# A DIRECT CONSTRUCTION OF A SLOW MANIFOLD FOR A SEMILINEAR WAVE EQUATION OF KLEIN–GORDON TYPE

HAIDAR MOHAMAD AND MARCEL OLIVER

Abstract. We study a semilinear wave equation whose linear part corresponds to the linear Klein–Gordon equation in the non-relativistic limit, augmented with a nonlinearity that is Fréchet-differentiable over the complex numbers. We show that this equation possesses an almost invariant manifold in phase space that generalizes the slow manifold which is known to exist for finite-dimensional Galerkin truncations of the system. This manifold is shown to be almost invariant to any algebraic order and can be constructed in the $H^{s-1} \times H^s$ phase space of the equation uniformly in the order of the approximation. In particular, we prove that the dynamics on this "slow manifold" shadows orbits of the full system over a finite interval of time.

## 1. Introduction

We study the semilinear wave equation

$$\varepsilon\, \partial_t^2 u - \mathrm{i}\, \partial_t u - \Delta u = g(u) \tag{1}$$

in the limit of small $\varepsilon$, where $u \colon [0,T] \times \mathbb{T} \to \mathbb{C}$, $\Delta u = \partial_x^2 u$ denotes the Laplacian, and $g$ is a Fréchet-differentiable function on $H^s(\mathbb{T})$ for some $s \in \mathbb{R}$. Written as a first order system of evolution equations

$$\partial_t u = v\,, \tag{2a}$$

$$\varepsilon\, \partial_t v = \mathrm{i}\, v + \Delta u + g(u)\,, \tag{2b}$$

the problem has a formal resemblance to the finite-dimensional two-scale fast-slow system

$$\dot q = p\,, \tag{3a}$$

$$\varepsilon\, \dot p = Jp - \nabla V(q)\,, \tag{3b}$$

where $q \colon [0,T] \to \mathbb{R}^{2d}$ is the vector of positions, $p$ is the vector of corresponding momenta, $J$ is a symplectic matrix, and $V$ is a smooth potential. In finite dimensions, there is an asymptotic separation of time scales: the linearized dynamics has $d$ "slow" eigenvalues of magnitude $O(1)$ and $d$ "fast" eigenvalues of $O(\varepsilon^{-1})$. It is straightforward to construct an almost invariant slow manifold to $O(\varepsilon^N)$ and it is well known how to achieve invariance up to exponentially small terms [4, 12, 16]. Similar results are possible with an infinite number of fast degrees of freedom [9]. In our context, however, Nekhoroshev theory is not applicable due to the lack of asymptotic separation of scales. Nonetheless, we are able to show that some remnant of typical conclusions of Nekhoroshev theory continues to hold true: We construct, to an arbitrary order $N$, a manifold in phase space that is invariant over $O(1)$-times.

To do so, we revisit the explicit construction of an $O(\varepsilon^{N+1})$ slow manifold $p = F_{\text{slow}}^N(q)$ in the finite dimensional case. It is based on the ansatz

$$F_{\text{slow}}^N(q) = \sum_{k=0}^{N} f_k(q)\, \varepsilon^k \,, \tag{4}$$

where the coefficient vector fields $f_k$ are determined by the condition that the evolution equation for the fast residual component $w = p - F_{\text{slow}}^N(q)$ is of order $O(\varepsilon^{N+1})$. This will directly lead to a shadowing result of the form

$$\sup_{t \leq T} \|q(t) - Q(t)\| \leq c\, \varepsilon^{N+1} \,, \tag{5}$$

where $q$ solves (3) with prepared initial data $p_0 = F_{\text{slow}}(q_0)$ and $Q$ solves the slow equation

$$\dot{Q} = F_{\text{slow}}^N(Q) \tag{6}$$

with $Q(0) = q_0$.

The situation for (1) is different, even in the linear case. Since the Laplacian is unbounded, there is no spectral gap between the families, parameterized by $\varepsilon$, of eigenvalues that remain bounded and those that diverge as $\varepsilon \to 0$. On the level of the asymptotic construction, we observe that an iterative construction of the slow vector field as, for example, described in [7] will involve composition with the Laplacian, i.e., lead to a loss of two derivatives in Sobolev space per iteration. This loss of derivatives is observed in related perturbation problems, e.g. [14, 18].

In this paper, we show, focusing on the equation (1), that it is possible to construct an infinite-dimensional analog of the approximate slow manifold in the example above. It is the graph of a nonlinear mapping which is the sum of a linear operator from $H^s(\mathbb{T})$ into $H^{s-1}(\mathbb{T})$ and an iteratively constructed nonlinear map from $H^s(\mathbb{T})$ into itself. This "slow manifold" (i) is invariant up to terms of $O(\varepsilon^{N+1})$ in the $H^s(\mathbb{T})$ topology and (ii) can be seen as a *regular* perturbation of a nonlinear Schrödinger equation which is the leading order non-relativistic limit of equation (1). (For a justification of the leading order asymptotics of the semilinear Klein–Gordon equation, see [11, 17].)

To be concrete, we note that (2) with $g = 0$ is easily block-diagonalized. Its eigenoperators must satisfy the characteristic equation

$$\varepsilon\, L_\pm^2 - \mathrm{i}\, L_\pm - \Delta = 0 \,, \tag{7}$$

so that

$$L_\pm = \mathrm{i}\, \frac{1 \pm \sqrt{1 - 4\varepsilon\Delta}}{2\varepsilon} \,. \tag{8}$$

Clearly, there is no spectral gap between the two eigenoperators. However,

$$\lim_{\varepsilon \to 0} L_-(\varepsilon) \to \mathrm{i}\, \Delta \tag{9}$$

in the strong operator topology of $H^s$ whereas the sequence $L_+(\varepsilon)u$ diverges for any nonzero $u \in H^s$. In this sense, we shall speak of the subspace of phase space associated with $L_-$ as the *slow subspace* and the subspace associated with $L_+$ as the *fast subspace*.

In the general case, i.e., when $g \neq 0$, we seek an almost invariant submanifold in the phase space on which the dynamics is given by an evolution equation of the form

$$\partial_t U = L_- U + F_{\text{slow}}^N(U) \,, \tag{10}$$

the infinite dimensional analog of the slow equation (6). The vector field $F^N_{\text{slow}}$ is again sought in the form (4), where the coefficient vector fields are constructed iteratively in Section 2 below. The novelty of our construction is that each of the coefficient vector fields $f_k$, under suitable assumptions, maps $H^s$ into itself. We note that, even in finite dimensions, we cannot expect the existence of a truly invariant slow manifold [12]. Classical Nekhoroshev theory [16] provides "almost stability" of the manifold over exponentially long times. Here, the result is much weaker, namely valid for times of order one for the slow limit system.

Our result can be paraphrased as follows. *Suppose $g$ is sufficiently smooth and complex Fréchet-differentiable. Let $u_0 \in H^s(\mathbb{T})$ for $s$ sufficiently large but otherwise fixed. Then for every $N$ there exists a slow vector field $F^N_{\text{slow}}$ such that when $U$ solves the slow equation* (10) *and $u$ solves the full system* (2) *consistently initialized with $v(0) = L_- u_0 + F^N_{\text{slow}}(u_0)$, then there is a time $T > 0$ which only depends on the time of existence of the slow equation such that the full solution $u$ exists on the same interval of time and there is a constant $C > 0$ such that, as $\varepsilon \to 0$,*

$$\sup_{0 \le t \le T} \|u(t) - U(t)\|_s \le C\, \varepsilon^{N+1} \,. \tag{11}$$

A precise statement of the result is formulated as Theorem 7 in Section 5. We emphasize that the time interval of validity is physically meaningful (i.e., says something about the solution of the original PDE) in the sense that every fixed, finite interval of time can be used so long as the limit dynamics remains bounded, then $\varepsilon$ can be chosen as to make the shadowing error as small as one likes. Thus, the slow solution is visible as a finite-time model for suitably initialized full dynamics.

Equation (1) is related to the semilinear Klein–Gordon equation

$$\frac{\hbar^2}{2mc^2}\, \partial_t^2 \psi - \frac{\hbar^2}{2m}\, \Delta\psi + \frac{mc^2}{2}\, \psi = f(|\psi|^2)\psi \,, \tag{12}$$

where $\hbar$ is the Planck constant, $c$ is the speed of light, $m$ and is the mass of the particle. As in [17], we choose units in which $m$ and $\hbar$ take the value 1 and set $\varepsilon = 1/(2c^2)$, so that $\varepsilon \to 0$ corresponds to the non-relativistic limit where $c \to \infty$. Then, inserting the modulated rotating wave ansatz

$$\psi = u \, \exp(\mathrm{i} mc^2 t/\hbar) \tag{13}$$

and rescaling space such that, for convenience, the prefactor in front of the Laplace operator takes the value 1, we obtain

$$\varepsilon\, \partial_t^2 u - \mathrm{i}\, \partial_t u - \Delta u = f(|u|^2)\, u \,. \tag{14}$$

Note, however, that $u \mapsto |u|^2$ is not Fréchet-differentiable in vector spaces over $\mathbb{C}$ so that, in general, $u \mapsto f(|u|^2)\, u$ is not Fréchet-differentiable over $\mathbb{C}$. Thus, the class of nonlinearities $g$ we consider for equation (1) generally does not correspond to a classical Klein–Gordon nonlinearity as in (14). It is open whether our result extends in some form to nonlinearities that are only differentiable over $\mathbb{R}$. Our construction, however, depends crucially on Fréchet-differentiability over the complex numbers.

Let us mention some related results. Masmoudi and Nakanishi [13] studied the non-relativistic limit of the Klein–Gordon equation in the form (12). They make a solution ansatz in the form of a pair of left and right rotating waves,

$$\psi = u_-\, \mathrm{e}^{\mathrm{i}c^2 t} + u_+\, \mathrm{e}^{-\mathrm{i}c^2 t} + o(1) \,, \tag{15}$$

and prove that this decomposition holds true for general initial data in the energy space $H^1$, where $u_-$ and $u_+$ satisfy a pair of coupled nonlinear Schrödinger equations with a certain averaged potential. Lu and Zhang [10] recently obtained a next-order correction in the same setting; a number of recent papers also use this ansatz as a starting point for developing asymptotics-preserving numerical schemes [1, 2, 3, 5, 6]. Thus, our result differs from these references in that we seek asymptotics to any order in a uniform functional setting at the expense that we are required to work with prepared initial data.

We have shown in an earlier paper [15] that when $g$ is a linear operator, the corresponding slow vector field can be constructed via the solution of an operator Sylvester equation. The construction here provides a different integral representation for the solution of the required order conditions, but leads to an equivalent statement in the linear case. Neither of the two expressions give rise to a straightforward numerical scheme. However, they enable estimates which we believe are an essential ingredient for proving a shadowing theorem for solutions of slow equations that arise from variational asymptotics, analogous to what is known in finite dimensions [7]. These variational slow equations are indeed computable with standard methods.

The remainder of the paper is structured as follows. In the next Section 2, we describe the iterative construction of the slow vector field and prove a result on the solvability of the order condition. The resulting estimates on the slow vector field are derived in Section 3. Section 4 discusses the Cauchy problems for the full system (1) and the slow equation (10). In the final Section 5, we prove the main result on the shadowing of orbits to the full equation by the slow dynamics.

## 2. The iterative construction

Our construction modifies the naive approach in two respects: we are using the *exact* splitting (8) into fast and slow invariant subspaces in the linear part of the equation, and we move certain next order terms into the order condition of the previous order.

To begin, we introduce a "fast" variable

$$w = \partial_t u - L_- u - F_{\text{slow}}^{N+1}(u) \,, \tag{16}$$

where $u$ solves (1) and $F_{\text{slow}}^{N+1}(u)$ is sought as an expansion in powers of $\varepsilon$ in the form (4). Differentiating (16) and noting that

$$\frac{\mathrm{i}}{\varepsilon} - L_- = L_+ \,, \tag{17}$$

we obtain equation (1) in $u$-$w$-variables:

$$\partial_t u = L_- u + F_{\text{slow}}^{N}(u) + \varepsilon^{N+1} \, f_{N+1}(u) + w \,, \tag{18a}$$

$$\begin{aligned} \partial_t w = L_+ w - DF_{\text{slow}}^{N+1}(u) w + L_+ F_{\text{slow}}^{N+1}(u) + \frac{1}{\varepsilon} \, g(u) \\ - DF_{\text{slow}}^{N+1}(u) L_- u - DF_{\text{slow}}^{N+1}(u) F_{\text{slow}}^{N+1}(u) \,. \end{aligned} \tag{18b}$$

We now seek coefficient vector fields $f_k(u)$ such that all but the first two terms on the right hand side of (18b) are eliminated up to an $\mathcal{O}(\varepsilon^{N+1})-$remainder. However, as we shall see, a naive approach will lead to loss of derivatives. Indeed, noting that $L_+$ is an operator of formal order $\mathcal{O}(\varepsilon^{-1})$, we would seek order conditions

$$\varepsilon\, L_+ f_0(u) + g(u) = 0\,, \tag{19a}$$

$$\varepsilon\, L_+ f_{k+1}(u) - Df_k(u) L_- u + \sum_{j+l=k} Df_j(u) f_l(u) = 0\,. \tag{19b}$$

Setting $K = (\varepsilon L_+)^{-1}$, we obtain the recursion

$$f_0(u) = -K g(u)\,, \tag{20a}$$

$$f_{k+1}(u) = K(Df_k(u)L_- u) - K \sum_{j+l=k} Df_j(u) f_l(u)\,. \tag{20b}$$

We note that $K$ is a positive, self-adjoint, compact operator. However, it is not uniformly compact in $\varepsilon$. In fact, it is easy to verify that, for $v \in H^s$,

$$\|L_- v\|_{s-1} \le \frac{1}{\sqrt{\varepsilon}}\, \|v\|_s\,, \tag{21a}$$

$$\|L_- v\|_{s-2} \le \|v\|_s\,, \tag{21b}$$

$$\|Kv\|_s \le \|v\|_s\,, \tag{21c}$$

$$\|Kv\|_{s+1} \le \frac{1}{\sqrt{\varepsilon}}\, \|v\|_s\,. \tag{21d}$$

Thus, the estimation of the first term on the right hand side of (20b) will lead to either loss of derivatives or loss of order. In particular, when estimating $u$ and $w$ in the same space, order is lost completely and (20) ceases to define a well-ordered asymptotic series.

To remedy this problem, we seek to include the "bad" term $Df_k(u)L_- u$ into the computation of the coefficient vector field at the same order $k$. This leads to the alternative order conditions

$$\varepsilon\, L_+ f_0(u) - \varepsilon\, Df_0(u) L_- u + g(u) = 0\,, \tag{22a}$$

$$\varepsilon\, L_+ f_{k+1}(u) - \varepsilon\, Df_{k+1}(u) L_- u + \sum_{j+l=k} Df_j(u) f_l(u) = 0\,. \tag{22b}$$

Thus, finding the $f_k$ requires us to solve a differential equation of the form

$$Df(u)Au + Bf(u) + g(u) = 0\,, \tag{23}$$

where $A$ and $B$ are two generally unbounded linear operators.

*Definition* 1. Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. We write $\mathcal{C}^n_{\mathrm{b}}(\mathcal{X},\mathcal{Y})$ to denote the space of functions from $\mathcal{X}$ to $\mathcal{Y}$ that are $n$ times continuously differentiable and such that the $n$-th derivative maps bounded subsets of $\mathcal{X}$ into bounded subsets of $\mathcal{L}^n(\mathcal{X},\mathcal{Y})$, the space of bounded $n$-linear forms from $\mathcal{X}^n$ to $\mathcal{Y}$. We abbreviate $\mathcal{C}^n_{\mathrm{b}}(\mathcal{X}) \equiv \mathcal{C}^n_{\mathrm{b}}(\mathcal{X},\mathcal{X})$.

Note that when $f \in \mathcal{C}^1_{\mathrm{b}}(\mathcal{X},\mathcal{Y})$, we have

$$f(u) - f(v) = \int_0^1 Df(u + t(v-u))(v-u)\,\mathrm{d}t\,. \tag{24}$$

Thus, $f \in \mathcal{C}^0_{\mathrm{b}}(\mathcal{X},\mathcal{Y})$ and, moreover, $f$ is Lipshitz with a constant that is uniform on bounded subsets of $\mathcal{X}$.

Solutions to (23) are provided by the following theorem.

**Theorem 2.** *Let $\mathcal{H}$ be a complex Hilbert space, $g\colon \mathcal{H} \to \mathcal{H}$ a Fréchet-differentiable function, and $A$ and $B$ two closed, densely defined normal operators having the same domain $\mathcal{D} = D(A) = D(B) \subset \mathcal{H}$. Assume the following.*

(i) *$g(\mathcal{D}) \subset \mathcal{D}$, $g \in \mathcal{C}^0_{\mathrm{b}}((\mathcal{D}, \|\cdot\|_B))$, and $Dg \in \mathcal{C}^0_{\mathrm{b}}(\mathcal{H}, \mathcal{L}(\mathcal{H}))$, where $\|\cdot\|_B = \|\cdot\| + \|B\cdot\|$ denotes the graph norm on $\mathcal{D}$.*

(ii) *The spectrum of $A$ is contained in the closed left complex half-plane.*

(iii) *The spectrum of $B$ is contained in the left complex half-plane with $\delta = \operatorname{dist}(\sigma(B), \mathrm{i}\mathbb{R}) > 0$.*

(iv) *$\mathrm{e}^{\tau B}$ and $\mathrm{e}^{\tau A}$ map $\mathcal{D}$ into itself and $\mathrm{e}^{tA}B = B\mathrm{e}^{tA}$ on $\mathcal{D}$ for all $t > 0$.*

*Then the function $f\colon \mathcal{D} \to \mathcal{D}$ defined by*

$$f(u) = \int_0^\infty \mathrm{e}^{\tau B}\, g(\mathrm{e}^{\tau A} u)\, \mathrm{d}\tau \tag{25}$$

*solves the differential equation*

$$Df(u)Au + Bf(u) + g(u) = 0 \tag{26}$$

*for $u \in \mathcal{D}$.*

*Proof.* Note first that the integral defining $f$ is well defined on $(\mathcal{D}, \|\cdot\|_B)$. Indeed, due to (iv), $\|\mathrm{e}^{tA}v\|_B = \|\mathrm{e}^{tA}v\| + \|\mathrm{e}^{tA}Bv\| \leq \|\mathrm{e}^{tA}\|_{\mathcal{L}(\mathcal{H})}\|v\|_B$ so that $\mathrm{e}^{\tau A}$ is uniformly bounded for $\tau \geq 0$ on $(\mathcal{D}, \|\cdot\|_B)$. Then, as $g \in \mathcal{C}^0_{\mathrm{b}}(\mathcal{D})$, for every $u \in \mathcal{D}$ there exists $M = M(u) > 0$ such that

$$\|g(\mathrm{e}^{\tau A}u)\|_B \leq M \tag{27}$$

and therefore

$$\|\mathrm{e}^{\tau B} g(\mathrm{e}^{\tau A}u)\|_B \leq \|\mathrm{e}^{\tau B}\|_{\mathcal{L}(\mathcal{D})}\, \|g(\mathrm{e}^{\tau A}u)\|_B \leq M \exp(-\tau\delta)\,. \tag{28}$$

This implies that (25) is well-defined as a Bochner integral and $f(\mathcal{D}) \subset \mathcal{D}$. Since $B\colon D(B) \to \mathcal{H}$ is closed, we can write

$$Bf(u) = \int_0^\infty B\mathrm{e}^{\tau B} g(\mathrm{e}^{\tau A}u)\, \mathrm{d}\tau\,. \tag{29}$$

For $\tau > 0$, write

$$F(\tau) = \mathrm{e}^{\tau B} g(\mathrm{e}^{\tau A}u)\,. \tag{30}$$

Hence,

$$\begin{aligned} Df(u)Au + Bf(u) &= \int_0^\infty \bigl(\mathrm{e}^{\tau B} Dg(\mathrm{e}^{\tau A}u)\mathrm{e}^{\tau A}Au + B\mathrm{e}^{\tau B} g(\mathrm{e}^{\tau A}u)\bigr)\, \mathrm{d}\tau \\ &= \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}\tau} F(\tau)\, \mathrm{d}\tau \\ &= \lim_{t\to\infty} F(t) - F(0) = -g(u)\,. \end{aligned} \tag{31}$$

Differentiation under the integral can be justified by noting that the Fréchet-derivative $D$ can be cast as a closed operator which can be moved inside of a Bochner integral by a lemma of Hille [8]. □

We apply Theorem 2 recursively with $A = -\varepsilon \mathrm{i} L_-$, $B = \varepsilon \mathrm{i} L_+$, and $g \equiv g_0$ initially to obtain the sequence of vector fields

$$f_k = \mathrm{i} \int_0^\infty \mathrm{e}^{\tau B} g_k(\mathrm{e}^{\tau A}u)\, \mathrm{d}\tau \tag{32a}$$

and

$$g_{k+1} = \sum_{j+m=k} Df_j f_m \tag{32b}$$

which solve the order conditions (22).

## 3. Functional properties of the iterative construction

**Lemma 3.** *Let $N \in \mathbb{N}$, $s \in \mathbb{R}$, and $g \in \mathcal{C}_{\mathrm{b}}^N(H^s)$. Then $f_k$ is well-defined by* (32a) *on $H^s$. Further, for all $k \in \{0, \dots, N\}$,*

$$g_k, f_k \in \mathcal{C}_{\mathrm{b}}^{N-k}(H^s)\,. \tag{33}$$

*If, moreover, $g \in \mathcal{C}_{\mathrm{b}}^{N+1}(H^{s-1})$, the functions $f_k$, solve the order equations* (22) *for $k = 0, \dots, N$.*

*Proof.* We proceed inductively. Since $g \in \mathcal{C}_{\mathrm{b}}^N(H^s)$ and the operator $D^{(N)}$ is closed, we conclude that (33) holds for $k = 0$. Let us now assume that (33) holds for some $k \geq 0$. Set $l = N - k - 1$ and write $V = (v_1, \dots, v_l)^T \in (H^s)^l$, so that

$$D^{(l)} f_{k+1}(u)V = \mathrm{i} \int_0^\infty \mathrm{e}^{\tau B}\, D^{(l)} g_{k+1}(\mathrm{e}^{\tau A}u)[\mathrm{e}^{\tau A}V]\,\mathrm{d}\tau\,. \tag{34}$$

Further, using the Leibniz rule,

$$D^{(l)}(Df_j f_m)V = \sum_{\pi \in \mathcal{P}(\mathbb{N}_l)} D^{(l+1-|\pi|)} f_j \left[D^{(|\pi|)} f_m V_\pi, V_{\mathbb{N}_l \setminus \pi}\right], \tag{35}$$

where $\mathcal{P}(\mathbb{N}_l)$ denotes the power set of $\mathbb{N}_l \equiv \{1, \dots, l\}$, $|\pi|$ is the cardinality of $\pi$, $S \in \mathcal{P}(\mathbb{N}_l)$, and $V_S$ is any vector $(v_{s_1}, \dots, v_{s_m})$ such that $\{s_1, \dots, s_m\} = S$ and $|S| = m$. The characterizations defining membership in $\mathcal{C}_{\mathrm{b}}^{N-k-1}(H^s)$ follow by using the induction hypothesis for $f_j$ and $f_m$. It follows, by differentiating $g_k = \sum_{j+m=k-1} Df_j f_m$ and using (35), that

$$g_k \in \mathcal{C}_{\mathrm{b}}^{N-k}(H^s) \subset \mathcal{C}_{\mathrm{b}}^0(H^s) \tag{36}$$

for $k = 0, \dots, N$. This proves (33).

We now assume further that $g \in \mathcal{C}_{\mathrm{b}}^{N+1}(H^{s-1})$. Then the previous assertion of the lemma implies that

$$g_k \in \mathcal{C}_{\mathrm{b}}^{N+1-k}(H^{s-1}) \tag{37}$$

for $k = 0, \dots, N$, which in turn implies

$$Dg_k \in \mathcal{C}_{\mathrm{b}}^{N-k}(H^{s-1}) \subset \mathcal{C}_{\mathrm{b}}^0(H^{s-1}) \tag{38}$$

for $k = 0, \dots, N$. Note that (36) together with (38) make the functions $g_k$ fulfill the first assumption of Theorem 2. Hence, applying this theorem successively with the recursively known functions $g_k$ and operators $A = -\varepsilon \mathrm{i} L_-$ and $B = \varepsilon \mathrm{i} L_+$, where $\mathcal{H} = H^{s-1}$ and $\mathcal{D} = H^s$, we conclude that the $f_k$ solve the order conditions as claimed. □

## 4. The Cauchy problems for the full and for the slow equation

In this section, we study the Cauchy problem for (1) and for the slow equation (10) on $\mathcal{C}([0,T],H^s(\mathbb{T}))$. We begin by noting that (1), written as a first order system in (2), takes the abstract form

$$\partial_t \Psi = \mathcal{A}\Psi + F(\Psi) \tag{39}$$

with

$$\Psi = \begin{pmatrix} u \\ \partial_t u \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} 0 & 1 \\ \Delta/\varepsilon & \mathrm{i}/\varepsilon \end{pmatrix}, \quad \text{and} \quad F(\Psi) = \begin{pmatrix} 0 \\ \varepsilon^{-1}\, g(u) \end{pmatrix}. \tag{40}$$

We cast (39) in its mild formulation

$$\Psi(t) = \mathrm{e}^{t\mathcal{A}}\Psi(0) + \int_0^t \mathrm{e}^{(t-s)\mathcal{A}}\, F(\Psi(s))\,\mathrm{d}s\,, \tag{41}$$

where

$$\mathrm{e}^{t\mathcal{A}} = \mathrm{e}^{\alpha t}\begin{pmatrix} \cos(tM) - \alpha M^{-1}\sin(tM) & M^{-1}\sin(tM) \\ \varepsilon^{-1}\Delta M^{-1}\sin(tM) & \cos(tM) + \alpha M^{-1}\sin(tM) \end{pmatrix} \tag{42}$$

with

$$M = \frac{\sqrt{1-4\varepsilon\Delta}}{2\varepsilon} \quad \text{and} \quad \alpha = \frac{\mathrm{i}}{2\varepsilon}\,. \tag{43}$$

We first study the action of the operator group $\mathrm{e}^{\mathcal{A}t}$ on the space $X_s = H^s \times H^{s-1}$ endowed with the norm $\|\cdot\|_{X_s}$ defined via

$$\|\Psi\|_{X_s}^2 = \|M\psi_1\|_{s-1}^2 + \|\psi_2\|_{s-1}^2 \tag{44}$$

for $\Psi = (\psi_1,\psi_2)^T \in X_s$.

**Lemma 4.** *For $t \in \mathbb{R}$ fixed, $\mathrm{e}^{t\mathcal{A}}$ maps $X_s$ into itself. Moreover,*

$$\|\mathrm{e}^{t\mathcal{A}}\Psi\|_{X_s} \leq 2\,\|\Psi\|_{X_s} \tag{45}$$

*and*

$$\lim_{t\to 0}\|\mathrm{e}^{t\mathcal{A}}\Psi - \Psi\|_{X_s} = 0 \tag{46}$$

*for all $\Psi \in X_s$.*

*Proof.* Let $\Psi = (\psi_1,\psi_2)^T \in X_s$. Then we have

$$\mathrm{e}^{t\mathcal{A}}\Psi = \mathrm{e}^{\alpha t}\begin{pmatrix} \cos(tM)\psi_1 - \alpha M^{-1}\sin(tM)\psi_1 + M^{-1}\sin(tM)\psi_2 \\ \varepsilon^{-1}\Delta M^{-1}\sin(tM)\psi_1 + \cos(tM)\psi_2 + \alpha M^{-1}\sin(tM)\psi_2 \end{pmatrix}. \tag{47}$$

It is clear that $M^{-1}\psi_2 \in H^s(\mathbb{T})$ and $\Delta M^{-1}\psi_1 \in H^{s-1}(\mathbb{T})$, so that $\mathrm{e}^{t\mathcal{A}}\Psi \in X_s$. Take the $X_s$-norm of (47) and note that $\|\Psi\|_{X_s}$ appears as a common factor on the right; this implies (45). Statement (46) follows from the dominated convergence theorem. □

**Proposition 5.** *Let $g \in \mathcal{C}^1_{\mathrm{b}}(H^s)$ for some $s \in \mathbb{R}$ and $\Psi_0 \in X_s$. Then there exists $T^* = T^*(\varepsilon) > 0$ such that (39) has a unique solution $\Psi \in \mathcal{C}([0,T^*),X_s)$ with $\Psi(0) = \Psi_0$. Moreover,*

$$\lim_{t\nearrow T^*}\|\Psi(t)\|_{X^s} = \infty \tag{48}$$

*if $T^* < \infty$.*

*Proof.* The proof is a standard contraction mapping argument on the mild formulation (41). Fix $R \geq 2\,\|\Psi_0\|_{X_s}$ and define

$$E_T = \bigl\{\Phi \in \mathcal{C}([0,T], X_s)\colon \sup_{t\leq T} \|\Phi(t)\|_{X^s} \leq 2R\bigr\}\,. \tag{49}$$

Then, since $g \in \mathcal{C}^1_{\mathrm{b}}(H^s)$, there exists $C_1 = C_1(\varepsilon, R) > 0$ such that

$$\|F(\Psi) - F(\Phi)\|_{X^s} \leq C_1\,\|\Psi - \Phi\|_{X^s} \tag{50}$$

for $\Psi, \Phi \in E_T$. We write $\Gamma(\Psi)$ to denote the right hand side of (41). For $\Psi \in E_T$, using (45) together with (50), we estimate

$$\|\Gamma(\Psi)\|_{X_s} \leq 2\,\|\Psi_0\|_{X_s} + T\,C_1 \sup_{t\leq T} \|\Psi(t)\|_{X^s} \leq R + T\,R\,C_1\,. \tag{51}$$

Thus, there exists $T_0 = T_0(R, \varepsilon) > 0$ such that $\sup_{t\leq T_0} \|\Gamma(\Psi(t))\|_{X^s} \leq 2R$; in other words, $E_{T_0} \subset \Gamma(E_{T_0})$. Further, for $\Psi, \Phi \in E_{T_0}$,

$$\sup_{t\leq T_0} \|\Gamma(\Psi(t)) - \Gamma(\Phi(t))\|_{X_s} \leq T_0\,C_1 \sup_{t\leq T_0} \|\Psi(t) - \Phi(t)\|_{X^s}\,. \tag{52}$$

Hence, there exists $T_1 = T_1(R, \varepsilon) \leq T_0$ such that $\Gamma$ is a strict contraction on $E_{T_1}$. Thus, $\Gamma$ has a unique fixed point $\Psi \in E_{T_1}$ which is the unique mild solution of (39). As (39) is invariant with respect to time translation, the fixed point argument can be restarted at initial time $t = T_1$, i.e., the time interval of existence is an open interval $[0, T^*)$, where $T^*$ is either infinite or $\Psi(t)$ blows up as $t \nearrow T^*$. □

We now turn to the slow equation (10). Its mild formulation reads

$$U(t) = \mathrm{e}^{tL_-}U_0 + \int_0^t \mathrm{e}^{(t-s)L_-}\,F^N_{\mathrm{slow}}(U(s))\,\mathrm{d}s\,. \tag{53}$$

Since $\mathrm{e}^{tL_-}$ is a unitary group, the situation here is easier than for (41), so that we can obtain an interval of existence which is independent of $\varepsilon$. The result is the following.

**Proposition 6.** *Let $\varepsilon_0 > 0$ and $g \in \mathcal{C}^{N+1}_{\mathrm{b}}(H^s)$ for some $s \in \mathbb{R}$ and $U_0 \in H^s(\mathbb{T})$. Then there exist $T^*_{\mathrm{slow}} > 0$ and $R > 0$ independent of $\varepsilon \in [0, \varepsilon_0]$ such that* (10) *has a unique solution $U \in \mathcal{C}([0, T^*_{\mathrm{slow}}], H^s(\mathbb{T}))$ with $U(0) = U_0$ and $\sup_{t\in[0,T^*_{\mathrm{slow}}]} \|U(t)\|_s \leq R$.*

*Proof.* As in the proof of Proposition 5, we use a fixed point argument on the mild formulation (53). We only detail the differences to the previous argument. We write $\Lambda(U)$ to denote the right hand side of (53) and, for $R \geq \|U_0\|_s$ fixed, define

$$E_T = \bigl\{V \in \mathcal{C}([0,T], H^s)\colon \sup_{t\leq T} \|V(t)\|_s \leq 2R\bigr\}\,. \tag{54}$$

Noting that $\mathrm{e}^{tL_-} \in \mathcal{L}(H^s(\mathbb{T}))$ is an isometry and using (33), we find that there exist $C_k = C_k(R) > 0$ such that

$$\begin{aligned}\|\Lambda(U)(t)\|_s &\leq \|U_0\|_s + T \sup_{t\leq T} \sum_{k=0}^N \varepsilon^k\,\|f_k(U(t))\|_s \\ &\leq \|U_0\|_s + T\,\sum_{k=0}^N \varepsilon^k\,C_k\end{aligned} \tag{55}$$

for all $U \in E_T$. Note that for $\varepsilon \in [0, \varepsilon_0]$, the term in the right-hand side of (55) does not depend on $\varepsilon$. Hence, there exists $T > 0$ independently of $\varepsilon$ such that

$\Lambda(E_T) \subset E_T$. Similarly, we can verify that there exists $T_1 \in (0, T]$ independent of $\varepsilon$ such that $\Lambda$ is a strict contraction on $E_{T_1}$. $\square$

## 5. Main theorem

We are now able to prove the main theorem on the shadowing of consistently initialized solutions of the full system by solutions of the slow systems over an $O(1)$-interval of time, determined by the time interval of existence of the slow system.

**Theorem 7.** *Fix $N \in \mathbb{N}$, $s \in \mathbb{R}$, and suppose $g \in \mathcal{C}_{\mathrm{b}}^{N+2}(H^s) \cap \mathcal{C}_{\mathrm{b}}^{N+2}(H^{s-1})$. Let $F^{N+1} = \sum_{k=0}^{N+1} \varepsilon^k f_k$, with coefficient vector fields given by (32a). Take $u_0 \in H^s$ and let $U$ denote the solution to the slow equation (10) with $U(0) = u_0$. Its uniform time of existence $T^*_{\mathrm{slow}} > 0$ and uniform $H^s$ bound $R > 0$ are given by Proposition 6. Then there exist $\varepsilon_1 > 0$ and $C$, only depending on $T^*_{\mathrm{slow}}$ and $R$, such that for all $\varepsilon \in (0, \varepsilon_1]$ the solution $u$ to the full Cauchy problem (1) with $u(0) = u_0$ and $\partial_t u(0) = L_- u_0 + F^N_{\mathrm{slow}}(u_0)$ exists on the same uniform time interval $[0, T^*_{\mathrm{slow}}]$ and*

$$\sup_{0 \le t \le T^*_{\mathrm{slow}}} \|u(t) - U(t)\|_s \le C\, \varepsilon^{N+1} \,. \tag{56}$$

*Proof.* Since $g \in \mathcal{C}_{\mathrm{b}}^{N+1}(H^s) \cap \mathcal{C}_{\mathrm{b}}^{N+2}(H^{s-1})$, Lemma 3 asserts that the coefficient functions $f_k \colon H^s \to H^s$ satisfy the order conditions (22) so that, continuing the argument from (18b),

$$\partial_t w = \left( \frac{\mathrm{i}}{\varepsilon} - L_- - DF^{N+1}_{\mathrm{slow}}(u) \right) w + \varepsilon^{N+1} R_N(u) \,, \tag{57}$$

where the remainder $R_N \colon H^s \to H^s$ is given by

$$R_N = - \sum_{k=N+1}^{2N+2} \varepsilon^{k-N-1} \sum_{\substack{j+l=k \\ j,l \le N+1}} Df_j f_l \,. \tag{58}$$

Moreover, $g \in \mathcal{C}_{\mathrm{b}}^{N+2}(H^s)$ implies, again by Lemma 3, that $f_j \in \mathcal{C}_{\mathrm{b}}^{N+2-j}(H^s)$ and $Df_j \in \mathcal{C}_{\mathrm{b}}^{N+1-j}(H^s, \mathcal{L}(H^s))$ for $j = 0, \dots, N+1$.

Let $\varepsilon_0$ as in Proposition 6. Then, for every $\varepsilon \le \varepsilon_0$, let $T(\varepsilon) \in (0, T^*_{\mathrm{slow}}]$ be maximal such that the solution $u$ to (1) with the stated initial conditions exists on $[0, T(\varepsilon)]$ with

$$\sup_{0 \le t \le T(\varepsilon)} \|u(t)\|_s \le 2R \,. \tag{59}$$

Such interval of time exists by Proposition 5. Thus, there exist constants $C_1$ and $C_2$ which depend only on $R$ such that the Duhamel formula for (57) implies that

$$\|w(t)\|_s \le \|w(0)\|_s + C_2\, \varepsilon^{N+1} + C_1 \int_0^t \|w(\tau)\|_s \,\mathrm{d}\tau \,. \tag{60}$$

Since $w(0) = 0$ by assumption, the Gronwall inequality implies that there exists a constant $C_3 = C_3(R, T^*_{\mathrm{slow}})$ such that

$$\sup_{0 \le t \le T(\varepsilon)} \|w(t)\|_s \le C_3\, \varepsilon^{N+1} \,. \tag{61}$$

Next, taking the difference between (10) and (18a),

$$\partial_t (u - U) = L_-(u - U) + F^N_{\mathrm{slow}}(u) - F^N_{\mathrm{slow}}(U) + w + \varepsilon^{N+1} f_{N+1}(u) \,, \tag{62}$$

and noting that the last two terms are $\mathcal{O}(\varepsilon^{N+1})$ in the sense stated above, we estimate

$$\begin{aligned} \|u - U\|_s \leq \|u(0) - U(0)\|_s + C_4 \int_0^t \|u - U\|_s \, \mathrm{d}\tau \\ + \int_0^t \|w + \varepsilon^{N+1} f_{N+1}(u)\|_s \, \mathrm{d}\tau \end{aligned} \tag{63}$$

where $C_4$ depends only on $R$ and $T^*_{\mathrm{slow}}$. Since $U(0) = u(0)$, the Gronwall inequality implies that there exists $C > 0$ depending only on $R$ and $T^*_{\mathrm{slow}}$ such that

$$\sup_{0 \leq t \leq T(\varepsilon)} \|u(t) - U(t)\|_s \leq C \, \varepsilon^{N+1} \,. \tag{64}$$

Now choose $\varepsilon_1 \in (0, \varepsilon_0]$ such that

$$\varepsilon_0^{N+1} < \frac{R}{C} \,. \tag{65}$$

Then, for $\varepsilon \in (0, \varepsilon_1]$,

$$\sup_{0 \leq t \leq T(\varepsilon)} \|u(t)\|_s \leq \sup_{0 \leq t \leq T(\varepsilon)} (\|u(t) - U(t)\|_s + \|U(t)\|_s) \leq C \, \varepsilon^{N+1} + R < 2R \,. \tag{66}$$

This proves that $T(\varepsilon) = T^*_{\mathrm{slow}}$, for otherwise $T(\varepsilon)$ could not be maximal, so that (56) holds as stated. □

## Acknowledgments

The authors thank Anton Kutsenko for useful discussions. The work was supported by German Research Foundation grant OL-155/6-1. MO further acknowledges support through German Research Foundation Collaborative Research Center TRR 181.

## References

[1] W. Bao, Y. Cai, and X. Zhao, *A uniformly accurate multiscale time integrator pseudospectral method for the Klein–Gordon equation in the nonrelativistic limit regime*, SIAM J. Numer. Anal., 52 (2014), pp. 2488–2511.

[2] W. Bao and X. Zhao, *A uniformly accurate (UA) multiscale time integrator Fourier pseudospectral method for the Klein–Gordon–Schrödinger equations in the nonrelativistic limit regime*, Numer. Math., 135 (2017), pp. 833–873.

[3] P. Chartier, N. Crouseilles, M. Lemou, and F. Méhats, *Uniformly accurate numerical schemes for highly oscillatory Klein–Gordon and nonlinear Schrödinger equations*, Numer. Math., 129 (2015), pp. 211–250.

[4] C. J. Cotter and S. Reich, *Semigeostrophic particle motion and exponentially accurate normal forms*, Multiscale Model. Simul., 5 (2006), pp. 476–496.

[5] X. Dong, Z. Xu, and X. Zhao, *On time-splitting pseudospectral discretization for nonlinear Klein–Gordon equation in nonrelativistic limit regime*, Commun. Comput. Phys., 16 (2014), pp. 440–466.

[6] E. Faou and K. Schratz, *Asymptotic preserving schemes for the Klein–Gordon equation in the non-relativistic limit regime*, Numer. Math., 126 (2014), pp. 441–469.

[7] G. A. Gottwald and M. Oliver, *Slow dynamics via degenerate variational asymptotics*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 470 (2014), p. 20140460.

[8] E. Hille, *Une généralisation du problème de Cauchy*, Ann. Inst. Fourier, 4 (1952), pp. 31–48.

[9] K. U. Kristiansen and C. Wulff, *Exponential estimates of symplectic slow manifolds*, J. Differ. Equations, 261 (2016), pp. 56–101.

[10] Y. Lu and Z. Zhang, *Higher order asymptotic analysis of the Klein-Gordon equation in the non-relativistic limit regime*, Asymptot. Anal., 102 (2017), pp. 157–175.

[11] S. Machihara, K. Nakanishi, and T. Ozawa, *Nonrelativistic limit in the energy space for nonlinear Klein-Gordon equations*, Math. Ann., 322 (2002), pp. 603–621.

[12] R. S. MacKay, *Slow manifolds*, in Energy Localisation and Transfer, T. Dauxois, A. Litvak-Hinenzon, R. S. MacKay, and A. Spanoudaki, eds., World Scientific, 2004, pp. 149–192.

[13] N. Masmoudi and K. Nakanishi, *From nonlinear Klein-Gordon equation to a system of coupled nonlinear Schrödinger equations*, Math. Ann., 324 (2002), pp. 359–389.

[14] K. Matthies and A. Scheel, *Exponential averaging for Hamiltonian evolution equations*, Trans. Amer. Math. Soc., 355 (2003), pp. 747–773.

[15] H. Mohamad and M. Oliver, *$H^s$-class construction of an almost invariant slow subspace for the Klein–Gordon equation in the non-relativistic limit*, J. Math. Phys., 59 (2018), pp. 051509, 8.

[16] N. Nekhoroshev, *An exponential estimate of the time of stability of a nearly-integrable Hamiltonian system*, Russ. Math. Surv., 32 (1977), pp. 1–65.

[17] M. Tsutsumi, *Nonrelativistic approximation of nonlinear Klein–Gordon equations in two space dimensions*, Nonlinear Anal., 8 (1984), pp. 637–643.

[18] C. Wulff and M. Oliver, *Exponentially accurate Hamiltonian embeddings of symplectic A-stable Runge–Kutta methods for Hamiltonian semilinear evolution equations*, Proc. Roy. Soc. Edinburgh Sect. A, 146 (2016), pp. 1265–1301.

School of Engineering and Science, Jacobs University, 28759 Bremen, Germany