I \subset (-T_\star, T^\star)$, *where* U *is the solution to* (4) *with initial data* $U(0) = U_0$.

A useful consequence of the uniqueness in light cones (Cauchy Theory in $H^1 \times L^2$) is the following *finite speed of propagation* property.

Finite Propagation Speed. *Let* $U = (u_1, u_2) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ *be a solution of* (4) *on* $(-\infty, T^\star]$. *There exists* C_0, *depending only on* $\|U(T^\star)\|_{H^1 \times L^2}$ *such that if there exist* $0 < \varepsilon$ *and* $M > 0$ *satisfying*

$$\int_{|x|>M} |\nabla u_1(T^\star)|^2 + |u_1(T^\star)|^2 + |u_2(T^\star)|^2 dx \leq \varepsilon,$$

then for any $t \in [-\infty, T^\star]$ *we have*

$$\int_{|x|>2M+(T^\star-t)} |\nabla u_1(t)|^2 + |u_1(t)|^2 + |u_2(t)|^2 dx \leq C_0 \varepsilon,$$

Proof. Let χ_M be a cutoff function such that

$$\chi_M(x) = \begin{cases} 1 & \text{for } |x| > 2M \\ 0 & \text{for } |x| < M \end{cases} \quad \text{and} \quad \|\nabla \chi_M\|_\infty < \frac{C_0}{M}.$$

Define $U_{T^\star,M} := U(T^\star)\chi_M$ and denote by U_M the associated solution of (4). By assumption, we have $\|U_{T^\star,M}\|_{H^1 \times L^2}^2 \leq \varepsilon$ and by the Cauchy Theory in $H^1 \times L^2$ the solution U_M exists on $\mathbb{R}$ and verifies for all $t \in \mathbb{R}$

$$\|U_M(t)\|_{H^1 \times L^2}^2 \leq C_0 \varepsilon.$$

However, by uniqueness on light cones, U_M and U coincide on $\{(t,x) \in \mathbb{R} \times \mathbb{R}^d : |x| > 2M + (T^\star - t)\}$, and for any $t \in (-\infty, T^\star)$ this implies

$$\int_{|x|>2M+(T^\star-t)} |\nabla u_1(t)|^2 + |u_1(t)|^2 + |u_2(t)|^2 dx \leq \|U_M(t)\|^2_{H^1 \times L^2} \leq C_0 \varepsilon,$$

which was the desired conclusion. □

Lorentz transform. Among the symmetries of (4), we already mentioned the phase shift and translation. We consider now the Lorentzian symmetry, defined as follows. Take $U(t,x) = (u_1, u_2)(t,x)$ and $v \in \mathbb{R}^d$ with $|v|$ smaller than the speed of light for (4), namely $|v| < 1$. The *Lorentz transform* $\mathscr{L}_v U$ of U is the function of (t,x) defined by

$$(\mathscr{L}_v U)(t,x) := \begin{pmatrix} u_1(\tau, y) \\ \gamma(u_2(\tau, y) - v\nabla_y u_1(\tau, y)) \end{pmatrix}$$

where τ and y are defined by

$$\begin{aligned} \tau = \tau(t,x) &:= \gamma(t - v \cdot x) = \frac{1}{\gamma} t - \gamma(x - vt) \cdot v, \\ y = y(t,x) &:= x - x_v + \gamma(x_v - vt) = x - vt + (\gamma - 1)(x - vt)_v. \end{aligned}$$

Here, the Lorentz parameter γ is defined by

$$\gamma := \frac{1}{\sqrt{1 - |v|^2}},$$

and the subscript v denote the orthogonal projection onto the vectorial line generated by v, that is

$$x_v := \frac{x \cdot v}{|v|^2} v.$$

It is simple algebra to verify that (4) is Lorentz invariant, in the sense that if U is a solution of (4), then so is $\mathscr{L}_v U$. Also note that the Lorentz transform is invertible with inverse $\mathscr{L}_{-v}$.

Standing waves. Take $\omega \in \mathbb{R}$. In the Hamiltonian formulation, a *standing wave* with frequency ω is a solution of (4) of the form $U(t,x) = e^{i\omega t}\Phi_\omega(x)$. Plugging this ansatz for U into (4), it is easy to see that $\Phi_\omega = \binom{\varphi_{\omega,1}}{\varphi_{\omega,2}}$ must be a critical point of $E + \omega Q$, hence a solution to the stationary elliptic system

$$\begin{cases} -\Delta\varphi_{\omega,1} + m\varphi_{\omega,1} - |\varphi_{\omega,1}|^{p-1}\varphi_{\omega,1} + i\omega\varphi_{\omega,2} = 0, \\ \varphi_{\omega,2} - i\omega\varphi_{\omega,1} = 0. \end{cases}$$

The solutions of this system are clearly of the form $\binom{\varphi_\omega}{i\omega\varphi_\omega}$, where φ_ω satisfies the scalar equation

$$-\Delta\varphi_\omega + (m - \omega^2)\varphi_\omega - |\varphi_\omega|^{p-1}\varphi_\omega = 0. \tag{5}$$

Solutions to (5) and their properties are well-known (see [3, 4, 11, 20] and the references therein). For every $\omega \in (-\sqrt{m}, \sqrt{m})$ there exists a unique, positive, and radial function $\varphi_\omega \in \mathscr{C}^2(\mathbb{R}^d)$ solution of (5). In addition, the function φ_ω is exponentially decaying at infinity: for any $\mu < (m - \omega^2)$ there exists $C(\mu, \omega) > 0$ such that

$$|\varphi_\omega(x)| \leq C(\mu, \omega) e^{-\sqrt{\mu}|x|} \qquad \text{for all } x \in \mathbb{R}^d. \tag{6}$$

Furthermore, any φ_ω satisfy the scaling property

$$\varphi_\omega(x) = (m - \omega^2)^{\frac{1}{p-1}} \tilde{\varphi}\left((m - \omega^2)^{\frac{1}{2}} x\right), \tag{7}$$

where $\tilde{\varphi}$ is the unique positive radial solution to $-\Delta\tilde{\varphi} + \tilde{\varphi} - |\tilde{\varphi}|^{p-1}\tilde{\varphi} = 0$. The function φ_ω is called *ground state*. In dimension $d \geq 2$, there exist infinitely many other solutions to (5), called *excited states*. In the sequel, we shall deal only with ground states solutions to (5). Indeed, our analysis deeply relies on properties of the ground states which do not hold for other solutions, in particular the stability of the associated standing waves (see Section 4 for details).

Remark 2. It is interesting to notice that, although the presence of the nonlinear term permits the existence of states with negative energy, the standing waves have always positive energies. Indeed, a straightforward computation assures that for a standing wave $U(t, x) = e^{i\omega t}\Phi_\omega(x)$ the corresponding energy is given by

$$E(\Phi_\omega) = \left(\frac{1}{2} - \frac{1}{p+1}\right)\left(\|\nabla\varphi_\omega\|_2^2 + m\|\varphi_\omega\|_2^2\right) + \left(\frac{1}{2} + \frac{1}{p+1}\right)\omega^2\|\varphi_\omega\|_2^2$$

The fact that $p > 1$ guarantees that $E(\Phi_\omega) > 0$.

Remark 3. The scaling property (7) guarantees that the energy of the ground states varies continuously with respect to ω. This fact implies that the multi-soliton solutions for (NLKG) behave at large time like a sum of solitons that are allowed to have different energies. A straightforward computation indeed gives

$$\|\varphi_\omega\|_2^2 = (m - \omega^2)^{\frac{4-d(p-1)}{2(p-1)}} \|\tilde{\varphi}\|_2^2$$
$$\|\nabla\varphi_\omega\|_2^2 = (m - \omega^2)^{\frac{p(2-d)+2+d}{2(p-1)}} \|\nabla\tilde{\varphi}\|_2^2.$$

Now $\tilde{\varphi}$ is solution of $-\Delta\tilde{\varphi} + \tilde{\varphi} - |\tilde{\varphi}|^{p-1}\tilde{\varphi} = 0$ such that, by means of Pohozaev identity,

$$\|\nabla\tilde{\varphi}\|_2^2 = \frac{d(p-1)}{2d - (d-2)(p+1)}\|\tilde{\varphi}\|_2^2.$$

Merging all this information we get the relation between energy and ω given by

$$E(\varphi_\omega) = g(\omega)\|\tilde{\varphi}\|_2^2,$$

where

$$g(\omega) = \frac{p-1}{2(p+1)}\left((m-\omega^2)^{\frac{p(2-d)+2+d}{2(p-1)}} + m(m-\omega^2)^{\frac{4-d(p-1)}{2(p-1)}}\right) + \left(\frac{p+3}{2(p+1)}\omega^2(m-\omega^2)^{\frac{4-d(p-1)}{2(p-1)}}\right),$$

which can be rewritten as

$$g(\omega) = \frac{(m(p-1)+2\omega^2)(m-\omega^2)^{\frac{2}{p-1}-\frac{d}{2}}}{(p+1)}.$$

The monotonicity of $g(\omega)$ when ω belongs to $\mathcal{O}_{\text{stab}}$ follows easily.

Solitons. Starting from a standing wave, one generates a new family of solutions to (4) simply by boosting them using the Lorentz transform. These new solutions are the *solitary waves* (or simply *solitons*). Precisely, take a frequency $|\omega| \le \sqrt{m}$, the profile $\Phi_\omega := \binom{\varphi_\omega}{i\omega\varphi_\omega}$ (where φ_ω is the ground state of (5)), a phase $\theta \in \mathbb{R}$, and a speed and a position $v, x_0 \in \mathbb{R}^d$ with $|v| < 1$. The associated soliton is

$$e^{i\frac{\omega}{\gamma}t+i\theta}\Phi_{\omega,v}(x-vt-x_0),$$

where the new profile $\Phi_{\omega,v}$ is given by

$$\Phi_{\omega,v}(x) = e^{-i\gamma\omega v\cdot x}\begin{pmatrix}\varphi_\omega(x+(\gamma-1)x_v)\\ \gamma(i\omega\varphi_\omega(x+(\gamma-1)x_v) - v\nabla\varphi_\omega(x+(\gamma-1)x_v))\end{pmatrix}. \tag{8}$$

By direct computation, and provided we have noticed that

$$E'\begin{pmatrix}u_1\\u_2\end{pmatrix} = \begin{pmatrix}-\Delta u_1 + mu_1 - |u_1|^{p-1}u_1\\ u_2\end{pmatrix},$$

$$Q'\begin{pmatrix}u_1\\u_2\end{pmatrix} = \begin{pmatrix}iu_2\\-iu_1\end{pmatrix} = iJ\begin{pmatrix}u_1\\u_2\end{pmatrix}, \qquad P'\begin{pmatrix}u_1\\u_2\end{pmatrix} = \begin{pmatrix}-\nabla u_2\\ \nabla u_1\end{pmatrix} = -J\nabla\begin{pmatrix}u_1\\u_2\end{pmatrix},$$

it is not difficult to see that $\Phi_{\omega,v}$ is a critical point of

$$S := E + \frac{\omega}{\gamma}Q + v\cdot P.$$

With all these preliminaries out of the way, we can go on with the proof of Theorem 1.

3. Existence of Multi-Solitons

This section contains the core of the proof of Theorem 1 assuming uniform estimates (Proposition 1) which are proved in Section 6.

Assume that $p < 1 + \frac{4}{d}$. Take $N \in \mathbb{N}$, $(\omega_j, \theta_j, v_j, x_j)_{j=1,\dots,N} \subset \mathcal{O}_{\text{stab}}$, Φ_j the associated Hamiltonian profiles (as in (8)), $v_\star$ and $\omega_\star$ as in (2), (3), (γ_j) the Lorentz parameters, (R_j) the corresponding solitons

$$R_j(t,x) := e^{i\frac{\omega_j}{\gamma_j}t+i\theta_j}\Phi_j(x - v_j t - x_j),$$

and R the sum of the solitons:

$$R(t,x) := \sum_{j=1}^{N} R_j(t,x).$$

Reformulated using the Hamiltonian expression (4) of (NLKG), our goal is to prove that there exists $\alpha = \alpha(d, N) > 0$, such that if $v_j \neq v_k$ for any $j \neq k$, then there exist $T_0 \in \mathbb{R}$ and a solution U to (4) existing on $[T_0, +\infty)$ and such that the following estimate holds for all $t \in [T_0, +\infty)$

$$\|U(t) - R(t)\|_{H^1\times L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}.$$

We are going to define a sequence of approximate multi-solitons and prove its convergence to the desired solution of (4). Take an increasing sequence of time $T^n \to +\infty$ and for each n let U_n be the solution to (4) obtained by solving (4) backward in time from T^n with final data $U_n(T^n) = R(T^n)$. Our proof will rely on two main ingredients. First we have uniform estimates for the sequence of approximate multi-solitons.

Proposition 1 (Uniform Estimates). *There exist* $\alpha = \alpha(d, N) > 0$, *and* $T_0 \in \mathbb{R}$ (*independent of* n) *such that for* n *large enough the solution* U_n *of* (4) *with* $U_n(T^n) = R(T^n)$ *exists on* $[T_0, T^n]$ *and satisfies for all* $t \in [T_0, T^n]$ *the estimate*

$$\|U_n(t) - R(t)\|_{H^1\times L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}. \tag{9}$$

Proposition 1 establishes that the approximate multi-solitons U_n all satisfy the desired estimate on time intervals of the form $[T_0, T^n]$, with T_0 independent of n. The proof of Proposition 1 is rather involved and we postpone it to Section 6 (useful informations for this proof are derived in Sections 4 and 5).

The second ingredient of the proof of Theorem 1 is an $H^1 \times L^2$−compactness property of the sequence of initial data of the approximate multi-solitons.

Lemma 2 (Compactness). *Let* T_0 *be given by Proposition* 1. *For any* $\varepsilon > 0$ *there exists* M_ε *such that for any* n *large enough* U_n *verifies*

$$\int_{|x|>M_\varepsilon} |\nabla U_{n,1}(T_0)|^2 + |U_{n,1}(T_0)|^2 + |U_{n,2}(T_0)|^2 dx \leq \varepsilon.$$

The argument for the proof of Lemma 2 is different from the Schrödinger equation case. Indeed, we benefit with the Klein-Gordon equation of the Finite Propagation Speed, which is not the case for Schrödinger equations where one has to us virial identities (see e.g., [29, Lemma 2]).

Proof of Lemma 2. The result is a consequence of the Finite Speed of Propagation and the uniform estimates of Proposition 1. Indeed, take $\varepsilon > 0$ and let $T^\star$ be such that $e^{-\alpha\sqrt{m-\omega_\star^2}v_\star T^\star} < \frac{\varepsilon}{2}$. Then it follows from Proposition 1 that for n large enough

$$\|U_n(T^\star) - R(T^\star)\|_{H^1\times L^2} \leq \frac{\varepsilon}{2}. \tag{10}$$

By exponential decay of the sum of solitons, there exists $\widetilde{M}_\varepsilon$ such that

$$\int_{|x|>\widetilde{M}_\varepsilon} |\nabla(R(T^\star)_1)|^2 + |(R(T^\star))_1|^2 + |(R(T^\star))_2|^2 dx \leq \frac{\varepsilon}{2}. \tag{11}$$

Combining (10) and (11), we get

$$\int_{|x|>\widetilde{M}_\varepsilon} |\nabla U_{n,1}(T^\star)|^2 + |U_{n,1}(T^\star)|^2 + |U_{n,2}(T^\star)|^2 dx \leq \varepsilon.$$

By Finite Speed of Propagation, this implies

$$\int_{|x|>\tilde{2}M_\varepsilon+(T^\star-T_0)} |\nabla U_{n,1}(T_0)|^2 + |U_{n,1}(T_0)|^2 + |U_{n,2}(T_0)|^2 dx \leq \varepsilon.$$

Setting $M_\varepsilon = 2\widetilde{M}_\varepsilon + (T^\star - T_0)$ finishes the proof. □

Proof of Theorem 1. With in hand our sequence of approximate multi-solitons satisfying the desired estimate, the only thing left to do is to prove that it actually converges to a solution of (4) satisfying the same estimate (9).

First of all, we show the convergence of initial data. Since U_n satisfies (9), the sequence $U_n(T_0)$ is bounded in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. Therefore there exists $U_0 \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ such that $U_n \rightharpoonup U_0$ weakly in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. We are going to prove that the previous convergence is strong in $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$ for any $0 < s < 1$. Take $\varepsilon > 0$. Using Lemma 2, we infer the existence of $M_\varepsilon > 0$ such that for n large enough

$$\int_{|x|>M_\varepsilon} |\nabla U_{n,1}(T_0)|^2 + |U_{n,1}(T_0)|^2 + |U_{n,2}(T_0)|^2 dx + \int_{|x|>M_\varepsilon} |\nabla U_{0,1}(T_0)|^2 + |U_{0,1}(T_0)|^2 + |U_{0,2}(T_0)|^2 dx \leq \frac{\varepsilon}{2}. \tag{12}$$

Define $\chi_\varepsilon : \mathbb{R}^d \to [0, 1]$ a cutoff function such that $\chi_\varepsilon(x) = 1$ if $|x| < M_\varepsilon$, $\chi_\varepsilon(x) = 0$ if $|x| > 2M_\varepsilon$, and $\|\nabla\chi_\varepsilon\|_\infty \leq 1$. We have

$$\|U_n(T_0) - U_0\|_{H^s\times H^{s-1}} \leq \|(U_n(T_0) - U_0)\chi_\varepsilon\|_{H^s\times H^{s-1}} + \|(U_n(T_0) - U_0)(1-\chi_\varepsilon)\|_{H^s\times H^{s-1}}$$

From the compactness (a proof of this fact is included in the Appendix, Lemma 17 for the reader's convenience) of the injection $H^s(\Omega) \hookrightarrow H^{s-\delta}(\Omega)$ when Ω is bounded and $\delta > 0$, we infer that, for n large enough and maybe up to a subsequence, we have

$$\|(U_n(T_0) - U_0)\chi_\varepsilon\|_{H^s\times H^{s-1}} \leq \frac{\varepsilon}{2}.$$

Moreover, by (12)

$$\|(U_n(T_0) - U_0)(1-\chi_\varepsilon)\|_{H^s\times H^{s-1}} \leq \|(U_n(T_0) - U_0)(1-\chi_\varepsilon)\|_{H^1\times L^2} \leq \frac{\varepsilon}{2}.$$

Combining the last three equations gives us

$$\|U_n(T_0) - U_0\|_{H^s\times H^{s-1}} \leq \varepsilon.$$

Hence $U_n(T_0)$ converges strongly to U_0 in $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$.

Let us now show that the solution U of (4) in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ with data $U(T_0) = U_0$ satisfies the required estimate. By Local Cauchy Theory of (4) in $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$, we have the strong convergence

$$U_n(t) \to U(t) \quad \text{in } H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$$

for any $t \in [T_0, +\infty)$. In addition, by uniqueness of the limit and since $U_n(t)$ is bounded in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ (by (9)), we have the weak convergence for $t \in [T_0, +\infty)$

$$U_n(t) \rightharpoonup U(t) \quad \text{in } H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d).$$

Therefore, by weak lower semi-continuity of the $H^1 \times L^2$-norm and (9), we have

$$\|U(t) - R(t)\|_{H^1\times L^2} \leq \liminf_{n\to+\infty} \|U_n(t) - R(t)\|_{H^1\times L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t},$$

which concludes the proof of Theorem 1. □

4. Properties of the Profiles

Since we will be working mainly within the Hamiltonian formulation of (NLKG), it will be convenient to characterize the soliton profiles using the conserved quantities. We already mentioned that the profile $\Phi_{\omega,v}$ is a critical point of the functional *action*

$$S := E + \frac{\omega}{\gamma} Q + v \cdot P,$$

or more explicitly a solution to

$$\begin{cases} -\Delta w_1 + m w_1 - |w_1|^{p-1} w_1 + i\frac{\omega}{\gamma} w_2 - v \cdot \nabla w_2 = 0, \\ w_2 - i\frac{\omega}{\gamma} w_1 + v \cdot \nabla w_1 = 0. \end{cases} \tag{13}$$

In this section, we are going to give some variational characterizations of $\Phi_{\omega,v}$ and study the Hessian $S''(\Phi_{\omega,v})$.

As far as we know, the variational characterizations given in the following Proposition 3 were never derived before, although they are expected in view of what happens in the scalar setting. The ideas on the relationships between different variational characterizations used further in this section were introduced by Jeanjean and Tanaka in [17, 18] (see also [2] for related results). We believe that these variational characterizations of the profile $\Phi_{\omega,v}$ are of independent interest.

For the purpose of constructing multi-solitons, the main result of this section is the coercivity property given in Lemma 8. The proof of this result relies on the variational characterization of the profiles as well as on their non-degeneracy, which is given by Lemma 7. We shall follow closely the presentation made in [21] for the standing waves of NLS.

4.1. *Variational Characterizations*

We define the *mountain pass level* by

$$MP := \inf_{\eta\in\Gamma}\sup_{s\in[0,1]} S(\eta(s))$$

where Γ is the set of admissible paths

$$\Gamma := \{\eta \in \mathscr{C}\left((0,1), H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)\right); \eta(0)=0, S(\eta(1))<0\}.$$

We define the Nehari constraint for $W \in H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$ by

$$I(W) := \langle S'(W), W\rangle$$

and the Nehari level by

$$NL := \min\{S(W);\ I(W)=0, W\neq 0\}.$$

We also define the *least energy level* by

$$LE := \min\{S(W);\ W\in H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d),\ W\not\equiv 0,\ S'(W)=0\}.$$

Proposition 3. *The profile $\Phi_{\omega,v}$ admits the following variational characterizations*:

$$S(\Phi_{\omega,v}) = MP = NL = LE.$$

Let us start by proving using mountain pass arguments that S admits a critical point. Then we will show that this critical point is at the mountain pass level and also at the least energy level and at the Nehari level and we will identify it with $\Phi_{\omega,v}$.

Lemma 4. *There exists $\Psi \in H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$ a non-trivial critical point of S, i.e.,*

$$\Psi\neq 0,\quad S'(\Psi)=0.$$

Before going further, we make the following useful observations on the formulation of S: for $W=(w_1,w_2)\in H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$ it is simple algebra to see that

$$\begin{aligned}S(W) &= \frac{1}{2}\|\nabla w_1\|_2^2 - \frac{1}{2}\|v\cdot\nabla w_1\|_2^2 + \frac{1}{2}\left(m-\frac{\omega^2}{\gamma^2}\right)\|w_1\|_2^2 - \frac{1}{2}\frac{\omega}{\gamma}v\cdot \mathrm{Im}\int_{\mathbb{R}^d} w_1\nabla\bar{w}_1\\ &\quad + \frac{1}{2}\left\|v\cdot\nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 - \frac{1}{p+1}\|w_1\|_{p+1}^{p+1}.\end{aligned}$$

We can remark further that if $\tilde{w}_1$ is such that $w_1(x) = e^{-i\omega\gamma v\cdot x}\tilde{w}_1(x + (\gamma - 1)x_v)$ then we have

$$S(W) = \frac{1}{\gamma}\left(\frac{1}{2}\|\nabla\tilde{w}_1\|_2^2 + \frac{1}{2}(m - \omega^2)\|\tilde{w}_1\|_2^2\right) + \frac{1}{2}\left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 - \frac{1}{p+1}\|w_1\|_{p+1}^{p+1}. \tag{14}$$

We shall also use the following Lemma at several occasions.

Lemma 5. *For any* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that for any* $W = (w_1, w_2) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ *we have*

$$\varepsilon\|w_1\|_{H^1}^2 + \|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\|_2^2 \geq \delta\|W\|_{H^1\times L^2}^2.$$

Proof. We only have to make $\|w_2\|_2^2$ appear. Write

$$w_2 = \alpha\left(v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1\right) + w_2^\perp,$$

where $\alpha \in \mathbb{R}$ and $\left(v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1, w_2^\perp\right)_2 = 0$. We have

$$\left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 = (1+\alpha)^2\left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1\right\|_2^2 + \|w_2^\perp\|_2^2 \tag{15}$$

$$\|w_2\|_2^2 = \alpha^2\left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1\right\|_2^2 + \|w_2^\perp\|_2^2.$$

There is a possible degeneracy in (15) if $\alpha = -1$, but we can compensate it by using a piece of $\|w_1\|_{H^1}^2$:

$$\begin{aligned}
&\frac{\varepsilon}{2}\|w_1\|_{H^1}^2 + \left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 \\
&\geq C\left((1+\alpha)^2 + \frac{\varepsilon}{2}\right)\left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1\right\|_2^2 + C\|w_2^\perp\|_2^2 \\
&\geq \widetilde{C}\|w_2\|_2^2.
\end{aligned} \tag{16}$$

The desired inequality is then a direct consequence of (16). □

Proof of Lemma 4. Step 1: Mountain-Pass geometry.

We claim that the functional S has a mountain-pass geometry, i.e.,

$$MP > 0.$$

We start by showing that Γ is not empty. Indeed, take $W = (w_1, w_2) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ and $s > 0$. Then using (14) we see that

$$S(sW) = \frac{s^2}{2}\left(\frac{1}{\gamma}\left(\frac{1}{2}\|\nabla\tilde{w}_1\|_2^2 + \frac{1}{2}(m-\omega^2)\|\tilde{w}_1\|_2^2\right) + \frac{1}{2}\left\|v\cdot\nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2\right) - \frac{s^{p+1}}{p+1}\|w_1\|_{p+1}^{p+1}.$$

Therefore if s is large enough we have $S(sW) < 0$, hence the path $s \mapsto S(\frac{s}{C}W)$ belongs to Γ provided C has been chosen large enough.

To show that $MP > 0$, it is enough to prove that there exists a function $f : \mathbb{R}^+ \to \mathbb{R}$ such that $f(s) > 0$ for s close to 0 and $S(W) \geq f(\|W\|_{H^1\times L^2})$. Using (14), the continuity of $w_1 \to \tilde{w}_1$ in $H^1(\mathbb{R}^d)$ and Sobolev embeddings, it is easy to see that there exists $\varepsilon > 0$ such that

$$S(W) \geq \frac{\varepsilon}{2}\|w_1\|_{H^1}^2 + \frac{1}{2}\left\|v\cdot\nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 - C\|w_1\|_{H^1}^{p+1}.$$

From Lemma 5 we infer that there exists $\tilde{\delta} > 0$ such that

$$S(W) \geq \tilde{\delta}\|W\|_{H^1\times L^2}^2 - C\|W\|_{H^1\times L^2}^{p+1}.$$

This implies that $S(W) > 0$ if $\|W\|_{H^1\times L^2}$ is small and there exist $C > 0$, $\delta > 0$ such that $S(W) > C > 0$ for $\|W\|_{H^1\times L^2} = \delta$. This implies $MP > 0$ and S has a mountain pass geometry.

Step 2: Existence of a Palais-Smale sequence.

From Ekeland variational principle (see e.g., [54]) and Step 1, we infer the existence of a Palais-Smale sequence $W_n = (w_{1,n}, w_{2,n})$ at the level MP, i.e.,

$$S(W_n) \to MP, \qquad S'(W_n) \to 0, \text{ as } n \to +\infty. \tag{17}$$

Step 3: Non-vanishing of the Palais-Smale sequence.

Assume by contradiction that the sequence W_n is vanishing, more precisely for any $R > 0$ we have

$$\lim_{n\to+\infty}\sup_{y\in\mathbb{R}^d}\int_{|x-y|<R}(|w_{1,n}|^2 + |w_{2,n}|^2)dx = 0.$$

Take $\varepsilon > 0$ and $R > 0$ and let n be large enough so that

$$\sup_{y\in\mathbb{R}^d}\int_{|x-y|<R}(|w_{1,n}|^2 + |w_{2,n}|^2)dx < \varepsilon.$$

Recall Lions' Lemma (see [25]): for any $w \in H^1(\mathbb{R}^d)$ we have

$$\|w\|_{p+1}^{p+1} \leq C\left(\sup_{y\in\mathbb{R}^d}\int_{|x-y|<R}|w|^2dx\right)^{p-1}\|w\|_{H^1}^2. \tag{18}$$

Therefore, for n large enough, and using (14) and Lemma 5, we get

$$S(W_n) \geq C\|W_n\|^2_{H^1\times L^2} - \varepsilon\|w_{1,n}\|^2_{H^1},$$

which implies (if ε has been chosen small enough) that (W_n) is bounded in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. This boundedness has two consequences: as $n \to +\infty$, we have

$$\langle S'(W_n), W_n\rangle \to 0 \quad \text{and} \quad \|w_{1,n}\|_{p+1} \to 0. \tag{19}$$

where the first limit is due to the fact that (W_n) is a Palais-Smale sequence (see (17)) and the second limit comes from (18). However, we have

$$\left(\frac{1}{p+1} - \frac{1}{2}\right)\|w_{1,n}\|^{p+1}_{p+1} = S(W_n) - \frac{1}{2}\langle S'(W_n), W_n\rangle$$

and therefore (19) implies

$$\lim_{n\to+\infty} S(W_n) = 0,$$

which enters in contradiction with $\lim_{n\to+\infty} S(W_n) = MP > 0$. Therefore the sequence (W_n) is non-vanishing.

Step 4: Convergence to a critical point

Since W_n is non-vanishing, there exists $R, \delta > 0$ and $(y_n) \subset \mathbb{R}^d$ such that for n large enough

$$\int_{|x-y_n|<R} |w_{1,n}|^2 dx > \delta. \tag{20}$$

If we substitute $W_n(\cdot - y_n)$ to W_n (keeping the same notation), the sequence (W_n) is still a Palais-Smale sequence and keeps the same properties. In particular, as known from Step 3, (W_n) is bounded in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$, and therefore we have the existence of $\Psi \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ such that $W_n \rightharpoonup \Psi$ weakly in $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. Since W_n is a Palais-Smale sequence and S' is continuous, we have $S'(\Psi) = 0$. Hence we only have to show that Ψ is non-trivial. This is a direct consequence of (20) and the compact injection $H^1(|x| < R) \hookrightarrow L^2(|x| < R)$. Hence Ψ is a non-trivial critical point of S and the proof of Lemma 4 is finished. □

We turn now to the variational characterizations of the critical point obtained in Lemma 4.

Lemma 6. *Take Ψ the critical point of S found in Lemma* 4. *The following equality is satisfied.*

$$S(\Psi) = MP = NL = LE.$$

Proof. Let us start by showing

$$S(\Psi) \leq MP. \tag{21}$$

Using $S'(\Psi) = 0$, we have

$$S(\Psi) = S(\Psi) - \frac{1}{p+1}\langle S'(\Psi), \Psi\rangle$$
$$= \left(\frac{1}{2} - \frac{1}{p+1}\right)\left(\frac{1}{\gamma}\left(\|\nabla\tilde{w}_1\|_2^2 + (m-\omega^2)\|\tilde{w}_1\|_2^2\right) + \left\| v\cdot\nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2\right) \leq \cdots \tag{22}$$

Recall that $\tilde{w}_1$ is such that $w_1(x) = e^{-i\omega\gamma v\cdot x}\tilde{w}_1(x + (\gamma-1)x_v)$ (see (14)). Using the weak lower semi-continuity of the norm, we can continue the inequality started in (22) by

$$\cdots \leq \left(\frac{1}{2} - \frac{1}{p+1}\right)\liminf_{n\to+\infty}\left(\frac{1}{\gamma}(\|\nabla\tilde{w}_{1,n}\|_2^2 + (m-\omega^2)\|\tilde{w}_{1,n}\|_2^2) + \left\| v\cdot\nabla w_{1,n} - i\frac{\omega}{\gamma}w_{1,n} + w_{2,n}\right\|_2^2\right)$$
$$= \liminf_{n\to+\infty}\left(S(W_n) - \frac{1}{p+1}\langle S'(W_n), W_n\rangle\right). \tag{23}$$

Since (W_n) is a Palais-Smale sequence we have

$$\lim_{n\to+\infty}\left(S(W_n) - \frac{1}{p+1}\langle S'(W_n), W_n\rangle\right) = MP,$$

and we can conclude from (22) and (23) that Ψ verifies (21).

We continue by showing that

$$MP \leq NL. \tag{24}$$

Take an element of the Nehari manifold $W \in H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$, $I(W) = 0$. The idea, as in [17, 18], is to construct a path in Γ so that $S(\eta(s))$ achieves its maximum when $\eta(s) = W$. It is easy to see that for C large enough the path η_C defined by $\eta_C(s) = CsW$ fulfills our needs. Indeed, we have

$$\frac{\partial}{\partial s}S(sW) = s\left(\frac{1}{\gamma}(\|\nabla\tilde{w}_1\|_2^2 + (m-\omega^2)\|\tilde{w}_1\|_2^2) + \left\| v\cdot\nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 - s^{p-1}\|w_1\|_{p+1}^{p+1}\right).$$

In particular, $\frac{\partial}{\partial s}S(sW)|_{s=1} = I(W) = 0$. Therefore $\frac{\partial}{\partial s}S(sW) > 0$ for $s \in (0,1)$ and $\frac{\partial}{\partial s}S(sW) < 0$ for $s > 1$. Hence the path $S(\eta_C(s))$ achieves its maximum when $s = \frac{1}{C}$ and $\eta(\frac{1}{C}) = W$. Therefore,

$$ML \leq S(W),$$

and since this is true for any W on the Nehari manifold this proves (24).

It is easy to see that

$$NL \leq LE. \tag{25}$$

Indeed, any solution W of (13) (i.e., any critical point of S) satisfies the Nehari identity $I(W) = 0$. Thus the infimum for NL is taken on a larger set than the infimum for LE, hence (25).

Finally, as a direct consequence of $S'(\Psi) = 0$ and the definition of LE we have

$$LE \leq S(\Psi). \tag{26}$$

Combining (21), (24), (25), (26) finishes the proof of Lemma 6. □

Proof of Proposition 3. In view of Lemmas 4 and 6, the only thing left to prove is that $\Psi = \Phi_{\omega,v}$. Let us first see the case $v = 0$. Since $\Psi = (\psi_1, \psi_2)$ is a critical point of S, we have $\psi_2 = i\omega\psi_1$. Therefore, since φ_ω is a ground state of (5), we have

$$\begin{aligned} S(\Psi) &= \frac{1}{2}\|\nabla\psi_1\|_2^2 + \frac{1}{2}(m - \omega^2)\|\psi_1\|_2^2 - \frac{1}{p+1}\|\psi_1\|_{p+1}^{p+1} \\ &\geq \frac{1}{2}\|\nabla\varphi_\omega\|_2^2 + \frac{1}{2}(m - \omega^2)\|\varphi_\omega\|_2^2 - \frac{1}{p+1}\|\varphi_\omega\|_{p+1}^{p+1} = S(\Phi_{\omega,0}). \end{aligned}$$

Therefore when $v = 0$, we indeed have $\Psi = \Phi_{\omega,0}$. Let us now treat the case $v \neq 0$. Let $\tilde{\psi}_1$ be such that $\psi_1(x) = e^{-i\omega\gamma v\cdot x}\tilde{\psi}_1(x + (\gamma - 1)x_v)$ and define $\tilde{\psi}_2 := i\omega\tilde{\psi}_1$. Then $\tilde{\Psi} := (\tilde{\psi}_1, \tilde{\psi}_2)$ is a solution to (13) with $v = 0$. Indeed, it is not hard to see that

$$-\Delta\psi_1 + i\frac{\omega}{\gamma}\psi_2 - v\cdot\nabla\psi_2 = e^{-i\omega\gamma v\cdot x}(-\Delta\tilde{\psi}_1 - \omega^2\tilde{\psi}_1).$$

Hence

$$S(\Psi) = \frac{1}{\gamma}(E + \omega Q)(\tilde{\Psi}) \geq \frac{1}{\gamma}(E + \omega Q)(\Phi_{\omega,0}) = S(\Phi_{\omega,v}).$$

This implies that $\Psi = \Phi_{\omega,v}$ for any v and finishes the proof of Proposition 3. □

4.2. *Kernel*

Lemma 7. *The following description holds for the kernel of* $S''(\Phi_{\omega,v})$:

$$\mathrm{Ker}(S''(\Phi_{\omega,v})) = \mathrm{Span}\{i\Phi_{\omega,v}, \nabla\Phi_{\omega,v}\}.$$

Proof. The inclusion $\supset$ is easy to obtain. Indeed, due to invariance by translation and phase shifts, for any $\theta \in \mathbb{R}$ and $y \in \mathbb{R}^d$ we have

$$S'(e^{i\theta}\Phi_{\omega,v}(\cdot + y)) = 0.$$

The result is obtained by deriving with respect to θ and y at $\theta = 0, y = 0$. The reverse inclusion is much more delicate. We shall rely on existing results for standing waves of NLS to prove it. First remark that if $W = (w_1, w_2)$ belongs to the kernel of $S''(\Phi_{\omega,v})$, then it satisfies

$$\begin{cases} -\Delta w_1 + m^2 w_1 - (p-1)|\Phi^1_{\omega,v}|^{p-3}\Phi^1_{\omega,v}\mathrm{Re}(\Phi^1_{\omega,v}\bar{w}_1) \\ \quad + |\Phi^1_{\omega,v}|^{p-1} w_1 + i\frac{\omega}{\gamma} w_2 - v\cdot\nabla w_2 = 0, \\ \qquad w_2 - i\frac{\omega}{\gamma} w_1 + v\cdot\nabla w_1 = 0. \end{cases}$$

Here, we have denoted by $\Phi^1_{\omega,v}$ the first component of $\Phi_{\omega,v}$, i.e.,

$$\Phi^1_{\omega,v} := e^{-i\gamma\omega v\cdot x}\varphi_\omega(x + (\gamma-1)x_v).$$

Take $\widetilde{W} := (\tilde{w}_1, \tilde{w}_2)$ such that

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = e^{-i\gamma\omega v\cdot x}\begin{pmatrix} \tilde{w}_1(x + (\gamma-1)x_v) \\ \gamma\tilde{w}_2(x+(\gamma-1)x_v) - \gamma v\nabla\tilde{w}_1(x+(\gamma-1)x_v) \end{pmatrix}.$$

It is a lengthy but straightforward computation to verify that $\widetilde{W}$ satisfies

$$\begin{cases} -\Delta\tilde{w}_1 + m^2\tilde{w}_1 - (p-1)|\Phi^1_{\omega,0}|^{p-3}\Phi^1_{\omega,0}\mathrm{Re}(\Phi^1_{\omega,0}\overline{\tilde{w}_1}) \\ \quad - |\Phi^1_{\omega,0}|^{p-1}\tilde{w}_1 + i\omega\tilde{w}_2 = 0, \\ \qquad \tilde{w}_2 - i\omega\tilde{w}_1 = 0. \end{cases}$$

Remembering now that $\Phi^1_{\omega,0} = \varphi_\omega$ and using the second equation to substitute in the first we get

$$\begin{cases} -\Delta\tilde{w}_1 + (m-\omega^2)\tilde{w}_1 - (p-1)|\varphi_\omega|^{p-1}\mathrm{Re}(\tilde{w}_1) - |\varphi_\omega|^{p-1}\tilde{w}_1 = 0, \\ \qquad \tilde{w}_2 = i\omega\tilde{w}_1. \end{cases} \tag{27}$$

Fortunately we arrive on a known ground: it is well-known since the celebrated work of Weinstein [52] and Kwong [20] (see [5] for a modern short proof of this result) that the only solutions to (27) are

$$\begin{pmatrix} \tilde{w}_1 \\ \tilde{w}_2 \end{pmatrix} \in \mathrm{Span}\left\{\begin{pmatrix} \nabla\varphi_\omega \\ i\omega\nabla\varphi_\omega \end{pmatrix}; \begin{pmatrix} i\varphi_\omega \\ -\omega\varphi_\omega \end{pmatrix}\right\} = \mathrm{Span}\{\nabla\Phi_{\omega,0}; i\Phi_{\omega,0}\}.$$

Coming back into the original variables, this implies that

$$W \in \mathrm{Span}\{\nabla\Phi_{\omega,v}; i\Phi_{\omega,v}\}$$

and finishes the proof. □

4.3. *Coercivity*

The proof of our result relies on the fact that the solitary waves we are considering are stable. In particular, we have at our disposal a coercivity property on the Hessian of the action S related to the soliton profile $\Phi_{\omega,v}$ which allows us to control the difference between a soliton and a function in a neighborhood of its orbit. The coercivity property is the following.

Lemma 8 (Coercivity). *Assume $p < 1 + \frac{4}{d}$ and let $\omega \in \mathbb{R}$ and $v \in \mathbb{R}^d$ be such that $\frac{1}{1+\frac{4}{p-1}-d} < \frac{\omega^2}{m} < 1$, and $|v| < 1$ (i.e., they are compatible with $\mathcal{O}_{\text{stab}}$ defined in* (1)) *and let $\Phi_{\omega,v}$ be the associated ground state. There exists δ such that for any $W \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ satisfying the orthogonality conditions*

$$(W, \nabla\Phi_{\omega,v})_2 = (W, iJ\Phi_{\omega,v})_2 = (W, i\Phi_{\omega,v})_2 = 0 \tag{28}$$

we have

$$H_{\omega,v}(W) \geq \delta \|W\|^2_{H^1 \times L^2}$$

where for brevity in notation we defined

$$H(W) := \langle S''(\Phi_{\omega,v})W, W\rangle.$$

Similar results date back to the work of Weinstein [52, 53] for NLS equations. These ideas were later generalized by Grillakis, Shatah and Strauss [14] in an abstract setting. More recently, Stuart [51] described precisely the orbital stability of solitons of NLKG using also a coercivity statement, but with different orthogonality conditions and a slightly more complicated proof.

Proof of Lemma 8. Step 1: Analysis of the spectrum of $S''(\Phi_{\omega,v})$.

We first remark that, due the exponential localization of $\Phi_{\omega,v}$, the operator $S''(\Phi_{\omega,v})$ is a compact perturbation of the self-adjoint operator

$$\mathbb{L} := \begin{pmatrix} -\Delta + m & 0 \\ 0 & 1 \end{pmatrix} + \frac{\omega}{\gamma}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} + v \cdot \begin{pmatrix} 0 & -\nabla \\ \nabla & 0 \end{pmatrix}, \quad D(\mathbb{L}) = H^2(\mathbb{R}^d) \times H^1(\mathbb{R}^d).$$

By Weyl's Theorem, $S''(\Phi_{\omega,v})$ and $\mathbb{L}$ share the same essential spectrum, that we now analyze. Observe that for $W = (w_1, w_2) \in H^2(\mathbb{R}^d) \times H^1(\mathbb{R}^d)$ we have

$$\langle \mathbb{L}W, W\rangle = \|\nabla w_1\|_2^2 + m\|w_1\|_2^2 + \|w_2\|_2^2 - 2\frac{\omega}{\gamma}(iw_1, w_2)_2 + 2v \cdot (\nabla w_1, w_2)_2,$$

which, after some factorizations (similar to those used in (14)), we can rewrite

$$\langle \mathbb{L}W, W\rangle = \frac{1}{\gamma}\left(\|\nabla \tilde{w}_1\|_2^2 + (m - \omega^2)\|\tilde{w}_1\|_2^2\right) + \left\|v \cdot \nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2$$

where $\tilde{w}_1$ is such that $w_1(x) = e^{-i\omega\gamma v\cdot x}\tilde{w}_1(x + (\gamma - 1)x_v)$. From Lemma 5, we see that there exists $\delta > 0$ such that for any $W \in H^2(\mathbb{R}^d) \times H^1(\mathbb{R}^d)$ we have

$$\langle \mathbb{L}W, W\rangle \geq \delta\|W\|^2_{H^1 \times L^2}.$$

This implies that the essential spectrum of $S''(\Phi_{\omega,v})$ is positive and away from 0. The rest of its spectrum consists in a finite number of isolated eigenvalues. It turns out that from the variational characterization of $\Phi_{\omega,v}$, we can infer that $S''(\Phi_{\omega,v})$ has Morse Index 1, i.e., it admits only one negative simple eigenvalue (see e.g., [1]). We denote this eigenvalue by $-\lambda < 0$, and Ψ an associated normalized eigenvector, i.e., $S''(\Phi_{\omega,v})\Psi = -\lambda\Psi$ and $\|\Psi\|_{L^2 \times L^2} = 1$.

Step 2: A positivity property. We prove now that if $W \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ *satisfies the orthogonality conditions* (28)*, then*

$$\langle S''(\Phi_{\omega,v})W, W\rangle > 0.$$

A particular vector associated with $S''(\Phi_{\omega,v})$ is $\Lambda_\omega \Phi_{\omega,v}$, where $\Lambda_\omega := \frac{\partial}{\partial\omega}$. Indeed, deriving $S'(\Phi_{\omega,v}) = 0$ with respect to ω we get

$$S''(\Phi_{\omega,v})\Lambda_\omega \Phi_{\omega,v} = -\frac{1}{\gamma}Q'(\Phi_{\omega,v}) = -\frac{1}{\gamma}iJ\Phi_{\omega,v}. \tag{29}$$

This implies, using (7),

$$\begin{aligned}
&\langle S''(\Phi_{\omega,v})\Lambda_\omega\Phi_{\omega,v}, \Lambda_\omega\Phi_{\omega,v}\rangle \\
&\quad = -\frac{1}{\gamma}\langle Q'(\Phi_{\omega,v}), \Lambda_\omega\Phi_{\omega,v}\rangle = -\frac{1}{\gamma}\Lambda_\omega Q(\Phi_{\omega,v}) \\
&\quad = \Lambda_\omega\left(\frac{\omega}{\gamma}\|\varphi_\omega\|_2^2\right) = \frac{(m-\omega^2)^{\frac{2}{p-1}-\frac{d}{2}}}{\gamma}\left(-\frac{2\omega^2\left(\frac{2}{p-1}+\frac{d}{2}\right)}{m-\omega^2} - 1\right)\|\tilde{\varphi}\|_2^2 < 0
\end{aligned} \tag{30}$$

where the last inequality follows from the fact that ω is compatible with $\mathcal{O}_{\text{stab}}$ (see the definition (1)). It is easy to verify that $\Lambda_\omega\Phi_{\omega,v}$ is orthogonal to the kernel of $S''(\Phi_{\omega,v})$, namely that we have

$$(\Lambda_\omega\Phi_{\omega,v}, i\Phi_{\omega,v})_2 = (\Lambda_\omega\Phi_{\omega,v}, \nabla\Phi_{\omega,v})_2 = 0.$$

Let us write the orthogonal decomposition of $\Lambda_\omega\Phi_{\omega,v}$ along the spectrum of $S''(\Phi_{\omega,v})$:

$$\Lambda_\omega\Phi_{\omega,v} = \alpha\Psi + \Pi, \tag{31}$$

where $\alpha \neq 0$ and Π is in the positive eigenspace of $S''(\Phi_{\omega,v})$, in particular

$$\langle S''(\Phi_{\omega,v})\Pi, \Pi\rangle \geq \delta\|\Pi\|^2_{H^1\times L^2}.$$

From (30) and (31), we infer that

$$-\lambda\alpha^2 + \langle S''(\Phi_{\omega,v})\Pi, \Pi\rangle = \langle S''(\Phi_{\omega,v})\Lambda_\omega\Phi_{\omega,v}, \Lambda_\omega\Phi_{\omega,v}\rangle < 0. \tag{32}$$

Take now $W \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ satisfying the orthogonality conditions (28). We also write the orthogonal decomposition of W along the spectrum of $S''(\Phi_{\omega,v})$:

$$W = \beta\Psi + \Xi, \tag{33}$$

where $\beta \in \mathbb{R}$ and Ξ is in the positive eigenspace of $S''(\Phi_{\omega,v})$. If $\beta = 0$, the conclusion follows, so we assume $\beta \neq 0$. Using (28), (29), (31) and (33), we have

$$0 = \langle S''(\Phi_{\omega,v})\Lambda_\omega\Phi_{\omega,v}, W\rangle = -\lambda\alpha\beta + \langle S''(\Phi_{\omega,v})\Pi, \Xi\rangle.$$

Note that on the positive spectral subspace of $S''(\Phi_{\omega,v})$, Cauchy-Schwartz inequality holds:

$$\langle S''(\Phi_{\omega,v})\Pi, \Xi\rangle^2 \leq \langle S''(\Phi_{\omega,v})\Pi, \Pi\rangle\langle S''(\Phi_{\omega,v})\Xi, \Xi\rangle.$$

Therefore,

$$\begin{aligned}\langle S''(\Phi_{\omega,v})W, W\rangle &= -\lambda\beta^2 + \langle S''(\Phi_{\omega,v})\Xi, \Xi\rangle \\ &\geq -\lambda\beta^2 + \frac{\langle S''(\Phi_{\omega,v})\Pi, \Xi\rangle^2}{\langle S''(\Phi_{\omega,v})\Pi, \Pi\rangle} > -\lambda\beta^2 + \frac{(-\lambda\alpha\beta)^2}{\lambda\alpha^2} = 0.\end{aligned}$$

Step 3. The coercivity property.

Assume by contradiction that there exists $(W_n = (w_1^n, w_2^n)) \subset H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ satisfying the orthogonality conditions (28) and such that

$$\|W_n\|_{H^1\times L^2} = 1 \quad \text{and} \quad \lim_{n\to+\infty} \langle S''(\Phi_{\omega,v})W_n, W_n\rangle = 0.$$

Recall that, as for (14), we have

$$\begin{aligned}\langle S''(\Phi_{\omega,v})W_n, W_n\rangle &= \frac{1}{\gamma}\left(\|\nabla\tilde{w}_1^n\|_2^2 + (m-\omega^2)\|\tilde{w}_1^n\|_2^2\right) + \left\|v\cdot\nabla w_1 - i\frac{\omega}{\gamma}w_1 + w_2\right\|_2^2 \\ &\quad - \int_{\mathbb{R}^d}\left((p-1)|\Phi_{\omega,v}^1|^{p-3}\mathrm{Re}(\Phi_{\omega,v}^1\bar{w}_1^n)^2 + |\Phi_{\omega,v}^1|^{p-1}|w_1^n|^2\right)dx.\end{aligned}$$

where $\tilde{w}_1^n$ is such that $w_1^n(x) = e^{-i\omega\gamma v\cdot x}\tilde{w}_1^n(x + (\gamma-1)x_v)$.

Since (W_n) is bounded in $H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$, there exists $W \in H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$ such that

$$W_n \rightharpoonup W \text{ as } n\to+\infty \text{ weakly in } H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d).$$

On one hand W must satisfy (28) and from Step 2 we have, if $W \neq 0$.

$$\langle S''(\Phi_{\omega,v})W, W\rangle > 0.$$

On the other hand, by weak convergence and exponential decay of $\Phi_{\omega,v}$ we have

$$\langle S''(\Phi_{\omega,v})W, W\rangle \leq \liminf_{n\to+\infty}\langle S''(\Phi_{\omega,v})W_n, W_n\rangle = 0.$$

Therefore W must be $W \equiv 0$. However, in this case it would implies

$$-\int_{\mathbb{R}^d}\left((p-1)|\Phi_{\omega,v}^1|^{p-3}\mathrm{Re}(\Phi_{\omega,v}^1\bar{w}_1^n)^2 + |\Phi_{\omega,v}^1|^{p-1}|w_1^n|^2\right)dx \to 0$$

and since $\|W_n\|_{H^1\times L^2} = 1$, we would have (using Lemma 5)

$$\langle S''(\Phi_{\omega,v})W_n, W_n\rangle \geq \delta > 0,$$

which is a contradiction. □

5. Modulation Theory

The use of a coercivity property similar to Lemma 8 but adapted to a multi-soliton (see Lemma 12) will require to deal with orthogonality conditions. These orthogonality conditions will obtained by modulation.

Given parameters (ε, L), consider a neighborhood of the sum of solitons

$$\mathcal{U}(\varepsilon, L) := \left\{ U \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d);\ \inf_{\substack{\xi_j > \xi_{j-1}+L \\ \vartheta_j \in \mathbb{R} \\ j=1,\dots,N}} \left\| U - \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(\cdot - \xi_j) \right\|_{H^1 \times L^2} < \varepsilon \right\}.$$

The main result of this section is the following.

Proposition 9 (Dynamical Modulation). *There exists $\tilde{\varepsilon}, \widetilde{L}, C, \widetilde{C} > 0$ such that for any $0 < \varepsilon < \tilde{\varepsilon}$ and $L > \widetilde{L}$ the following property is verified.*

Let $U(t, x) = (u_1, u_2)(t, x)$ be a solution of (4) satisfying on a time interval I

$$U \in \mathcal{U}(\varepsilon, L), \quad \text{for all } t \in I.$$

For $j = 1, \dots, N$, there exist (unique) $\mathscr{C}^1$ functions

$$\tilde{\theta}_j : I \to \mathbb{R}, \quad \tilde{\omega}_j : I \to (-\sqrt{m}, \sqrt{m}), \qquad \tilde{x}_j : I \to \mathbb{R}^d,$$

such that if we define $\widetilde{R}_j(t)$ and $\Upsilon(t)$ by

$$\widetilde{R}_j(t) = e^{i\tilde{\theta}_j(t)} \Phi_{\tilde{\omega}_j(t), v_j}(\cdot - \tilde{x}_j(t)), \qquad \Upsilon(t) = U(t) - \sum_{j=1}^{N} \widetilde{R}_j(t), \tag{34}$$

then Υ satisfies for all $t \in I$ the orthogonality conditions

$$(\Upsilon, i\widetilde{R}_j)_2 = (\Upsilon, iJ\widetilde{R}_j)_2 = (\Upsilon, \nabla \widetilde{R}_j)_2 = 0, \quad j = 1, \dots, N. \tag{35}$$

Moreover, for all $t \in I$ we have

$$\|\Upsilon\|_{H^1 \times L^2} + \sum_{j=1}^{N} |\tilde{\omega}_j - \omega_j| \le \widetilde{C}\varepsilon, \quad \tilde{x}_{j+1} - \tilde{x}_j > \frac{L}{2}, \ j = 1, \dots, N-1,$$

and the derivatives in time verify

$$\sum_{j=1}^{N} \left(|\partial_t \tilde{\omega}_j| + \left| \partial_t \tilde{\theta}_j - \frac{\tilde{\omega}_j}{\gamma_j} \right|^2 + |\partial_t \tilde{x}_j - v_j|^2 \right) < C \left(\|\Upsilon\|_2^2 + e^{-3\alpha\sqrt{m-\omega_\star^2} v_\star t} \right). \tag{36}$$

The proof of Proposition 9 relies on the following Lemma. Note that this lemma is valid for time-independent functions.

Lemma 10 (Static Modulation). *There exist $\widetilde{L}, \widetilde{C}, \tilde{\varepsilon} > 0$ such that for any $L > \widetilde{L}$, $0 < \varepsilon < \tilde{\varepsilon}$, the following property is verified.*

For $j = 1, \dots, N$, there exist (unique) $\mathscr{C}^1$ functions

$$\tilde{\theta}_j : \mathscr{U}(\varepsilon, L) \to \mathbb{R}, \quad \tilde{\omega}_j : \mathscr{U}(\varepsilon, L) \to (-\sqrt{m}, \sqrt{m}), \quad \tilde{x}_j : \mathscr{U}(\varepsilon, L) \to \mathbb{R}^d,$$

such that if we define $\tilde{\Phi}_j$ and Υ by

$$\tilde{\Phi}_j = e^{i\tilde{\theta}_j}\Phi_{\tilde{\omega}_j, v_j}(\cdot - \tilde{x}_j), \quad \Upsilon = U - \sum_{j=1}^{N} \tilde{\Phi}_j$$

then Υ satisfies the orthogonality conditions

$$(\Upsilon, i\tilde{\Phi}_j)_2 = (\Upsilon, iJ\tilde{\Phi}_j)_2 = (\Upsilon, \nabla\tilde{\Phi}_j)_2 = 0, \quad j = 1, \dots, N. \tag{37}$$

Moreover,

$$\|\Upsilon\|_{H^1 \times L^2} + \sum_{j=1}^{N} |\tilde{\omega}_j - \omega_j| \le \widetilde{C}\varepsilon, \quad \tilde{x}_{j+1} - \tilde{x}_j > \frac{L}{2}, \ j = 1, \dots, N-1.$$

In the proofs, we will use the notation Λ_{ω_j} for the scaling operator, i.e.,

$$\Lambda_{\omega_j}\tilde{\Phi}_j := \frac{\partial}{\partial\omega} e^{i\tilde{\theta}_j}\Phi_{\omega, v_j}(\cdot - \tilde{x}_j)\Big|_{\omega = \tilde{\omega}_j}.$$

Proof. We start by proving the lemma in a ball. Take $\varepsilon > 0$, $L > 0$, $(\vartheta_j)_{j=1,\dots,N} \subset \mathbb{R}$ and $(\xi_j)_{j=1,\dots,N} \subset \mathbb{R}^d$ such that $\xi_{j+1} > \xi_j + L$ for $j = 1, \dots, N-1$. Let $\mathscr{B}(\varepsilon)$ denote the ball of $H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ defined by

$$\mathscr{B}(\varepsilon) = \left\{ U \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d); \left\| U - \sum_{j=1}^{N} e^{i\vartheta_j}\Phi_j(x - \xi_j) \right\|_{H^1 \times L^2} < \varepsilon \right\}.$$

Define $\mathfrak{p}_0 := (\vartheta_1, \dots, \vartheta_N, \omega_1, \dots, \omega_N, \xi_1, \dots, \xi_N)$, and let $\mathfrak{P} \subset \mathbb{R}^N \times \mathbb{R}^N \times (\mathbb{R}^d)^N$, be a neighborhood of $\mathfrak{p}_0$. We denote by

$$\mathfrak{p} = (\theta_1, \dots, \theta_N, \varpi_1, \dots, \varpi_N, y_1, \dots, y_N)$$

a generic element of $\mathfrak{P}$. We define the functional $F : \mathfrak{P} \times \mathscr{B}(\varepsilon) \to (\mathbb{R}^{(d+2)})^N$ by

$$F(\mathfrak{p}, U) := \begin{pmatrix} F_{k,1}(\mathfrak{p}, U) \\ F_{k,2}(\mathfrak{p}, U) \\ F_{k,3}(\mathfrak{p}, U) \end{pmatrix}_{k=1,\dots,N},$$

where for $k = 1, \dots, N$ we have set

$$F_{k,1}(\mathfrak{p}, U) = \left(U - \sum_{j=1}^{N} e^{i\theta_j}\tau_{y_j}\Phi_{\varpi_j, v_j}, ie^{i\theta_k}\tau_{y_k}\Phi_{\varpi_k, v_k} \right)_2,$$

$$F_{k,2}(\mathfrak{p}, U) = \left(U - \sum_{j=1}^{N} e^{i\theta_j}\tau_{y_j}\Phi_{\varpi_j, v_j}, ie^{i\theta_k}\tau_{y_k}J\Phi_{\varpi_k, v_k} \right)_2,$$

$$F_{k,3}(\mathfrak{p}, U) = \left(U - \sum_{j=1}^{N} e^{i\theta_j} \tau_{y_j} \Phi_{\varpi_j, v_j}, e^{i\theta_k} \tau_{y_k} \nabla \Phi_{\varpi_k, v_k} \right)_2 .$$

Here, τ_y is the translation by y, i.e., $\tau_y v(x) = v(x-y)$. We clearly have

$$F\left(\mathfrak{p}_0, \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(x-\xi_j) \right) = 0.$$

The lemma inside the ball will follow from the Implicit Function Theorem if we prove that

$$\frac{\partial F}{\partial \mathfrak{p}} \left(\mathfrak{p} = \mathfrak{p}_0, U = \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(x-\xi_j) \right) \text{ is invertible.} \tag{38}$$

The computation of the derivative is not very hard. Many terms will be made small using the exponential decay of the profiles. Other will cancel due to orthogonality. We will essentially be left with a diagonal matrix with nonzero entries, hence the invertibility. We give only some representative calculations. Let's start by

$$\frac{\partial F_{k,1}}{\partial \theta_j} \left(\mathfrak{p}_0, \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(x-\xi_j) \right) = -(ie^{i\vartheta_j} \tau_{\xi_j} \Phi_{\omega_j, v_j}, ie^{i\vartheta_k} \tau_{\xi_k} \Phi_{\omega_k, v_k})_2.$$

When $j = k$, we readily have

$$\frac{\partial F_{k,1}}{\partial \theta_k} \left(\mathfrak{p}_0, \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(x-\xi_j) \right) = -\|\Phi_{\omega_k, v_k}\|_2^2.$$

Assume now $j \neq k$. Then, by exponential decay (see (6)), we have

$$\begin{aligned} |\tau_{\xi_j} \Phi_{\omega_j, v_j} \tau_{\xi_k} \Phi_{\omega_k, v_k}| &\leq Ce^{-\frac{\sqrt{m-\omega_j^2}}{2}|x-\xi_j|} e^{-\frac{\sqrt{m-\omega_k^2}}{2}|x-\xi_k|} \\ &\leq Ce^{-\frac{\sqrt{m-\omega_j^2}}{4}|x-\xi_j|} e^{-\frac{\sqrt{m-\omega_k^2}}{4}|x-\xi_k|} e^{-\frac{\sqrt{\min\{m-\omega_j{}^2, m-\omega_k{}^2\}}}{4}|\xi_j-\xi_k|}. \end{aligned}$$

Therefore, since $|\xi_j - \xi_k| > L$, this implies

$$\left| \frac{\partial F_{k,1}}{\partial \theta_j} \left(\mathfrak{p}_0, \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(x-\xi_j) \right) \right| \leq Ce^{-\frac{\sqrt{\min\{m-\omega_j^2, m-\omega_k^2\}}}{4} L}. \tag{39}$$

This quantity can be made as small as we need by increasing the value of L. For the derivative with respect to ϖ_j, we have

$$\frac{\partial F_{k,1}}{\partial \varpi_j} \left(\mathfrak{p}_0, \sum_{j=1}^{N} e^{i\vartheta_j} \Phi_j(x-\xi_j) \right) = -(e^{i\theta_j} \tau_{\xi_j} \Lambda_{\omega_j} \Phi_{\omega_j, v_j}, ie^{i\theta_k} \tau_{\xi_k} \Phi_{\omega_k, v_k})_2.$$

When $j \neq k$, this quantity can be made small as in (39). For $j = k$, since $\varphi_\omega \in \mathbb{R}$, we simply have

$$\frac{\partial F_{k,1}}{\partial \varpi_k}\left(\mathfrak{p}_0, \sum_{j=1}^N e^{i\vartheta_j}\Phi_j(x-\xi_j)\right) = -\frac{1}{\gamma}(\Lambda_{\omega_j}\Phi_{\omega_k}, i\Phi_{\omega_k}) = 0.$$

All other computations follow from similar arguments and we finally find that

$$\frac{\partial F}{\partial \mathfrak{p}}\left(\mathfrak{p}_0, \sum_{j=1}^N e^{i\vartheta_j}\Phi_j(x-\xi_j)\right) = DIAG + O(e^{-\frac{\sqrt{m-\omega_\star^2}}{4}L})$$

where $DIAG$ is a diagonal matrix with nonzero entries on the diagonal. Therefore, for L large enough, we have the desired invertibility property (38) and the Implicit Function Theorem implies the result inside the ball. Since any $U \in \mathcal{U}(\varepsilon, L)$ belongs to some ball $\mathcal{B}(\varepsilon)$, the existence part follows in the cylinder $\mathcal{U}(\varepsilon, L)$. To show uniqueness, one has to prove that the functions obtained are independent of the ball chosen, we leave the details of this argument to the reader. □

Proof of Proposition 9. The first part of the statement follows from Lemma 10 (except the regularity that follows from other regularization arguments, see [28]), hence the main thing to check is (36). We first write the equation verified by Υ. Recall that U satisfies $\partial_t U = JE'(U)$. We replace U by $\sum_{j=1}^N \widetilde{R}_j(t) + \Upsilon(t)$ in the previous equation to get

$$\partial_t \Upsilon + \sum_{j=1}^N \left(i\partial_t\tilde{\theta}_j\widetilde{R}_j + \partial_t\tilde{\omega}_j\Lambda_{\tilde{\omega}_j}\widetilde{R}_j - \partial_t\tilde{x}_j\cdot\nabla\widetilde{R}_j\right) = JE'\left(\sum_{j=1}^N \widetilde{R}_j(t) + \Upsilon(t)\right)$$

such that it follows

$$\begin{aligned}\partial_t \Upsilon + \sum_{j=1}^N \left(i\left(\partial_t\tilde{\theta}_j - \frac{\tilde{\omega}_j}{\gamma_j}\right)\widetilde{R}_j + \partial_t\tilde{\omega}_j\Lambda_{\tilde{\omega}_j}\widetilde{R}_j - (\partial_t\tilde{x}_j - v_j)\cdot\nabla\widetilde{R}_j\right)\\ = \mathfrak{L}\Upsilon + \mathfrak{N}(\Upsilon) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t})\end{aligned} \tag{40}$$

where $\mathfrak{L}$ is the linearized operator defined by

$$\mathfrak{L} := J\left(\begin{pmatrix}-\Delta + m & 0\\ 0 & 1\end{pmatrix} - \sum_{j=1}^N \begin{pmatrix}(p-1)|\widetilde{R}_j|^{p-3}\widetilde{R}_j\mathrm{Re}(\widetilde{R}_j\bar{\cdot}) + |\widetilde{R}_j|^{p-1} & 0\\ 0 & 0\end{pmatrix}\right)$$

and $\mathfrak{N}(\Upsilon)$ is the remaining nonlinear part. To write this equation, we have used Lemma 18, the fact that

$$E'\left(\sum_{j=1}^N \widetilde{R}_j\right) = \sum_{j=1}^N E'(\widetilde{R}_j) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}),$$

$$\mathfrak{L} = JE''\left(\sum_{j=1}^N \widetilde{R}_j\right) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t})$$

and that $\widetilde{R}_j$ is a critical point of $E + \frac{\tilde{\omega}_j}{\gamma_j} Q + v_j \cdot P$. An analogous computation is derived in all details in Lemma 14.

Take now the scalar product of (40) with $iJ\widetilde{R}_k$. By using Lemma 18, the definition of $\widetilde{R}_k$, and the orthogonality conditions (35) it follows that

$$\begin{aligned}
(\nabla \widetilde{R}_k, iJ\widetilde{R}_k)_2 &= (i\widetilde{R}_k, iJ\widetilde{R}_k)_2 = 0 \\
(\Lambda_{\tilde{\omega}_j} R_j, iJ\widetilde{R}_k)_2 &= (\nabla \widetilde{R}_j, iJ\widetilde{R}_k)_2 = (i\widetilde{R}_j, iJ\widetilde{R}_k)_2 = O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}) \quad \text{if } j \neq k \\
(\partial_t \Upsilon, iJ\widetilde{R}_k)_2 &= -(\Upsilon, \partial_t iJ\widetilde{R}_k)_2.
\end{aligned}$$

Therefore

$$\partial_t \tilde{\omega}_k (\Lambda_{\tilde{\omega}_k} Q(\widetilde{R}_k)) = (\mathfrak{L}\Upsilon, iJ\widetilde{R}_k)_2 + (\Upsilon, \partial_t iJ\widetilde{R}_k)_2 + O(\|\Upsilon\|^2_{H^1\times L^2}) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}), \tag{41}$$

where the term $\Lambda_{\tilde{\omega}_k} Q(\widetilde{R}_k)$ comes from

$$(\Lambda_{\tilde{\omega}_k} \widetilde{R}_k, iJ\widetilde{R}_k)_2 = (\widetilde{R}_k, iJ\Lambda_{\tilde{\omega}_k} \widetilde{R}_k)_2 = \Lambda_{\tilde{\omega}_k} Q(\widetilde{R}_k).$$

Note that by exponential localization

$$(\mathfrak{L}\Upsilon, iJ\widetilde{R}_k)_2 = (\Upsilon, \mathfrak{L}^\star iJ\widetilde{R}_k)_2 = (\Upsilon, (JE''(\widetilde{R}_k))^\star iJ\widetilde{R}_k)_2 + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

We want to use the fact that $i\widetilde{R}_k$ belongs to the kernel of $S''(\widetilde{R}_k)$ (see Lemma 7), namely that

$$\left(J \left(E'' + \frac{\tilde{\omega}_k}{\gamma_k} Q'' + v_k \cdot P'' \right) (\widetilde{R}_k) \right)^\star iJ\widetilde{R}_k = 0.$$

To this aim, we use the definition (34) of $\widetilde{R}_k$, to compute the following time derivative and make the missing parts appear.

$$\begin{aligned}
(\Upsilon, \partial_t iJ\widetilde{R}_k)_2 = &\frac{\tilde{\omega}_k}{\gamma_k} (\Upsilon, (JQ''(\widetilde{R}_k))^\star iJ\widetilde{R}_k)_2 + v_k \cdot (\Upsilon, (JP''(\widetilde{R}_k))^\star iJ\widetilde{R}_k)_2 \\
&+ \left(\frac{\tilde{\omega}_k}{\gamma_k} - \partial_t \tilde{\theta}_k \right) (\Upsilon, J\widetilde{R}_k)_2 + (v_k - \partial_t \tilde{x}_k) \cdot (\Upsilon, iJ\nabla \widetilde{R}_k)_2 \\
&+ \partial_t \tilde{\omega}_k (\Upsilon, iJ\Lambda_{\tilde{\omega}_k} \widetilde{R}_k)_2.
\end{aligned}$$

Therefore, (41) gives

$$\begin{aligned}
\partial_t \tilde{\omega}_k (\Lambda_{\tilde{\omega}_k} Q(\widetilde{R}_k)) = &\left(\frac{\tilde{\omega}_k}{\gamma_k} - \partial_t \tilde{\theta}_k \right) (\Upsilon, J\widetilde{R}_k)_2 + (v_k - \partial_t \tilde{x}_k)(\Upsilon, iJ\nabla \widetilde{R}_k)_2 \\
&+ \partial_t \tilde{\omega}_k (\Upsilon, iJ\Lambda_{\tilde{\omega}_k} \widetilde{R}_k)_2 + O(\|\Upsilon\|^2_{H^1\times L^2}) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).
\end{aligned} \tag{42}$$

Take the scalar product of (40) with $\frac{\partial}{\partial x_j}\widetilde{R}_k$ for $j = 1, \ldots, d$ to get from similar arguments

$$(v_k - \partial_t \tilde{x}_k)\frac{1}{d}\left\|\nabla\widetilde{R}_k\right\|_2^2 = (\mathfrak{L}\Upsilon, \nabla\widetilde{R}_k)_2 + (\Upsilon, \partial_t\nabla\widetilde{R}_k)_2 + O(\|\Upsilon\|^2_{H^1\times L^2}) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{43}$$

Conversely to what happened for (42), we do not expect to have a cancellation on the linear term. We just estimate it by

$$(\mathfrak{L}\Upsilon, \nabla\widetilde{R}_k)_2 = (\Upsilon, \mathfrak{L}^\star\nabla\widetilde{R}_k)_2 \le C\|\Upsilon\|_2.$$

Therefore (43) gives

$$(v_k - \partial_t \tilde{x}_k)\frac{1}{d}\left\|\nabla\widetilde{R}_k\right\|_2^2 = O(\|\Upsilon\|_{H^1\times L^2}) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{44}$$

Finally, take the scalar product of (40) with $i\widetilde{R}_k$ and argue as previously to obtain

$$(\partial_t\tilde{\theta}_k - \tilde{\omega}_k)\left\|\widetilde{R}_k\right\|_2^2 = O(\|\Upsilon\|_{H^1\times L^2}) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{45}$$

Putting together (42), (44) and (45) we obtain a differential system for the modulation equations vector $Mod(t) := (\partial_t\tilde{\omega}_j, \partial_t\tilde{\theta}_j - \tilde{\omega}_j, \partial_t\tilde{x}_j - v_j)_{j=1,\ldots,N}$ of the form

$$A \cdot Mod(t) = M(\Upsilon) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}),$$

where $|M(\Upsilon)| \le C\|\Upsilon\|_{H^1\times L^2}$. As long as the modulation parameter do not vary too much and $\|\Upsilon\|_{H^1}$ remains small, A is invertible (it is of the form $DIAG + small$ with $DIAG$ a diagonal nondegenerate matrix) and we can deduce that

$$|Mod(t)| \le C\|\Upsilon\|_{H^1\times L^2} + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{46}$$

Coming back now to (42), it is now easy to see that in fact we can improve in part the previous estimate into

$$\sum_{j=1}^{N}|\partial_t\tilde{\omega}_j| \le C\|\Upsilon\|^2_{H^1\times L^2} + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{47}$$

This improvement is due to our choice of orthogonality conditions. Combining (46) and (47) gives the desired result. □

6. Uniform Estimates

This Section is devoted to the proof of Proposition 1. We essentially follow the same line as in the Schrödinger case [29]. Since our approximate multi-solitons have final data $U_n(T_n) = R(T_n)$, they satisfy the desired estimate at least on some interval $[T_n - \delta, T_n]$. Thus the idea is to reduce things to a bootstrap argument: proposition 1 is a consequence of the following proposition.

Proposition 11 (Bootstrap). *There exist $\alpha = \alpha(d, N) > 0$, and $T_0 \in \mathbb{R}$ (independent of n) such that for n large enough the following bootstrap property holds. For $t^\dagger \in [T_0, T^n]$, if U_n satisfies for all $t \in [t^\dagger, T^n]$ the estimate*

$$\|U_n(t) - R(t)\|_{H^1\times L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}, \tag{48}$$

then it will also satisfies for all $t \in [t^\dagger, T^n]$ the better estimate

$$\|U_n(t) - R(t)\|_{H^1\times L^2} \leq \frac{1}{2} e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}. \tag{49}$$

Let us quickly indicate how to obtain Proposition 1 from Proposition 11.

Proof of Proposition 1. Proposition 11 implies Proposition 1 by means of a classical continuity argument (see e.g., [29]). First, let us notice that the map $t \mapsto U_n(t) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ is continuous. Second, let us define

$$t^\star := \inf\{\tau \in [T_0, T^n] \text{ such that (48) holds for all } t \in [\tau, T^n]\}.$$

Recall that $U_n(T^n) = R(T^n)$, therefore we have $T_0 \leq t^\star < T^n$. Our purpose is to show that $t^\star = T_0$. Let us suppose that $t^\star > T_0$. Thanks to (49) we get

$$\|U_n(t) - R(t)\|_{H^1\times L^2} \leq \frac{1}{2} e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t},$$

for all $t \in [t^\star, T^n]$. By continuity it exists $\delta_1 > 0$ such that

$$\|U_n(t) - R(t)\|_{H^1\times L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t},$$

for all $t \in [t^\star - \delta_1, T^n]$. This contradicts the definition of $t^\star$ and finishes the proof. □

Hence now we only have to prove Proposition 11. For the rest of the paper, we make the following assumption.

Bootstrap Assumption. *Let $T_0 > 0$ to be determined later and assume that there exists $t^\dagger \in [T_0, T^n]$ such that U_n satisfies for all $t \in [t^\dagger, T^n]$ the estimate*

$$\|U_n(t) - R(t)\|_{H^1\times L^2} \leq e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}. \tag{50}$$

We want to prove that in fact (50) holds with the better constant $\frac{1}{2}$ on the left hand side.

To prove Proposition 11, we need a way to control the difference between the sum of solitons and the approximate multi-soliton U_n. If there is only one soliton, it is known since the ground work of Weinstein [52] that the coercivity property of the hessian of the action functional (Lemma 8) provides a mean to control the difference between a soliton and a solution close to the orbit of the soliton. As in [7, 8, 29, 34], we are going to generalize such a property to the case of N solitons. To that purpose, we define *localized versions* of the conservation laws around each solitons and prove that a coercivity property also holds for the functional action related to the multi-solitons.

6.1. *Bootstrap*

In this section, we prove Proposition 11, assuming three intermediate Lemmas proved in the later sections.

First of all, we begin by selecting a particular direction of propagation. Define the application $\Omega : \mathbb{S}^{d-1} \to \mathbb{R}$ by

$$\Omega(e) := \prod_{j\neq k} |(v_j - v_k, e)_{\mathbb{R}^d}|, \quad e \in \mathbb{S}^1.$$

Let e_1 be such that

$$\Omega(e_1) = \max\{\Omega(e), e \in \mathbb{S}^{d-1}\} > 0. \tag{51}$$

Here, the sup is a max since we are maximizing a continuous function on a compact set. Let us prove the last inequality. We have

$$\Omega^{-1}(\{0\}) = \bigcup_{j\neq k} \{e \in \mathbb{S}^{d-1}, (v_j - v_k, e)_{\mathbb{R}^d} = 0\}.$$

Each set composing the union on the right is of 0 Lesbegue measure, therefore so is $\Omega^{-1}(\{0\})$. Hence $\mathbb{S}^{d-1} \setminus \Omega^{-1}(\{0\}) \neq \emptyset$ and this proves (51). We can complete e_1 into an orthonormal basis $(e_1, \ldots, e_d)$ of $\mathbb{R}^d$ and we infer from (51) that there exists $\tilde{\alpha} > 0$ such that for any $j \neq k$,

$$|(v_j - v_k, e_1)_{\mathbb{R}^d}| \geq \tilde{\alpha}|v_j - v_k|. \tag{52}$$

Since (4) is rotation-invariant, we can assume that $(e_1, \ldots, e_d)$ is the canonical basis of $\mathbb{R}^d$. Calling v_j^1 the first component of the j-th velocity vector, up to reindexing the solitons, we can assume that

$$v_1^1 < v_2^1 < \cdots < v_N^1.$$

The localization works as follows. We first define a partition of unity $(\phi_j)_{j=1,\ldots,N}$: take ψ a cutoff function such that

$$\psi(s) = 0 \quad \text{for } s < -1 \quad \text{and} \quad \psi(s) = 1 \quad \text{for } s > 1,\ 0 \leq \psi'(s) \leq 1 \text{ in } [-1, 1],$$
$$\exists C > 0,\ \forall s \in \mathbb{R},\ |\psi'(s)| \leq C\sqrt{\psi(s)}. \tag{53}$$

Define

$$\psi_1(t, x) = 1, \quad \psi_j(t, x) = \psi\left(\frac{1}{\sqrt{t}}(x^1 - m_j t)\right), \quad m_j = \frac{1}{2}(v_{j-1}^1 + v_j^1).$$
$$\phi_j = \psi_j - \psi_{j+1} \quad \text{for } j = 1, \ldots, N-1 \text{ and } \phi_N = \psi_N.$$

We consider the following *localized action* functional for $W = (w_1, w_2) \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$

$$\mathcal{S}(t, W) = \sum_{j=1}^N S_j(t, W) = \sum_{j=1}^N E_j(t, W) + \frac{\omega_j}{\gamma_j} Q_j(t, W) + v_j \cdot P_j(t, W),$$

where for $j = 1, \dots, N$ we have defined the *localized energies, charges and momenta* by

$$
\begin{aligned}
E_j(t, W) &:= \frac{1}{2}\int_{\mathbb{R}^d} |\nabla w_1|^2 \phi_j dx + \frac{m}{2}\int_{\mathbb{R}^d} |w_1|^2 \phi_j dx \\
&\quad + \frac{1}{2}\int_{\mathbb{R}^d} |w_2|^2 \phi_j dx - \frac{1}{p+1}\int_{\mathbb{R}^d} |w_1|^{p+1} \phi_j dx, \\
Q_j(t, W) &:= \operatorname{Im} \int_{\mathbb{R}^d} w_1 \bar{w}_2 \phi_j dx, \\
P_j(t, W) &:= \operatorname{Re} \int_{\mathbb{R}^d} \nabla w_1 \bar{w}_2 \phi_j dx.
\end{aligned}
$$

Since U_n verifies (50), we can assume that T_0 is large enough, so that U_n satisfies the hypotheses of Proposition 9 and thus there exists a modulated sum of solitons $\widetilde{R} = \sum_{j=1}^N \widetilde{R}_j$ and Υ_n verifying the orthogonality conditions (35) such that

$$
\begin{aligned}
U_n(t) &= \widetilde{R}(t) + \Upsilon_n(t), \\
\|\Upsilon_n\|_{H^1 \times L^2} &\leq C e^{-\alpha\sqrt{m-\omega_\star^2} v_\star t}. \qquad (54)
\end{aligned}
$$

Let us define the localized linearized action for $\widetilde{R}$ by

$$
\mathscr{H}_n(\Upsilon_n(t), \Upsilon_n(t)) := \sum_{j=1}^N \langle S_j''(\widetilde{R}_j(t))\Upsilon_n(t), \Upsilon_n(t)\rangle.
$$

It turns out that $\mathscr{H}_n$ is inheriting the coercivity property of the hessian of the action around a single soliton (Lemma 8).

Lemma 12 (Coercivity). *There exists $C > 0$ such that for all $t \in [T_0, T^n]$ the localized Hessian verifies*

$$
\mathscr{H}_n(\Upsilon_n(t), \Upsilon_n(t)) \geq C\|\Upsilon_n(t)\|_{H^1 \times L^2}^2.
$$

In addition, since $\mathscr{S}(t, U_n(t))$ is made of localized versions of conserved quantities, it varies slowly.

Lemma 13 (Almost conservation). *For $t \in [t^\star, T^n]$, we have*

$$
\left|\frac{\partial}{\partial t}\mathscr{S}(t, U_n(t))\right| \leq o(e^{-2\alpha\sqrt{m-\omega_\star^2} v_\star t}).
$$

We also have the following Taylor-like expansion for $\mathscr{S}(t, U_n(t))$.

Lemma 14 (Taylor-like expansion). *The action $\mathscr{S}(t, U_n(t))$ satisfies for $t \in [t^\star, T^n]$*

$$
\begin{aligned}
\mathscr{S}(t, U_n(t)) &= \sum_{j=1}^N \left(E(R_j(t)) + \frac{\omega_j}{\gamma_j} Q(R_j(t)) + v_j \cdot P(R_j(t)) \right) \\
&\quad + \mathscr{H}_n(\Upsilon_n(t), \Upsilon_n(t)) + o(e^{-2\alpha\sqrt{m-\omega_\star^2} v_\star t}).
\end{aligned}
$$

With Lemmas 12, 13 and 14 in hand, we can now conclude the proof of Proposition 11.

Proof of Proposition 11. The first step is to show that

$$\|\Upsilon_n\|^2_{H^1\times L^2} = o\left(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}\right). \tag{55}$$

Indeed, thanks to Lemma 13 we obtain

$$\mathscr{S}(t, U_n(t)) \leq \mathscr{S}(T^n, U_n(T^n)) + o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{56}$$

Now notice that $\sum_{j=1}^N \left(E(R_j(t)) + \frac{\omega_j}{\gamma_j}Q(R_j(t)) + v_j\cdot P(R_j(t))\right)$ is a time independent quantity. Therefore,

$$\begin{aligned}
&\sum_{j=1}^N \left(E(R_j(t)) + \frac{\omega_j}{\gamma_j}Q(R_j(t)) + v_j\cdot P(R_j(t))\right)\\
&\quad = \sum_{j=1}^N \left(E(R_j(T_n)) + \frac{\omega_j}{\gamma_j}Q(R_j(T_n)) + v_j\cdot P(R_j(T_n))\right) = \mathscr{S}(T_n, U_n(T_n)).
\end{aligned}$$

Combined with Lemma 14 and (56), this implies

$$\mathscr{H}(\Upsilon_n(t), \Upsilon_n(t)) = \mathscr{S}(t, U_n(t)) - \mathscr{S}(T_n, U_n(T_n)) + o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}) = o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

By Lemma 12 we get

$$C\|\Upsilon_n\|^2_{H^1\times L^2} \leq \mathscr{H}(\Upsilon_n(t), \Upsilon_n(t)) \leq o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

Hence (55) is proved. Now we have

$$\|U_n - R\|^2_{H^1\times L^2} \leq 2\|\widetilde{R} - R\|^2_{H^1\times L^2} + 2\|\Upsilon_n\|^2_{H^1\times L^2},$$

such that, by (55) and (36) we infer

$$\begin{aligned}
\|U_n - R\|^2_{H^1\times L^2} &\leq C\left(\sum_{j=1}^N |\tilde{\omega}_j(t) - \omega_j|^2 + |\tilde{\theta}_j(t) - \theta_j|^2 + |\tilde{x}_j(t) - x_j|^2\right) + o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t})\\
&\leq o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}).
\end{aligned}$$

Choosing t large enough we have

$$\|U_n - R\|_{H^1\times L^2} \leq \frac{1}{2}e^{-\alpha\sqrt{m-\omega_\star^2}v_\star t}.$$

This concludes the proof. □

6.2. *Coercivity*

From now on and until the end of this paper, the subscript n is removed when there is no possible confusion. For example, U_n is now denoted simply by U.

We first prove Lemma 12.

Proof of Lemma 12. From Lemma 8, we already know that for any $j = 1, \dots, N$ we have

$$\langle S''(\widetilde{R}_j(t))\Upsilon(t), \Upsilon(t)\rangle \geq C\|\Upsilon(t)\|^2_{H^1\times L^2}$$

where the dependency of S in j is understood (recall that $S = E + \frac{\tilde{\omega}_j}{\gamma_j}Q + v_j \cdot P$). We remark that

$$\begin{aligned}\langle Q_j''(\widetilde{R}_j)\Upsilon(t), \Upsilon(t)\rangle &= \operatorname{Im}\int_{\mathbb{R}^d} \Upsilon_1(t)\overline{\Upsilon}_2(t)\phi_j dx \\ &= \operatorname{Im}\int_{\mathbb{R}^d} (\sqrt{\phi_j}\Upsilon_1(t)\overline{(\sqrt{\phi_j}\Upsilon_2(t))}dx = \langle Q''(\widetilde{R}_j)\sqrt{\phi_j}\Upsilon(t), \sqrt{\phi_j}\Upsilon(t)\rangle.\end{aligned}$$

Similar computations can be performed for the momentum and the 0-order part of the energy. We deal with the gradient part by means of the classical IMS localization formula (see e.g., [50]):

$$\|\nabla\Upsilon_1\|_2^2 = \sum_{j=1}^{N}\left(\left\|\nabla\left(\sqrt{\phi_j}\Upsilon_1\right)\right\|_2^2 - \left\|\left|\nabla\left(\sqrt{\phi_j}\right)\right|\Upsilon_1\right\|_2^2\right).$$

Straightforward computations using the definition of the cutoff functions ϕ_j and (53) imply that

$$\left\|\nabla\left(\sqrt{\phi_j}\right)\right\|_\infty \leq \frac{C'}{\sqrt{t}}.$$

This implies that

$$\left\|\nabla\left(\sqrt{\phi_j}\right)\Upsilon_1\right\|_2^2 \leq \frac{C'}{\sqrt{t}}\|\Upsilon\|^2_{H^1\times L^2}.$$

Combining these informations, we infer that

$$\begin{aligned}\mathscr{H}(\Upsilon(t), \Upsilon(t)) &= \sum_{j=1}^{N}\langle S_j''(\widetilde{R}_j(t))\Upsilon(t), \Upsilon(t)\rangle \\ &= \sum_{j=1}^{N}\langle S''(\widetilde{R}_j(t))\sqrt{\phi_j}\Upsilon(t), \sqrt{\phi_j}\Upsilon(t)\rangle - \left\|\nabla\left(\sqrt{\phi_j}\right)\Upsilon_1\right\|_2^2 \\ &\geq \left(C - \frac{C'}{\sqrt{t}}\right)\|\Upsilon(t)\|^2_{H^1\times L^2} \geq \frac{C}{2}\|\Upsilon(t)\|^2_{H^1\times L^2},\end{aligned}$$

where the last inequality follows from the fact that $t \geq T_0$ and T_0 can be chosen so that $T_0 \geq \left(\frac{2C'}{C}\right)^2$. □

6.3. *Almost Conservation*

In this section, we prove Lemma 13. Recall that we have assumed that $U \equiv U_n$ verifies the bootstrap assumption (50). We start with a preliminary lemma.

Lemma 15. *Let ϕ be a $\mathscr{C}^1$ function of the variable x_1 such that ϕ and ϕ' are bounded. Then for all t in the time interval of existence of U we have*

$$\frac{\partial}{\partial t}\mathrm{Im}\int_{\mathbb{R}^d} u_1\bar{u}_2\phi(x_1)dx = \mathrm{Im}\int_{\mathbb{R}^d} \partial_{x_1}u_1\bar{u}_1\phi'(x_1)dx,$$

$$\frac{\partial}{\partial t}\mathrm{Re}\int_{\mathbb{R}^d} \partial_{x_j}u_1\bar{u}_2\phi(x_1)dx = -\mathrm{Re}\int_{\mathbb{R}^d} \partial_{x_j}u_1\partial_{x_1}\bar{u}_1\phi'(x_1)dx \quad (j \neq 1),$$

$$\frac{\partial}{\partial t}\mathrm{Re}\int_{\mathbb{R}^d} \partial_{x_1}u_1\bar{u}_2\phi(x_1)dx = \int_{\mathbb{R}^d}\bigg(-|\partial_{x_1}u_1|^2 + \frac{1}{2}(|\nabla u_1|^2 + m|u_1|^2 - |u_2|^2) - \frac{1}{p+1}|u_1|^{p+1}\bigg)\phi'(x_1)dx.$$

Proof. The results follows from elementary computations using the fact that U is a solution to (4). □

Proof of Lemma 13. Let $U = (u_1, u_2)$ and let us start by looking at the derivative of the localized charge. By Lemma 15 we have

$$\begin{aligned}\left|\frac{\partial}{\partial t}\mathrm{Im}\int_{\mathbb{R}^d} u_1\bar{u}_2\psi_j dx\right| &= \left|\mathrm{Im}\int_{\mathbb{R}^d}\partial_{x_1}u_1\bar{u}_1\partial_{x_1}\psi_j dx + \mathrm{Im}\int u_1\bar{u}_2\partial_t\psi_j dx\right| \\ &= \frac{1}{\sqrt{t}}\left|\mathrm{Im}\int_{\mathbb{R}^d}\left(\partial_{x_1}u_1\bar{u}_1 - \frac{m_j}{2}u_1\bar{u}_2\right)\psi'\left(\frac{x^1 - m_j t}{\sqrt{t}}\right)dx\right| \\ &\leq \frac{C}{\sqrt{t}}\int_{\widetilde{A}_j}(|\partial_{x_1}u_1\bar{u}_1| + |u_1\bar{u}_2|)dx,\end{aligned}$$

where $\widetilde{A}_j := \{x \in \mathbb{R}^d;\ \psi_j'(x) \neq 0\}$. Remembering that $\phi_j = \psi_j - \psi_{j+1}$ for $j = 1, \dots, N-1$ and $\phi_N = \psi_N$, and defining $A_j := \{x \in \mathbb{R}^d;\ \phi_j' \neq 0\}$ we have

$$\left|\frac{\partial}{\partial t}\mathrm{Im}\int_{\mathbb{R}^d} u_1\bar{u}_2\phi_j dx\right| \leq \frac{C}{\sqrt{t}}\int_{A_j}(|\partial_{x_1}u_1\bar{u}_1| + |u_1\bar{u}_2|)dx,$$

and then

$$\left|\frac{\partial}{\partial t}\mathrm{Im}\int u_1\bar{u}_2\phi_j dx\right| \leq \frac{C}{\sqrt{t}}\left(\|u_1\|^2_{H^1(A_j)} + \|u_2\|^2_{L^2(A_j)}\right).$$

Now notice that $\|U\|^2_{H^1(A_j)\times L^2(A_j)} \leq 2\|U - R\|^2_{H^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)} + 2\|R\|^2_{H^1(A_j)\times L^2(A_j)}$. Thanks to Lemma 18 we have

$$\|R\|^2_{H^1(A_j)\times L^2(A_j)} \leq \|R_j\|^2_{H^1(A_j)\times L^2(A_j)} + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

By using the properties of our partition of unity and the decay of the profile of R_j it follows that

$$\|R_j\|^2_{H^1(A_j)\times L^2(A_j)} \leq \int_{|x^1|\geq \frac{\tilde{\alpha}}{2}v_\star t} Ce^{-\frac{1}{2}\sqrt{m-\omega_\star^2}|x|}dx \leq Ce^{-\frac{\tilde{\alpha}}{4}\sqrt{m-\omega_\star^2}v_\star t}.$$

Recall that $\tilde{\alpha}$ stems from (52). We conclude thanks to the bootstrap assumption (50) that

$$\left|\frac{\partial}{\partial t}Q_j(t, U(t))\right| = o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}) \tag{57}$$

if we choose $\alpha < \frac{\tilde{\alpha}}{8}$. Now we focus on the derivative of the localized momenta. We start with the first component of the momentum. By Lemma 15 we have

$$\begin{aligned}
&\left|\frac{\partial}{\partial t}\mathrm{Re}\int \partial_{x_1}u_1\bar{u}_2\psi_j dx\right| \\
&\quad\leq \frac{1}{\sqrt{t}}\int_{\mathbb{R}^d}\left[\partial_{x_1}u_1\bar{u}_2 - \left(-\frac{m_1}{2}|\partial_{x_1}u_1|^2 + \frac{|u_2|^2}{2} - \frac{|\nabla u_1|^2}{2} - \frac{m|u_1|^2}{2}\right)\right. \\
&\qquad \left. - \frac{|u_1|^{p+1}}{p+1}\right]\psi'\left(\frac{x_1 - m_1 t}{\sqrt{t}}\right)dx \\
&\quad\leq \frac{C}{\sqrt{t}}\left[\|u_1\|^2_{H^1(A_j)} + \|u_2\|^2_{L^2(A_j)} + \|u_1\|^{p+1}_{L^{p+1}(A_j)}\right] \\
&\quad\leq \frac{C}{\sqrt{t}}\left[\|u_1\|^2_{H^1(A_j)} + \|u_2\|^2_{L^2(A_j)} + \|u_1\|^{p+1}_{H^1(A_j)}\right].
\end{aligned}$$

Now we argue as for the derivative of the localized charge. The other components of the momentum can be estimated in a similar fashion. Thus we have

$$\left|\frac{\partial}{\partial t}P_j(t, U(t))\right| = o(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}). \tag{58}$$

Remark now that

$$\mathscr{S}(t, U(t)) = E(U(t)) + \sum_{j=1}^{N}\frac{\omega_j}{\gamma}Q_j(t, U(t)) + v_j \cdot P_j(t, U(t)).$$

Since E is a conserved quantity, combining (57) and (58) gives the desired result. □

6.4. *The Taylor Expansion*

We now prove Lemma 14. We start by an estimate on the modulation parameters.

Lemma 16. *For any $t \in [t^\star, T^n]$, we have*

$$\sum_{j=1}^{N}|\tilde{\omega}_j(t) - \omega_j| = O(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

Proof. Recall that $U = \sum_{j=1}^{N} \widetilde{R}_j + \Upsilon$. Thanks to the interaction estimates given by Lemma 18 and the orthogonality conditions it follows

$$Q_j(t, U) = Q(\widetilde{R}_j) + \operatorname{Im} \int_{\mathbb{R}^d} \Upsilon_1 \overline{\Upsilon}_2 \phi_j dx + O(e^{-3\alpha\sqrt{m-\omega_\star^2} v_\star t}). \tag{59}$$

We already computed the time-derivative of Q_j during the proof of Lemma 13 (see (57)), and it implies

$$|Q_j(t, U(t)) - Q_j(T^n, U(T_n))| = o(e^{-2\alpha\sqrt{m-\omega_\star^2} v_\star t}). \tag{60}$$

Thanks to the scaling property (7) of the profile we get

$$\begin{aligned} &Q(\widetilde{R}_j(t)) - Q(\widetilde{R}_j(T_n)) \\ &= \gamma_j \left(-\tilde{\omega}_j(t)(m - \tilde{\omega}_j(t)^2)^{\frac{2}{p-1}-\frac{d}{2}} + \tilde{\omega}_j(T^n)(m - \tilde{\omega}_j(T^n)^2)^{\frac{2}{p-1}-\frac{d}{2}} \right) \|\tilde{\varphi}(x)\|_2^2, \\ &= \gamma_j \left(-\tilde{\omega}_j(t) \left(m - \tilde{\omega}_j(t)^2\right)^{\frac{2}{p-1}-\frac{d}{2}} + \tilde{\omega}_j \left(m - \tilde{\omega}_j^2\right)^{\frac{2}{p-1}-\frac{d}{2}} \right) \|\tilde{\varphi}(x)\|_2^2, \end{aligned}$$

where the last inequality is due to the fact that $\tilde{\omega}_j(T^n) = \omega_j$. By simple Taylor expansion in frequencies we conclude

$$\begin{aligned} &Q(\widetilde{R}_j(t)) - Q(\widetilde{R}_j(T_n)) \\ &= \gamma_j \left[-(m - \omega_j^2)^{\frac{2}{p-1}-\frac{d}{2}} + 2\omega_j^2 (\frac{2}{p-1} - \frac{d}{2})(m - \omega_j^2)^{\frac{2}{p-1}-\frac{d}{2}-1} \right] (\tilde{\omega}_j(t) - \omega_j) \|\tilde{\varphi}(x)\|_2^2 \\ &\quad + o(\tilde{\omega}_j(t) - \omega_j). \end{aligned} \tag{61}$$

Since ω_j is part of the set $\mathcal{O}_{\text{stab}}$ (see (1)) we have

$$-(m - \omega_j^2)^{\frac{2}{p-1}-\frac{d}{2}} + 2\omega_j^2 \left(\frac{2}{p-1} - \frac{d}{2} \right) (m - \omega_j^2)^{\frac{2}{p-1}-\frac{d}{2}-1} > 0.$$

Combining the bootstrap assumption (50), and (59)–(61) gives the desired result. □

Proof of Lemma 14. The first step consists in splitting the action using $U = \sum_{j=1}^{N} \widetilde{R}_j + \Upsilon$. We start with the energy part. We have

$$\sum_{j=1}^{N} E_j(t, U) = E(U) = E(\widetilde{R} + \Upsilon) = E(\widetilde{R}) + E'(\widetilde{R})\Upsilon + \frac{1}{2}\langle E''(\widetilde{R})\Upsilon, \Upsilon \rangle + o(\|\Upsilon\|_{H^1 \times L^2}^2).$$

We treat the 0 order term first. By Lemma 18, we have

$$\begin{aligned} E(\widetilde{R}) &= \frac{1}{2} \left\| \nabla \left(\sum_{j=1}^{N} \widetilde{R}_{j,1} \right) \right\|_2^2 + \frac{m}{2} \left\| \sum_{j=1}^{N} \widetilde{R}_{j,1} \right\|_2^2 + \frac{1}{2} \left\| \sum_{j=1}^{N} \widetilde{R}_{j,2} \right\|_2^2 - \frac{1}{p+1} \left\| \sum_{j=1}^{N} \widetilde{R}_{j,1} \right\|_{p+1}^{p+1} \\ &= \sum_{j=1}^{N} \left(\frac{1}{2} \|\nabla \widetilde{R}_{j,1}\|_2^2 + \frac{m}{2} \|\widetilde{R}_{j,1}\|_2^2 + \frac{1}{2} \|\widetilde{R}_{j,2}\|_2^2 - \frac{1}{p+1} \|\widetilde{R}_{j,1}\|_{p+1}^{p+1} \right) + O(e^{-3\alpha\sqrt{m-\tilde{\omega}_\star^2} v_\star t}) \end{aligned}$$

where $\tilde{\omega}_\star = \max\{|\tilde{\omega}_j|; j = 1, \ldots, N\}$. In short, we have

$$E(\widetilde{R}) = \sum_{j=1}^{N} E(\widetilde{R}_j) + O(e^{-3\alpha\sqrt{m-\tilde{\omega}_\star^2}v_\star t}).$$

Now notice that

$$e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t} - e^{-3\alpha\sqrt{m-\tilde{\omega}_\star^2}v_\star t} = \frac{-\omega_\star e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}(3\alpha v_\star t)}{\sqrt{m-\omega_\star^2}}(\tilde{\omega}_\star - \omega_\star) + o(|\tilde{\omega}_\star - \omega_\star|),$$

such that, thanks to Lemma 16, we get

$$E(\widetilde{R}) = \sum_{j=1}^{N} E(\widetilde{R}_j) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

Using similar arguments we have

$$E'(\widetilde{R})\Upsilon = \sum_{j=1}^{N} E'(\widetilde{R}_j)\Upsilon + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}),$$

$$\langle E''(\widetilde{R})\Upsilon, \Upsilon\rangle = \sum_{j=1}^{N} \langle E''(\widetilde{R}_j)\Upsilon, \Upsilon\rangle + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

The proof follows the same steps for the localized charges and momenta: we have

$$Q_j(U) = \sum_{j=1}^{N} \left(Q(\widetilde{R}_j) + Q'(\widetilde{R}_j)\Upsilon + \frac{1}{2}\langle Q_j''(\widetilde{R}_j)\Upsilon, \Upsilon\rangle \right) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}),$$

$$P_j(U) = \sum_{j=1}^{N} \left(P(\widetilde{R}_j) + P'(\widetilde{R}_j)\Upsilon + \frac{1}{2}\langle P_j''(\widetilde{R}_j)\Upsilon, \Upsilon\rangle \right) + O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

The second step consists in expanding $\tilde{\omega}_j$ around ω_j using Lemma 16. Remembering that $\widetilde{R}_j$ is a critical point of $E + \frac{\tilde{\omega}_j}{\gamma_j}Q + v_j \cdot P$, we infer

$$E'(\widetilde{R}_j) + \frac{\omega_j}{\gamma_j}Q'(\widetilde{R}_j) + v_j \cdot P'(\widetilde{R}_j) = \frac{\omega_j - \tilde{\omega}_j}{\gamma_j}Q'(\widetilde{R}_j).$$

From Lemma 16, (50) and (54), it follows that

$$\left| \frac{\omega_j - \tilde{\omega}_j}{\gamma_j} Q'(\widetilde{R}_j)\Upsilon \right| \leq O(e^{-2\alpha\sqrt{m-\omega_\star^2}v_\star t})\|\Upsilon\|_{H^1\times L^2} \leq O(e^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

The only thing left to see is to remove the tildes corresponding to modulation. We have

$$\sum_{j=1}^{N} \left(E(\widetilde{R}_j) + \frac{\omega_j}{\gamma_j}Q(\widetilde{R}_j) + v_j \cdot P_j(\widetilde{R}_j) \right)$$
$$= \sum_{j=1}^{N} \left(E(R_j) + \frac{\omega_j}{\gamma_j}Q(R_j) + v_j \cdot P_j(R_j) + O((\tilde{\omega}_j - \omega_j)^2) \right),$$

where we have used the fact that

$$(\tilde{\omega}_j - \omega_j)(E' + \frac{\omega_j}{\gamma_j}Q' + v_j \cdot P')(R_j)\frac{\partial R_j}{\partial \omega} = 0.$$

Thanks to Lemma 16 we have

$$\sum_{j=1}^{N} |\tilde{\omega}_j - \omega_j|^2 \leq O(e^{-4\alpha\sqrt{m-\omega_\star^2}v_\star t}).$$

Gathering all these informations we get the desired result. □

A Appendix

Lemma 17 (Rellich-Kondrachov in H^s). *Let Ω be a bounded open set, $s \geq 0$ and $\varepsilon > 0$, and $u_n \in H^s(\mathbb{R}^d)$ be a bounded sequence such that supp $u_n \subset \Omega$. Then there exists $u \in H^s$ such that $\|u_n - u\|_{H^{s-\varepsilon}} = o(1)$.*

Proof. Let u_n be a bounded sequence in $H^s(\Omega)$ weakly converging to $u \in H^s$, we shall prove that, up to subsequences, $\|u_n - u\|_{H^{s-\varepsilon}} = o(1)$. By Plancherel identity we have

$$\|u_n - u\|^2_{H^{s-\varepsilon}} = \int_{|\xi|\leq R} (1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 d\xi + \int_{|\xi|>R} (1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 d\xi.$$

We have

$$\int_{|\xi|>R} (1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 d\xi \leq \frac{1}{(1+R^2)^{\varepsilon}} \int (1+|\xi|^2)^{s}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 d\xi \leq \frac{2}{(1+R^2)^{\varepsilon}} \|u_n\|^2_{H^s}.$$

In addition Ω is bounded and by weak convergence we have $\hat{u}_n(\xi) \to \hat{u}(\xi)$. To conclude it suffices to show that

$$\int_{|\xi|\leq R} (1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 d\xi = o(1). \tag{62}$$

Notice that

$$\|\hat{u}_n\|_{L^\infty(\Omega)} \leq \|u_n\|_{L^1(\Omega)} \leq \mu(\Omega)^{\frac{1}{2}}\|u_n\|_{L^2(\Omega)} \leq \mu(\Omega)^{\frac{1}{2}}\|u_n\|_{H^s}$$

and hence $(1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2$ is dominated by $C(1+|R|^2)^{s-\varepsilon}$ such that (62) holds.

Now, fix $\delta > 0$ and choose $R > 0$ and N sufficiently large such that

$$\int_{|\xi|>R} (1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 d\xi \leq \frac{\delta}{2},$$

and for all $n \geq N$

$$\int_{|\xi|\leq R}(1+|\xi|^2)^{s-\varepsilon}|\hat{u}_n(\xi)-\hat{u}(\xi)|^2 d\xi \leq \frac{\delta}{2},$$

i.e., $\|u_n - u\|_{H^{s-\varepsilon}} \leq \delta$. □

Lemma 18 (Interactions estimates). *There exists $f \in L_t^\infty L_x^1(\mathbb{R}, \mathbb{R}^d) \cap L_t^\infty L_x^\infty(\mathbb{R}, \mathbb{R}^d)$ such that if $j \neq k$*

$$|R_j R_k| + |R_j \nabla R_k| + |\nabla R_j \nabla R_k| + |R_j|\phi_k + |\nabla R_j|\phi_k \leq Ce^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t} f(t,x)$$

$$\left| |R|^{p+1} - \sum_{l=1}^{N} |R_l|^{p+1} \right| + \left| |R|^{p-1}R - \sum_{l=1}^{N} |R_l|^{p-1}R_l \right| + \left| |R|^{p-1} - \sum_{l=1}^{N} |R_l|^{p-1} \right|$$
$$\leq Ce^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t} f(t,x).$$

Proof. We start proving that there exists $f \in L_t^\infty L_x^1(\mathbb{R}, \mathbb{R}^d) \cap L_t^\infty L_x^\infty(\mathbb{R}, \mathbb{R}^d)$ such that if $j \neq k$

$$|R_j R_k| \leq Ce^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t} f(t,x).$$

Thanks to (6) (the Lorenz transform gives indeed only a contraction along the direction of propagation) we know that

$$|R_j| \leq Ce^{-\frac{1}{2}\sqrt{m-\omega_j^2}|x-v_j t|} \leq Ce^{-\frac{1}{2}\sqrt{m-\omega_\star^2}|x-v_j t|}$$
$$|R_k| \leq Ce^{-\frac{1}{4}\sqrt{m-\omega_k^2}|x-v_j t|} \leq Ce^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x-v_k t|}.$$

By a simple change of variable we get

$$|R_j||R_k| \leq Ce^{-\frac{1}{2}\sqrt{m-\omega_\star^2}|x|} e^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x-(v_k-v_i)t)|},$$

such that, thanks to the following inequality

$$|x-(v_k-v_j)t| \geq |(v_k-v_j)t| - |x| \geq v_\star t - |x|$$

we conclude

$$|R_j||R_k| \leq Ce^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x|} e^{-\frac{1}{4}\sqrt{m-\omega_\star^2}v_\star t}.$$

Taking $3\alpha \leq \frac{1}{4}$ we get the desired estimate. The estimates for $|R_j \nabla R_k|$ and $|\nabla R_j \nabla R_k|$ follow analogously.

Now we shall prove that if $j \neq k$

$$|R_j|\phi_k \leq Ce^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t} f(t,x).$$

Let us suppose without any lack of generality that $j < k-1$. Notice that

$$|R_j|\phi_k \leq C^{-\frac{1}{2}\sqrt{m-\omega_\star^2}|x-v_j t|}\chi_{[\frac{1}{2}(v_{k-1}+v_k)t-\sqrt{t},\, \frac{1}{2}(v_{k+1}+v_k)t+\sqrt{t}]}$$

that implies

$$|R_j|\phi_k \le Ce^{-\frac{1}{2}\sqrt{m-\omega_\star^2}|x-v_j t|}\chi_{[(v_j+v_\star)t-\sqrt{t},v_k t+\sqrt{t}]}.$$

By a simple change of variable we get

$$|R_j|\phi_k \le Ce^{-\frac{1}{2}\sqrt{m-\omega_\star^2}|x|}\chi_{[v_\star t-\sqrt{t},(v_k-v_j)t+\sqrt{t}]}.$$

Now, for $t \ge \max\{\frac{4}{v_\star^2}, 1\}$, it follows

$$|R_j|\phi_k \le Ce^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x|}e^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x|}\chi_{[\frac{1}{2}v_\star t,(v_k-v_j+1)t]} \le Ce^{-\frac{1}{8}\sqrt{m-\omega_\star^2}v_\star t}e^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x|}. \quad (63)$$

Now for $\alpha < \frac{1}{24}$ we conclude

$$|R_j|\phi_k \le Ce^{-3\alpha\sqrt{m-\omega_\star^2}v_\star t}e^{-\frac{1}{4}\sqrt{m-\omega_\star^2}|x|}.$$

The case $j \ge k-1$, $j \ne k$, follows identically as well as the estimates concerning the gradient. The second part of the lemma follows from the inequality

$$(|a+b|^p - |a|^p - |b|^p) \le C(|a||b|^{p-1} + |a|^{p-1}|b|) \quad \text{with } p > 0$$

that derives from the elementary inequality

$$(|1+t|^p - 1 - |t|^p) \le C(|t| + |t|^{p-1}) \quad \text{with } p > 0.$$

By arguing as before we get the desired estimates. □

Funding

J.B is supported by FIRB2012 "Dinamiche dispersive: analisi di Fourier e metodi variazionali" and PRIN2009 "Metodi Variazionali e Topologici nello Studio di Fenomeni non Lineari", M.G by the PRIN2009 grant "Critical Point Theory and Perturbative Methods for Nonlinear Differential Equations", S.L.C by the French ANR project ESONSE.

References

[1] Ambrosetti, A., Malchiodi, A. (2007). *Nonlinear Analysis and Semilinear Elliptic Problems*. Cambridge: Cambridge University Press.

[2] Bellazzini, J., Visciglia, N. (2010). Max-min characterization of the mountain pass energy level for a class of variational problems. *Proc. Amer. Math. Soc.* 138:3335–3343.

[3] Berestycki, H., Lions, P.-L. (1983). Nonlinear scalar field equations I. *Arch. Ration. Mech. Anal.* 82:313–346.

[4] Berestycki, H., Lions, P.-L. (1983). Nonlinear scalar field equations II. *Arch. Ration. Mech. Anal.* 82:347–375.

[5] Chang, S.-M., Gustafson, S., Nakanishi, K., Tsai, T.-P. (2007). Spectra of linearized operators for NLS solitary waves. *SIAM J. Math. Anal.* 39:1070–1111.

[6] Combet, V. (2011). Multi-soliton solutions for the supercritical gKdV equations. *Comm. Part. Diff. Eqs.* 36:380–419.

[7] Côte, R., Le Coz, S. (2011). High-speed excited multi-solitons in nonlinear Schrödinger equations. *J. Math. Pures Appl. (9)* 96:135–166.

[8] Côte, R., Martel, Y., Merle, F. (2011). Construction of multi-soliton solutions for the L^2-supercritical gKdV and NLS equations. *Rev. Mat. Iberoam.* 27:273–302.

[9] Côte, R., Muñoz, C. (2014). Multi-solitons for nonlinear Klein-Gordon equations. *Forum of Mathematics Sigma* 2:e15.

[10] Côte, R., Zaag, H. (2013). Construction of a multi-soliton blow-up solution to the semilinear wave equation in one space dimension. *Comm. Pure Appl. Math.* 66:1541–1581.

[11] Gidas, B., Ni, W.M., Nirenberg, L. (1979). Symmetry and related properties via the maximum principle. *Comm. Math. Phys.* 68:209–243.

[12] Ginibre, J., Velo, G. (1985). The global Cauchy problem for the nonlinear Klein-Gordon equation. *Math. Z.* 189:487–505.

[13] Ginibre, J., Velo, G. (1989). The global Cauchy problem for the nonlinear Klein-Gordon equation. II. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 6:15–35.

[14] Grillakis, M., Shatah, J., Strauss, W.A. (1987). Stability theory of solitary waves in the presence of symmetry I. *J. Funct. Anal.* 74:160–197.

[15] Grillakis, M., Shatah, J., Strauss, W.A. (1990). Stability theory of solitary waves in the presence of symmetry. II. *J. Func. Anal.* 94:308–348.

[16] Jeanjean, L., Le Coz, S. (2009). Instability for standing waves of nonlinear Klein-Gordon equations via mountain-pass arguments. *Trans. Amer. Math. Soc.* 361:5401–5416.

[17] Jeanjean, L., Tanaka, K. (2003). A note on a mountain pass characterization of least energy solutions. *Adv. Nonlinear Stud.* 3:445–455.

[18] Jeanjean, L., Tanaka, K. (2003). A remark on least energy solutions in $\mathbb{R}^N$. *Proc. Amer. Math. Soc.* 131:2399–2408.

[19] Killip, R., Stovall, B., Visan, M. (2014). Blowup behaviour for the nonlinear Klein–Gordon equation. *Math Ann.* 358:289–350.

[20] Kwong, M.K. (1989). Uniqueness of positive solutions of $\Delta u - u + u^p = 0$ in R^n. *Arch. Rational Mech. Anal.* 105:243–266.

[21] Le Coz, S. (2009). Standing waves in nonlinear Schrödinger equations. In: Emmrich, E., Wittbold, P., eds., *Analytical and Numerical Aspects of Partial Differential Equations*. Berlin: Walter de Gruyter, pp. 151–192.

[22] Le Coz, S., Li, D., Tsai, T.-P. Fast-moving finite and infinite trains of solitons for nonlinear Schrödinger equations. Preprint, arXiv:1304.3049.

[23] Le Coz, S., Tsai, T.-P. Infinite soliton and kink-soliton trains for nonlinear Schrödinger equations. Preprint, arXiv:1309.7846.

[24] Lindblad, H., Sogge, C.D. (1995). On existence and scattering with minimal regularity for semilinear wave equations. *J. Funct. Anal.* 130:357–426.

[25] Lions, P.-L. (1984). The concentration-compactness principle in the calculus of variations. The locally compact case. I. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 1:109–145.

[26] Liu, Y., Ohta, M., Todorova, G. (2007). Instabilité forte d'ondes solitaires pour des équations de Klein-Gordon non linéaires et des équations généralisées de Boussinesq [Strong instability of solitary waves for nonlinear Klein-Gordon equations and generalized Boussinesq equations]. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 24:539–548.

[27] Martel, Y. (2005). Asymptotic N-soliton-like solutions of the subcritical and critical generalized Korteweg-de Vries equations. *Amer. J. Math.* 127:1103–1140.
[28] Martel, Y., Merle, F. (2001). Instability of solitons for the critical generalized Korteweg-de Vries equation. *Geom. Funct. Anal.* 11:74–123.
[29] Martel, Y., Merle, F. (2006). Multi solitary waves for nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 23:849–864.
[30] Martel, Y., Merle, F. (2011). Description of two soliton collision for the quartic gKdV equation. *Ann. of Math. (2)* 174:757–857.
[31] Martel, Y., Merle, F. (2011). Inelastic interaction of nearly equal solitons for the quartic gKdV equation. *Invent. Math.* 183:563–648.
[32] Martel, Y., Merle, F. (2011c). Review of long time asymptotics and collision of solitons for the quartic generalized Korteweg-de Vries equation. *Proc. Roy. Soc. Edinburgh Sect. A* 141:287–317.
[33] Martel, Y., Merle, F., Tsai, T.-P. (2002). Stability and asymptotic stability in the energy space of the sum of N solitons for subcritical gKdV equations. *Comm. Math. Phys.* 231:347–373.
[34] Martel, Y., Merle, F., Tsai, T.-P. (2006). Stability in H^1 of the sum of K solitary waves for some nonlinear Schrödinger equations. *Duke Math. J.* 133:405–466.
[35] Merle, F. (1990). Construction of solutions with exactly k blow-up points for the Schrödinger equation with critical nonlinearity. *Comm. Math. Phys.* 129:223–240.
[36] Miura, R.M. (1976). The Korteweg-de Vries equation: A survey of results. *SIAM Rev.* 18:412–459.
[37] Nakamura, M., Ozawa, T. (2001). The Cauchy problem for nonlinear Klein-Gordon equations in the Sobolev spaces. *Publ. Res. Inst. Math. Sci.* 37:255–293.
[38] Nakanishi, K., Schlag, W. (2011). Global dynamics above the ground state energy for the focusing nonlinear Klein-Gordon equation. *J. Diff. Eqs.* 250:2299–2333.
[39] Nakanishi, K., Schlag, W. (2011). *Invariant Manifolds and Dispersive Hamiltonian Evolution Equations*. Zurich Lectures in Advanced Mathematics. Zürich: European Mathematical Society.
[40] Nakanishi, K., Schlag, W. (2012). Global dynamics above the ground state for the nonlinear Klein-Gordon equation without a radial assumption. *Arch. Ration. Mech. Anal.* 203:809–851.
[41] Nakanishi, K., Schlag, W. (2012). Invariant manifolds around soliton manifolds for the nonlinear Klein-Gordon equation. *SIAM J. Math. Anal.* 44:1175–1210.
[42] Ohta, M., Todorova, G. (2005). Strong instability of standing waves for nonlinear Klein-Gordon equations. *Discrete Contin. Dyn. Syst.* 12:315–322.
[43] Ohta, M., Todorova, G. (2007). Strong instability of standing waves for the nonlinear Klein-Gordon equation and the Klein-Gordon-Zakharov system. *SIAM J. Math. Anal.* 38:1912–1931 (electronic).
[44] Perelman, G. (1997). Some results on the scattering of weakly interacting solitons for nonlinear Schrödinger equations. In: *Spectral Theory, Microlocal Analysis, Singular Manifolds*. Mathematical Topics, Vol. 14. Berlin: Akademie Verlag, pp. 78–137.
[45] Perelman, G. (2004). Asymptotic stability of multi-soliton solutions for nonlinear Schrödinger equations. *Comm. Part. Diff. Eqs.* 29:1051–1095.

[46] Rodnianski, I., Schlag, W., Soffer, A. (2003). Asymptotic stability of n-soliton states of nonlinear Schrödinger equations. Unpublished manuscript.
[47] Shatah, J. (1983). Stable standing waves of nonlinear Klein-Gordon equations. *Comm. Math. Phys.* 91:313–327.
[48] Shatah, J. (1985). Unstable ground state of nonlinear Klein-Gordon equations. *Trans. Amer. Math. Soc.* 290:701–710.
[49] Shatah, J., Strauss, W.A. (1985). Instability of nonlinear bound states. *Comm. Math. Phys.* 100:173–190.
[50] Sigal, I.M. (1982). Geometric methods in the quantum many-body problem. Nonexistence of very negative ions. *Comm. Math. Phys.* 85:309–324.
[51] Stuart, D.M.A. (2001). Modulational approach to stability of non-topological solitons in semilinear wave equations. *J. Math. Pures Appl. (9)* 80:51–83.
[52] Weinstein, M.I. (1985). Modulational stability of ground states of nonlinear Schrödinger equations. *SIAM J. Math. Anal.* 16:472–491.
[53] Weinstein, M.I. (1986). Lyapunov stability of ground states of nonlinear dispersive evolution equations. *Comm. Pure Appl. Math.* 39:51–67.
[54] Willem, M. (1996). Minimax theorems. Progress in Nonlinear Differential Equations and Their Applications, Vol. 24. Boston: Birkhäuser Boston Inc.
[55] Zakharov, V.E., Shabat, A.B. (1972). Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Šoviet Physics JETP* 34:62–69.