Quartic normal forms for the periodic nonlinear Schrödinger equation with dispersion management

Panayotis Panayotaros
Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526, USA
Received 7 May 2003; received in revised form 26 November 2003; accepted 27 January 2004
Communicated by C.K.R.T. Jones

Abstract
We investigate Birkhoff normal forms for the periodic nonlinear Schrödinger equation with dispersion management. The normalization we describe is related to averaging arguments considered in the literature, and has the advantage of producing fewer resonant couplings between high spatial frequency modes. One consequence is that the normal form equations have invariant subspaces of large but finite dimension, where we can find several classes of periodic orbits. The formal arguments apply to other related dispersive systems, and to normal forms of high order. We also present a rigorous version of the normal form calculation and show that solutions of the quartic normal form equations remain close to solutions of the full system over a time that is inversely proportional to a small nonlinearity parameter.

© 2004 Elsevier B.V. All rights reserved.
PACS: 42.65.Sf; 45.20.Jj
Keywords: Nonlinear Schrödinger equations; Parametric forcing; Normal forms

1. Introduction

The nonlinear Schrödinger equation with dispersion management $u_t = -d(t)u_{xx} + 2i\gamma |u|^2u$, where $d(t)$ is a periodic real valued function, and $\gamma$ is real models the propagation of signals in an optical transmission line whose dispersive properties vary along the line. The initial condition $u(x,0)$ is interpreted as the emitted signal, while $u(x,t)$ is the signal at a distance $t$ from the origin of the transmission line. Varying (“managing”) the dispersion can lead to more robust propagation of signals of small amplitude (see e.g. [4,10,23]), and the idea has attracted considerable experimental attention in recent years. Theoretical studies have focused on nonlinear effects, modeled to lowest order by the cubic NLS + DM system above (see e.g. [5,20]).

The equation has been mainly studied on the line (e.g. with decay boundary conditions), and we will here consider the periodic case. Theoretical and numerical studies of the NLS + DM equation suggest that its dynamics is nontrivial, and the periodic system is an example of an infinite dimensional Hamiltonian system of independent interest. Also,
the periodic theory can be directly compared to simulations that use periodic boundary conditions and spatial discretizations that preserve the Hamiltonian structure. Such simulations can be of heuristic value for studying the equation on line, e.g. in studies of the evolution of well-localized, rapidly decaying solutions, and in finding orbits that have variational characterizations. We must however emphasize the possible differences in the dynamics of the periodic and unbounded problems, especially in questions of persistence of approximate solutions and of stability.

We will concentrate on the practically interesting parameter regime where the average $\delta$ of the “dispersion management function” $d(t)$, and the nonlinearity parameter $\gamma$ are comparable and small in absolute value. We also assume that the frequency $\Omega$ of the dispersion management function $d(t)$ is at least of $O(1)$ in absolute value, and we will be interested in solutions of $O(1)$.

One of the approaches in studying the NLS + DM equation on the line has been to formally derive an averaged model equation that is autonomous (see [5,21], also [1]). The averaged equation has the structure of a Schrödinger equation with a nonlocal cubic nonlinearity. Numerical studies of the averaged equation have indicated the existence of localized periodic solutions with a Gaussian-like single pulse profile, referred to as DM solitons (see e.g. [12,17]). More recently the existence of periodic solutions for the averaged equation on the line was also proved by variational methods (for $\delta < 0$, see [14,25]), and by a bifurcation argument (for $\delta \neq 0$, see [8]). The variational characterization of the DM soliton for $\delta < 0$ also implies nonlinear stability, and there are also error estimates for the averaged equations (see [14]). Also note that the problem of averaging of a dispersive system coupled to a fast oscillator has been also considered in a more abstract setting in [16].

In this work we investigate an alternative but related asymptotic theory for the periodic problem. Our first goal is to modify the approach of [5,21], and produce a simpler normal form equation. In particular, to recover the averaging theory of [5,21] we split the full system into an unperturbed and a perturbation part where the unperturbed part has only one fast oscillating degree of freedom (with frequency $\Omega$). We therefore have a very resonant problem, and the resulting normal form equations are not sparse enough. We here add to the unperturbed part of the system the high frequency oscillations corresponding to the high spatial frequency part of the averaged dispersive term. The resulting Birkhoff normal forms are more involved, but leave fewer resonant terms coupling high spatial frequency modes. A main observation is that, under a mild condition on $\Omega$, the quartic normal form equation has a finite dimensional invariant subspace. The dimension of this subspace may be large (of $O(|\delta|^{1/2})$), but some aspects of its dynamics are easier to analyze. For instance, elementary arguments imply the existence of several classes of periodic orbits, some of which are analogous to the orbits considered in the literature for the averaged NLS + DM equation on the line. The argument leading to the existence of invariant subspaces applies to higher order normal form equations, and to NLS + DM equations with higher dispersion (considered recently in [14,16]). A rough interpretation of the normalization we describe is that averaging over the fast oscillations of the high spatial frequency modes leads to invariant subspaces for the slower low spatial frequency motions.

In the second part of the paper we show that the solutions of the alternative quartic normal form we construct remain $O(\gamma)$ close to solutions of the full system over a time of $O(1/|\gamma|^2)$. The main assumption is that the initial conditions for the normal form system are of $O(1)$ in an appropriate Sobolev norm (the precise statement is in Section 4). Our approach follows the spirit of the formal calculations, where we consider the NLS + DM equation as an autonomous Hamiltonian system in an extended phase space. The transformation theory for infinite dimensional Hamiltonian systems has been developed by many authors (see e.g. [6]). Some points that require attention here is the low regularity of the dispersion management function (required by the applications), and the fact that the transformation in the extended phase is not close to the identity in some directions. The error estimates also use the fact that solutions of the full system cannot grow too much over the $O(1/|\gamma|^2)$ time interval of interest. Such control may not available for longer times, and it is not clear at present whether we can extend the error estimates to the higher order normal forms. The work of [16] suggests that long time regularity could lead to averaging results over times that are exponentially long in $|\gamma|^{-1}$.
The paper is organized as follows. In Section 2 we introduce the Hamiltonian structure of the NLS + DM system and establish the notation used in the formal calculations. We also emphasize the parameters of the problem. In Section 3 we formally construct Birkhoff normal forms and study some properties of the normal form systems. In Section 4 we give a rigorous version of the first order normal form calculation and estimate the distance between solutions of the quartic normal form system and the dispersion managed NLS equation.

2. Hamiltonian structure

We consider the initial value problem for the nonautonomous equation

\[ \dot{u}_t = i\delta(t)u_{xx} - 2i\gamma|u|^2u \]  
(2.1)

with \( u(x, t) \) a complex valued function satisfying periodic boundary conditions \( u(x, t) = u(x + 2\pi, t) \). The “dispersion management” function \( d(t) \) and the parameter \( \gamma \) are real. As remarked in Section 1 the “time” \( t \) in (2.1) is the distance from the point where we emit the signal, while the “spatial variable” \( x \) of (2.1) is physically the time. The “initial condition” \( u(x, 0) \) for (2.1) is the signal we send, and is assumed to be \( 2\pi \)-periodic. Also, the function \( d(t) \) in (2.1) will be \( T \)-periodic, and we decompose it as

\[ d(t) = \delta + \tilde{d}(t) \]

with \( \delta = \frac{1}{T} \int_0^T d(s) \, ds \),

(2.2)

the average. Letting \( \Omega = 2\pi/T \) we assume that \( |\Omega| \geq O(1) \). We will further assume that \( |\delta| \sim |\gamma| \ll 1 \). Note that since the system is nonautonomous we should consider initial conditions \( u(x, t_0), t_0 \in \mathbb{R} \). Equivalently, we here fix \( t_0 = 0 \) and handle the general case by appropriately shifting \( d(t) \).

Remark 2.1. The parameters \( \gamma, \delta, \Omega \) are assumed dimensionless. Some physically interesting special cases of the parameter regime we consider are: (i) \( |\Omega| \sim |\gamma|^{-1} \), (ii) \( |\Omega| \sim |\gamma|^{-1} \) with \( |\Omega|/h \) of \( O(1) \), and (iii) \( |\Omega| \sim |\gamma|^{-1} \) with \( |\Omega|/h \ll 1 \), where \( h \) is the amplitude of \( \tilde{d}(t) \). Also of interest is the case where (iv) \( h \ll |\gamma| \).

Remark 2.2. The dimensional analysis leading to the above parameters is derived for the equation on the line. The periodic problem is often used in numerical studies of the problem on the line, and this is reasonable for solutions that have a negligible amplitude (e.g. below \( 10^{-16} \) for double precision calculations) outside a region of \( O(1) \) size.

By (2.1) we can rescale such a region to have length \( 2\pi \) by rescaling the “spatial” variable \( x \) to \( \alpha x \) and the amplitude of the dispersion function \( d \) to \( a^2d \) without changing qualitatively the assumptions on the relative sizes of the parameters in the previous remark.

It is easy to see that Eq. (2.1) has the structure of a nonautonomous Hamiltonian system. To perform normal form calculations it will be convenient to first rewrite (2.1) using certain “amplitude” variables, and then make the system autonomous by introducing an additional angle variable. For the first step, we denote the Fourier transform of \( u(x, t) \) by \( u_k(t) \), and use the notation of (2.2) to define the variables \( a_k(t), k \in \mathbb{Z} \) by

\[ a_k(t) = u_k(t) e^{\alpha_k \tilde{\Lambda}(t)} \quad \text{with} \quad \alpha_k = k^2, \quad \tilde{\Lambda}(t) = \int_0^t \tilde{d}(s) \, ds. \]

(2.3)

From (2.1), the variables \( u(k, t) \) then evolve according to

\[ \dot{a_k} = -2i\gamma \sum_{k_1, k_2, k_3 \in \mathbb{Z}} a_{k_1} a_{k_2} a_{k_3} e^{i(k_1 + k_2 - k_3)} e^{-i\alpha_{k_1} + i\alpha_{k_2} - i\alpha_{k_3}} \tilde{\Lambda}(t), \quad k \in \mathbb{Z}. \]

(2.4)
where \( \delta_r = 1 \) if \( r = 0 \), and 0 otherwise. The initial condition is \( u_0(0) = u_1(0), k \in \mathbb{Z} \). By the definition of \( \tilde{\Lambda} \) the right hand side of (2.4) is \( T \)-periodic. The time dependence is therefore absorbed in the nonlinear term. To eliminate the explicit dependence on time in (2.4) we consider an angle \( \phi \in [0, 2\pi) \), and add to (2.4) the equation

\[
\dot{\phi} = \Omega \quad \text{with} \quad \phi(0) = \phi_0 = 0.
\]  

(2.5)

We also define the function \( \Lambda \) by \( \Lambda(\phi) = \tilde{\Lambda}(\phi/T) \), and note that \( \Lambda \) is \( 2\pi \)-periodic with zero average. Then, the nonautonomous system (2.4) is equivalent to the autonomous system consisting of (2.4) with \( \Lambda(t) \) replaced by \( \Lambda(\phi) \), and (2.5). Adding an "action" variable \( J \in \mathbb{R} \), we further define the Poisson bracket \( \{,\} \) on pairs of functions \( F,G \) of the variables \( a_k,a_\ast_k,k \in \mathbb{Z} \) and \( \phi,J \) by

\[
\{F,G\} = -i \sum_{k \in \mathbb{Z}} \left( \frac{\partial F}{\partial a_k} \frac{\partial G}{\partial a_\ast_k} - \frac{\partial F}{\partial a_\ast_k} \frac{\partial G}{\partial a_k} \right) + \frac{\partial F}{\partial J} \frac{\partial G}{\partial \phi} - \frac{\partial F}{\partial \phi} \frac{\partial G}{\partial J}.
\]  

(2.6)

A straightforward calculation then shows the following proposition.

**Proposition 2.3.** The evolution equation for the variables \( a_k,k \in \mathbb{Z} \), and \( J,\phi \) above is the Hamiltonian system

\[
a_k \to \{a_k,H\}, \quad k \in \mathbb{Z}, \quad \phi = [\phi, H], \quad J = [J, H],
\]  

(2.7)

where the Hamiltonian \( H \) is

\[
H = \delta \sum_{k \in \mathbb{Z}} |a_k|^2 - \Omega J + \gamma \sum_{k_1,k_2,k_3,k_4 \in \mathbb{Z}} e^{i\phi} a_{k_1}a_{k_2}a_{k_3}a_{k_4} I(k_1,k_2,k_3,k_4,n),
\]  

(2.8)

and the coefficients \( I(k_1,k_2,k_3,k_4,n) \) are given by

\[
I(k_1,k_2,k_3,k_4,n) = f_{\omega}(n) h_{k_1-k_2,k_3-k_4,n},
\]  

(2.9)

\[
m = a_{k_1} + a_{k_2} - a_{k_3} - a_{k_4}, \quad f_{\omega}(n) = (2\pi)^{-1/2} \int_0^{2\pi} e^{-in\Lambda(\phi)} e^{-in\phi} d\phi.
\]  

(2.10)

Note the equation for \( J \) gives as the rate of change of the "energy" \( H + \Omega J \) (up to a factor \( \Omega^{-1} \)).

**Remark 2.4.** The Hamiltonian \( H \) in (2.8) shows that the parameter range we are considering describes a system with small dispersion and small nonlinearity. The "weakly nonlinear" parameter regime where \( |\delta| \sim |\Omega| \sim O(1) \) and \( |\gamma| \ll 1 \) will be considered elsewhere. Note that the weakly nonlinear regime poses some interesting problems related to the works of [2,7,19] on weakly perturbed 1D NLS equations (with Dirichlet boundary conditions). As we see in the next section, the special parameter ranges (iii) and (iv) of Remark 2.1 also lead to "near-integrable" systems (see [1,22], respectively, for the two limits).

The above setup can be generalized to other dispersion relations \( a_{k_0} \), and to the case where the parameter \( \gamma \) is replaced by a time-dependent real function. Also, we can easily extend the formalism to quasi-periodic dispersion management functions \( d(t) \). A practically important example where the coefficients \( f_{\omega}(n) \) in (2.10) can be evaluated in closed form is the piecewise constant \( T \)-periodic dispersion management function

\[
d(t) = \delta + \tilde{\Lambda} \sin \Omega t,
\]  

(2.11)

with \( \tilde{\Lambda} + B(T - t) = 0 \), i.e. the average of \( d(t) \) over \([0,T] \) is \( \delta \). Another interesting example is the real analytic dispersion management function

\[
d(t) = \delta + \tilde{\Lambda} \sin \Omega t,
\]  

(2.12)
whose coefficients \( \hat{f}_m(n) \) are Bessel functions. For general dispersion management functions we can also obtain some information about the coefficients \( \hat{f}_m(n) \) by asymptotic arguments, e.g. in the large \(|m|\) limit. The coefficients are discussed further in the next section.

3. Quartic Birkhoff normal forms

In this section we split the Hamiltonian into two parts, the “unperturbed part” and the “perturbation”, and seek to simplify the perturbation part by a near-identity canonical transformation. Near-identity canonical transformations smoothly connected to the identity can be constructed by composing time-1 maps of Hamiltonian vector fields, and we will consider transformations leading to the well-known Birkhoff normal forms. The assumed ranges of parameters \( \Omega, \delta, \gamma \) suggest two possible splittings of the Hamiltonian. First, since \(|\Omega|\) is (at least) of \( O(1) \), and \(|\delta| \) and \(|\gamma| \) are small we can take the “unperturbed part” of the Hamiltonian \( H \) of (2.8) to be \(-\Omega J \). The resulting Birkhoff normal form equations are (the periodic analogues of) the averaged equations of [5,21] (see also [1]). We briefly rederive this averaging theory using the language of normal forms below. Our main goal here is to investigate an alternative splitting of the Hamiltonian \( H \) where the unperturbed part consists of \(-\Omega J \) plus the quartic terms of \( H \) that describe oscillations with frequencies that are at least of \( O(1) \). We will show that the normal form equations derived using the second splitting can have finite dimensional invariant subspaces.

To recover the averaging theory of [5,21] we write the Hamiltonian \( H \) of (2.8) as \( H = -\Omega J + H_2 + H_4 \) with \( H_2 \) and \( H_4 \) the quadratic and quartic parts, respectively. We seek a function \( \psi_1 \) such that the canonical transformation obtained by the time-1 map \( \Phi^{1}_{\psi_1} \) of the Hamiltonian flow of \( \psi_1 \) simplifies the “perturbation part” \( H_2 + H_4 \). Specifically, we formally write

\[
H \circ \Phi^{1}_{\psi_1} = \exp(\text{Ad}_{\psi_1}H) = -\Omega J + H_2 + H_4 + [\psi_1, -\Omega J] + Y_1
\]

with \( Y_1 \) representing the remaining terms. By the definition of the Poisson bracket we see that each monomial

\[
y(k_1, k_2, k_3, k_4, n)a_{k_1}a_{k_2}a_{k_3}^*a_{k_4}^* e^{i n \phi},
\]

in \( H_4 \) is eliminated by a monomial

\[
y(\Omega)^{-1}H(k_1, k_2, k_3, k_4, n)a_{k_1}a_{k_2}a_{k_3}^*a_{k_4}^* e^{i n \phi},
\]

in \( \psi_1 \). Consequently, the resonance condition for the part \( H_4 \) is

\[
n\Omega = 0, \quad k_1 + k_2 - k_3 - k_4 = 0, \quad \hat{f}_{\psi_1}(n) \neq 0, \quad k_1, \ldots, k_4, n \in \mathbb{Z}
\]

with \( m = a_{k_1} + a_{k_2} - a_{k_3} - a_{k_4} \). We immediately see that the resonant part \( \tilde{H}_4 \) of \( H_4 \) is

\[
\tilde{H}_4 = \gamma \sum_{k_1, k_2, k_3, k_4 \in \mathbb{Z}} a_{k_1}a_{k_2}a_{k_3}^*a_{k_4}^* \hat{f}_{\psi_1}(0)(0)_{k_1, k_2, k_3, k_4}.
\]

We also easily see that the quadratic part \( H_2 \) is resonant. Choosing

\[
\psi_1 = \gamma \sum_{k_1, k_2, k_3, k_4 \in \mathbb{Z}, n \in \mathbb{Z}} a_{k_1}a_{k_2}a_{k_3}^*a_{k_4}^* \hat{f}_{\psi_1}(k_1, k_2, k_3, k_4, n) / n\Omega,
\]

the quartic normal form Hamiltonian is \( H_2 + \tilde{H}_4 \).
Regarding the structure of \( H_2 + \tilde{H}_2 \), let \( A_m \) be the set of all integers \( k_1, k_2, k_3, k_4 \) satisfying \( k_1 + k_2 - k_3 - k_4 = 0 \), and \( m = \omega_{k_1} + \omega_{k_2} - \omega_{k_3} - \omega_{k_4} \). The level sets \( A_m \) can be parameterized explicitly by two integers (we omit this here). The parameterization of \( \Lambda(\phi) \) becomes especially simple and we can write \( \tilde{H}_2 = \tilde{H}_{k,4} + \tilde{H}_{k,0} \) with

\[
\tilde{H}_{k,4} = 2y f_0(0) \sum_{k_1, k_2 \in \mathbb{Z}} |\omega_{k_1}|^2 |\omega_{k_2}|^2.
\]

(3.7)

\[
\tilde{H}_{k,0} = y \sum_{m \in \mathbb{Z}: m \neq 0} f_0(0) \sum_{k_1, k_2, k_3, k_4 \in \mathbb{Z}} a_{k_1} a_{k_2} a_{k_3} a_{k_4}.
\]

(3.8)

The part \( \tilde{H}_{k,4} \) is integrable (in the sense of Poincaré).

To indicate the structure of the coefficients \( f_0(0) \), we first consider the piecewise constant dispersion management function of (2.11) with \( \tau = T/2, \Lambda = h, \) and \( B = -h \). Then

\[
f_0(0) = \frac{3}{\sqrt{2\pi}} e^{-\frac{1}{2h\Omega^2}} - 1 \frac{\Omega}{h m}, \quad m \in \mathbb{Z}^*,
\]

(3.9)

and \( \tilde{f}_0(0) = \sqrt{2\pi} \). We can see that \( \mathcal{N}_{k,4} \) can only vanish for the discrete values of \( h \) where \( (h/\Omega) \in \pi\mathbb{Z} \), and also in the limit \( h \to +\infty \). In comparison, the coefficients \( f_0(0) \) for the real analytic dispersion management function of (2.12) with \( \Lambda = h \) are

\[
\tilde{f}_0(0) = \sqrt{2\pi} J_0(|m|h\Omega^{-1}), \quad m \in \mathbb{Z}^*
\]

(3.10)

with \( J_0 \) the Bessel function of order 0. For \( h \to +\infty \), the coefficients \( \tilde{f}_0(0), m \neq 0, \) of (3.10) decay as \( |m|h\Omega|^{-1.5} \). For more general integrable dispersion management functions, the definition of the coefficients \( f_0(0) \) in (2.10) implies that \( \tilde{f}_0(0) = \sqrt{2\pi} \). The decay of the coefficients in the amplitude \( h \) of \( \Lambda(\phi) \) and in \( |m| \neq 0 \) can be found by a stationary phase argument. For instance, for \( \Lambda(\phi) \) twice differentiable with nondegenerate critical points we expect a \( |m|h\Omega|^{-1.5} \) decay, while for \( \Lambda(\phi) \) Lipschitz but not differentiable we have \( |m|h\Omega|^{-1} \) decay (i.e. as in the two examples above). For \( h\Lambda(\phi) \) Hölder continuous with exponent less than unity, e.g. for \( d(t) \) unbounded but integrable, the coefficients \( \tilde{f}_0(0) \) can decay even faster as \( |m|h\Omega \to \infty \). Thus the nonintegrable part of \( \tilde{H}_2 \) decays faster in \( |m| \) and \( h \) for the more singular dispersion management functions. In the limit \( h \to 0, H_2 + \tilde{H}_2 \) reduces to the Hamiltonian of the cubic NLS. In that limit, it is most natural to view the NLS + DM system as a small perturbation of the cubic NLS equation (see [11,22,23]).

A main advantage of the normal form equations above is that they are straightforward to compute and extend to higher orders. Higher order resonance conditions are also trivial and do not involve any small divisors. Note that higher order calculations can be further simplified by assuming that \( |\delta| \sim |\gamma| \sim |\Gamma|^{-1} \), this is consistent with the scales of physical interest in the problem. Although the quartic and higher order normal form systems obtained by this theory are autonomous, their dynamics are still difficult to analyze (we discuss some known results for the line below).

In the alternative normalization below we try to eliminate more terms by better controlling the high spatial frequency nonlinear interactions. In particular, fix \( N = |\delta|^{-1/2} \) and let

\[
H_2 = \tilde{H}_2 + h_2
\]

(3.11)

with

\[
\tilde{H}_2 = \delta \sum_{|k| > N} a_k |a_k|^2, \quad h_2 = \delta \sum_{|k| \leq N} a_k |a_k|^2.
\]

(3.12)

The "unperturbed part" of the Hamiltonian, denoted by \( h_0 \), will now be \( h_0 = \tilde{H}_2 - \Omega J \) and will thus contain only the fast oscillators in \( H_2 \), i.e. the ones with frequencies that are greater than unity. The "perturbation part" will
be \( h_2 + H_4 \), i.e. it will contain "small" terms of \( O(\delta) \) and \( O(\gamma) \). As before, we seek to eliminate the lowest order nonresonant part of \( h_2 + H_4 \) by a canonical transformation \( \Phi_H^1 \), that is the time-1 map of an appropriate function \( x_1 \). We will have

\[
H \circ \Phi_H^1 = \exp(\mathcal{A}_d, H) = h_0 + h_2 + H_4 + [x_1, h_0] + R_1
\]

with \( R_1 \) the remainder. It is easy to see that \( h_2 \) is resonant, and that each monomial

\[
y^n(\tilde{k}_1, k_2, k_3, k_4, n) a_{k_1} a_{k_2} a_{k_3} a_{k_4}^* e^{i\omega},
\]

in \( H_4 \) is eliminated by a monomial

\[
y^n(3k_1 - \tilde{k}_1, k_2, k_3, k_4, n) a_{k_1} a_{k_2} a_{k_3} a_{k_4}^* e^{i\omega},
\]

in \( x_1 \), where

\[
\tilde{a}_k = \begin{cases} 0 & \text{if } |k| \leq N, \\ a_k & \text{if } |k| > N. \end{cases}
\]

The resonance conditions for the quartic terms are therefore

\[
n\Omega - \delta \tilde{n} = 0, \quad k_1 + k_2 - k_3 - k_4 = 0, \quad \tilde{f}_n(n) \neq 0, \quad k_1, \ldots, k_4, n \in \mathbb{Z}.
\]

(3.17)

with \( \tilde{\Omega} = \tilde{a}_{\tilde{k}_1} + \tilde{a}_{\tilde{k}_2} - \tilde{a}_{\tilde{k}_3} - \tilde{a}_{\tilde{k}_4} \). We will not give a full analysis of (3.17), although it is quite remarkable that we can here obtain a complete picture of the resonances. The idea is to first parameterize the level sets of \( \tilde{\Omega} \) in \( \tilde{\mathbb{Z}} \) subject to \( k_1 + k_2 - k_3 - k_4 = 0 \), and then examine the level sets of \( n\Omega - \delta \tilde{n} \) in \( \mathbb{Z} \). Instead of this we will consider a partial normalization that eliminates only a subset of the nonresonant mode interactions. The resulting "partial" normal form, denoted by \( h_0 + h_2 + N_4 \), is simple to produce and gives some interesting insights into the dynamics of the system.

The main observation is that the subspace spanned by the modes \( a_k \) with \( |k| \leq N \) can be invariant under the evolution of \( h_2 + h_4 + N_4 \). To see this we consider "low" and "high" frequency mode index sets \( U_L = \{ k \in \mathbb{Z} : |k| \leq N \} \) and \( U_H = \mathbb{Z} \times U_L \), and decompose \( \mathbb{Z}^4 \) into disjoint products of the two sets. We use the notation

\[
U_{LLL} = U_L \times U_L \times U_L \times U_L, U_{LLHL} = U_L \times U_L \times U_H \times U_L, U_{LLLL} = U_L \times U_L \times U_L \times U_L, \text{ etc., and subdivide } \mathbb{Z}^4 \text{ into 2^4 disjoint subregions. Accordingly, we also decompose } H_4 \text{ into 2^4 parts coupling different combinations of quartets of low and high index modes.}
\]

We now seek to eliminate all quartic nonresonant terms involving only low modes, and quartic nonresonant interactions involving one high frequency mode and three low frequency modes. The corresponding parts of \( H_4 \) will be denoted by \( H_{4_{LLL}} \), \( H_{4_{LLHL}} \), \( H_{4_{LHL}} \), \( H_{4_{LLL}} \) respectively. To eliminate \( H_{4_{LLL}} \), we use the function \( x_{LLL} \), given by the expression of (3.6), with summation over \( k_1, \ldots, k_4 \) in \( U_L \). The resonant part of \( H_{4_{LLL}} \) is

\[
N_{4_{LLL}} = \gamma \sum_{k_1, k_2, k_3, k_4 \in U_L} a_{k_1} a_{k_2} a_{k_3} a_{k_4}^* \tilde{f}_n(0) a_{k_1} a_{k_2} a_{k_3} a_{k_4}^* e^{i\omega},
\]

and we denote the nonresonant part of \( H_{4_{LLL}} \) by \( H_{4_{LLL}} \). To eliminate \( H_{4_{LLHL}} \), \( H_{4_{LHL}} \), \( H_{4_{LLL}} \) we examine the resonance condition (3.17) in \( U_{LLHL}, \ldots, U_{LHL} \), respectively. For instance, in \( U_{LLHL} \) (3.17) reduces to

\[
n\Omega + \delta k_4 = 0, \quad k_1 + k_2 - k_3 - k_4 = 0, \quad k_1, k_2, k_3 \in U_L, \quad k_4 \in U_H, \quad n \in \mathbb{Z}.
\]

(3.19)

with \( \tilde{f}_n(n) \neq 0 \). We see that resonant terms can only involve high modes with indices \( k_4 \in [N, 3N] \). Then

\[
|\Omega| > 10|\delta|N^2.
\]

(3.20)
Proposition 3.2. Let $\Omega$ be a periodic solution that has a variational interpretation. The dynamics on this subspace can be complicated and interesting. Of special interest are the second procedure. We examine Hamilton’s equations for $H_{1,L,L,L}$.

Proof. We have the autonomous system
\[ \dot{q}_k = -\frac{\partial}{\partial p_k} (h_2 + N_{4,L,L,L}), \quad k \in U_L, \tag{3.24} \]
and we look for solutions of the form $A_k = e^{it} \phi_k$, $k \in U_L$, with $\lambda$ real. The equation for the $A_k = q_k + ip_k$ has the structure of the constrained variational problem
\[ \frac{\lambda}{\delta q_k} \frac{\partial \mathcal{L}}{\partial q_k} = \frac{\partial \mathcal{L}}{\partial p_k}, \quad \frac{\partial \mathcal{L}}{\partial p_k} = \frac{\partial \mathcal{L}}{\partial q_k}, \quad k \in U_L. \tag{3.25} \]
with
\[ I = \sum_{k \in \mathbb{Z}} (\nu_k^2 + \nu_k^4), \quad V = h_2 + N_{kLLL}. \quad (3.26) \]

Note that \( I \) and \( V \) are both invariant under the circle action generated by the Hamiltonian flow of \( I \) and that the quotient of the sphere \( S^2 \) is \( \mathbb{CP}^{N-1} \) by the circle action is \( \mathbb{CP}^{N-1} \). On the other hand, by a proposition of Weinstein [24] and Moser [13] a smooth function on \( \mathbb{CP}^{N-1} \) must have at least \( V \) critical points.

Remark 3.3. In the argument above we can also look for solutions of (3.24) that have the more general form
\[ g_k = e^{i\varphi_k} a_k, \quad k \in \mathbb{Z}, \]
with \( a_k \) satisfying \( \sum_{j=1}^{N} + b_{k_j} - b_{k_j} = 0 \) for all \( k_1, k_2, k_3, k_4 \) in \( \mathbb{Z} \). Some choices of \( a_k \) however, for instance \( g_k = k \), do not lead to compact level sets \( I = C \), and the existence of solutions of (3.25) is not guaranteed.

We can alternatively prove Proposition 3.2 by noting that (3.24) has families of exact “monochromatic” solutions of the form \( A_0 e^{i\varphi_0} \) with \( k_0 = -N, -N+1, \ldots, N \). For \( |A_0|^2 = C \) these solutions are critical points of \( V \) on \( I = C \). At the same time some of the critical points of \( V \) are expected to coincide with the DM soliton (i.e. “bump-like”) solutions found in simulations with periodic boundary conditions (see e.g. [14,15], unpublished work in progress).

The numerical results therefore suggest that the number of families of critical points of \( V \) on the level hypersurfaces of \( I \) is greater that \( 2N + 1 \).

The rapidly decaying, Gaussian-like DM soliton solutions found numerically in the periodic problem motivate some comparisons with results on solutions for the equation on the line. The methods employed there seek solutions that decay at infinity and concern the analogue of the quartic normal form system with Hamiltonian \( H_2 + H_4 \) in (3.5) for the line (i.e. with summation over the \( k_i \) replaced by integration). In the case \( h_4 < 0 \) the DM soliton is interpreted as a minimum (for \( \varphi > 0 \), maximum for \( \varphi < 0 \) of \( H_2 + H_4 \) over functions with fixed \( L_2 \) norm (see [21]). An existence proof based on this variational characterization of the DM soliton is in [25]. The functional is extremized in the Sobolev space \( H^1(\mathbb{R}, C) \). The analogous solution here is possibly the minimum (for \( \varphi > 0 \), maximum for \( \varphi < 0 \) of \( V \) over the spheres \( I = C \) in (3.26). This scenario is consistent with numerical calculations where we see that for \( \varphi < 0 \) the value of \( V \) for the numerical DM soliton is smaller than the value of \( V \) of all the monochromatic solutions (on the same level surface of \( I \)). The assumption \( h_4 < 0 \) is crucial in the variational argument of [25], but irrelevant in Proposition 5.2, where some of the critical points of \( V \) may be related to the localized solutions shown to exist in [8] for \( h_4 \neq 0 \) (by a bifurcation argument). For \( h_4 > 0 \) our comparisons of \( V \) for the numerical DM solitons and the monochromatic waves suggest that the DM solitons are not global extrema of \( V \). It should be emphasized that this numerical work is still in progress and will be reported elsewhere. (We have also recently seen a variational argument of [9] for \( \delta = 0 \), obtained from \( h_4 \to 0^+ \), with \( \nu \) fixed. Note that the second normal form construction we presented is inapplicable for \( \delta = 0 \), while the first averaging procedure is still meaningful.)

Remark 3.4. The comparison of the solutions of (3.24) with solutions of an analogous equation on the line is natural for the DM solitons seen numerically in the periodic problem, but is heuristic. Note for example that solutions of (3.24) such as the monochromatic waves have no clear variational interpretation on the line. Following Remark 2.2, it will be interesting to study the localized solutions of (3.24) as \( |\delta| \) decreases.

Conservation of the quartic normal form Hamiltonian \( h_2 + N_{kLLL} \) and the \( L^2 \) norm on the invariant subspace \( M_k \) implies that the minima and maxima of \( V \) in Proposition 3.2 are nonlinearly stable on the invariant subspace \( M_k \). We therefore still have the possibility of high frequency instabilities for more general initial conditions. In
comparison, the variational characterization of the DM soliton solution for \( \delta \gamma < 0 \) in [21] implies nonlinear stability (see [14,25]). Note however that in (3.11) and (3.12), we can split \( H_{\delta} \) into \( H_{\delta} + h_{\delta} \) differently, by making \( N \) larger. Assuming \( |\Omega| > 10|\beta|N^2 \), i.e. assuming larger \( |\Omega| \), and following the arguments above verbatim we have an invariant subspace \( M_L \) with larger dimension. Thus, if the solutions of Proposition 3.2 are inside the domain where we expect the second normal form procedure to be valid (see Section 4, and the remark at the end of that section) possible high frequency instabilities will not be detected in spectral numerical simulations with \( K \leq \dim(M_L) \) modes.

The splitting of the quadratic Hamiltonian into low and high frequency parts can be also used for more general dispersion relations \( \omega_k \) (a cubic dispersion is considered in [14,15]). For instance, let \( N_1, M_1 > 0 \) with \( \delta M_1 \) of \( O(1) \), and assume that \( \omega_k : \mathbb{R}^+ \rightarrow \mathbb{R} \) is strictly increasing for \( k > N_1 \), diverges as \( k \rightarrow +\infty \), and is bounded by \( M_1 \) for \( k \leq N_1 \). Also extend \( \omega_k \) to be even or odd. For such a dispersion relation the arguments above apply with minor modifications, and lead to the existence of an invariant subspace with periodic and quasi-periodic solutions.

It is also straightforward to find conditions under which the subspace \( M_L \) above is invariant for higher order partial normal form systems. First, we see that we can always eliminate the angle dependent part of the Hamiltonian that couples only low modes. Also, the resonance conditions for \( 2n \)-wave interactions coupling one high mode with \( 2n-1 \) low modes have a simple form that is similar to (3.21). Note that higher order terms always couple an even number of modes, and are sums of monomials of the form \( a_1 \cdots a_{2n-1} \omega_{k_1}^a \cdots \omega_{k_{2n-1}}^a \) with \( k_1 + \cdots + k_n - k_{n+1} - \cdots - k_{2n} = 0 \). A resonant monomial with indices \( k_1, \ldots, k_{2n-1} \in U_{/>}, k_n \in U_{/\leq} \) must satisfy \( 2\Omega - \delta k_n^2 = 0 \), and by \( |k_n| \leq (2n-1)N \), we can avoid resonances and small divisors by requiring that \( |\Omega| \approx \beta (2n-1)^2 N^{2+1} \). Similar considerations apply to all terms coupling \( 2n-1 \) low modes with one high mode. We thus see that the formal argument is the same for higher orders, but \( |\Omega| \) must be assumed larger, increasing quadratically in the order of the normal form. Alternatively, resonant interactions between \( 2n-1 \) low modes and one high mode can be avoided by assuming that \( \Omega/\delta \) is irrational. In that case one must however examine possible small divisors.

4. Error estimates for the normal form equations

In this section we estimate the distance between solutions of the full system (2.1), (2.4) and the quartic normal form equation by making the formal calculations of the previous section rigorous.

To state the error estimates for the quartic (partial) normal form equation we write the full system (2.4) as

\[
\dot{u} = L_0 u + F(u, t), \quad u(0) = u_0
\]  

(4.1)

with \( L = i\delta a_{\Omega, t} \), and \( F(a, t) \) the nonlinearity. We also write Hamilton’s equation for the quartic normal form Hamiltonian \( h_0 + h_2 + h_4 \) in (3.22) as

\[
\dot{h} = Lh + \tilde{G}(h, t), \quad h(0) = h_0.
\]  

(4.2)

We are here writing the normal form equation as a nonautonomous system; the action component is omitted, and \( t = \Omega^{-1} \phi \). The two equations can be considered in the Sobolev spaces \( H^s, s > 1/2 \), of complex valued \( 2\pi \)-periodic functions. The norm of a function \( u \) in \( H^s \) will be

\[
|u|^2 = \sum_{k \in \mathbb{Z}} (1 + |k|^s)|u_k|^2
\]  

(4.3)

with \( u_k \) the Fourier coefficients of \( u \). First, we have the following basic local existence theorem.
Proposition 4.1. Let $s > 1/2$, $\beta > 0$, with $\beta \sim O(1)$, and assume that $d(t)$ is locally absolutely integrable. Consider the initial value problems of (4.1) and (4.2) with initial conditions satisfying $\|a_k\|_s, \|b_k\|_s \leq O(1)$. Then for $|y| > 0$ there exists a positive constant $C = C(\|a_k\|_s, \|b_k\|_s, \beta) \sim O(1)$ and a time $t_1 \geq C|y|^{-1}$ for which (4.1), (4.2) have unique solutions $a(t), b(t) \in R(t_1, y_0, \beta)$, where $R(t_1, y_0, \beta) = \{y(t) \in C^2([0, t_1], H^s): \|y(t) - y_0\|_s \leq \beta\}$.

Remark 4.3. The constants $C_{\phi}$ are obtained readily following the arguments of Lemma 4.4 below. Since the Lipschitz constant $L_F$ is precisely the one for the map $u \mapsto |y|^2 u$ in the cubic nonlinear Schrödinger equation. Similar considerations apply to the map $G$ (4.2). Lipschitz constants for the part of $F$ that is eliminated by the canonical transformation $\Phi^I_{s}$ are obtained readily following the arguments of Lemma 4.4 below. Since the Lipschitz constants for $F$ and $G$ and the size of the initial conditions of (4.1), (4.2) are close, we choose for convenience to state the local existence theorem for the two initial value problems with the same constants $C, \beta$. We now state our estimate for the distance between the solutions of the full system (4.1) and the quartic normal form Eq. (4.2).

Theorem 4.2. Let $s > 1$, and let $0 < |\beta| \leq c_3|y|$ for some $c_3 < O(1)$. Also assume that $|\bar{\Omega}| > 0|\Upsilon^1| N$ and that the (dilated) periodic dispersion management function $d(t)\Omega^{-1}$ is in $L^2$. Consider a solution $b(t)$ of (4.2) with initial condition $\|b_k\|_s \leq \rho$, with $x = s + 1$, $\rho_0 \sim O(1)$, and a solution $a(t)$ of (4.1) with initial condition $a_0 = b_0$. Then there exist constants $C_0, C_1$ of $O(1)$, and $y_0$ for which $|y| \leq y_0$ implies

$$\|a(t) - b(t)\|_s \leq C_0|y|, \quad \forall t \in [0, C_3|y|^{-1}]$$

(4.5)

The constants $C_0, C_1, y_0$ depend on $x, |\bar{\Omega}|, c_3$, and $\rho_0$.

Remark 4.3. The constants $C_0, C_1, y_0$ do not change significantly as $|\bar{\Omega}|$ diverges. This is seen in Lemma 4.4 below. Also, the constants are independent of the amplitude $b(t)$ of the dispersion management function $d(t)$. Thus the estimate applies to all the physically interesting parameter ranges of $|\bar{\Omega}|$ and $b(t)$ discussed in Remark 2.1, but there are no ranges leading to significant improvements of the error estimate.

The main ingredients of the proof of Theorem 4.2 are estimates for the canonical transformation $\Phi^I_{s}$ defined formally in the previous section, and a bound for the Hamiltonian vector field of the remainder $R_1$ in (3.22). We start by making the Hamiltonian structure of (4.1) precise, and proceed with an outline of the argument.

We consider the Sobolev spaces $H^s$ of $2\pi$-periodic complex valued functions, viewed as real Hilbert spaces with the inner product

$$\langle u, v \rangle_s = \Re \sum_{k \in \mathbb{Z}} (1 + |k|^2) u_k \bar{v}_k$$

(4.6)

where $u_k, v_k$ denote the Fourier coefficients of $u, v$, respectively. The norm in $H^s$ is given by (4.3), while the ball of radius $\rho$ around the origin is denoted by $B^s(\rho)$. We also let $|u, v| = |\langle u, v \rangle|$. Functions on $H^s$ can be extended
the complexification \( H' \) of \( H \) by letting the real and imaginary parts of the Fourier coefficients \( a_k \) of \( u \in H \) become complex.

For \( f: B'(\rho) \to \mathbb{R} \) Fréchet \( C^1 \) smooth in \( B'(\rho) \), we define the gradient map \( \nabla f : B'(\rho) \to H' \) by \( \langle \nabla f(u), v \rangle = Df(u)v \), with \( D \) the Fréchet derivative, and \( v \in H' \). Also, \( \mathcal{J} = -i \) defines a symplectic structure in \( H' \), and we denote the Hamiltonian vector field \( \mathcal{J} f \) of a \( C^1 \) function \( f \) by \( \mathcal{V}f \). The time-dependent Hamiltonian for the system of (2.4) is \( h(a, t) = H(a, 2t, J) + \mathcal{J} f \), with \( H \) as in (2.8). If \( d(t) \) is integrable, \( h : H' \times \mathbb{R} \to \mathbb{R}, \) \( r \geq 1 \), is \( C^1 \) in the variable \( a \), for all \( r \in \mathbb{R} \), and (4.1) can be written as the nonautonomous Hamiltonian system

\[
\dot{u} = \mathcal{J} h(a, t) = V_h(a, t).
\]

(4.7)

The extended phase space will be the set of \( x = [x_1, x_2, J] \in H' \times \mathbb{R} \times \mathbb{R} \), where the second and third components correspond to the angle \( \phi \) and action \( J \), respectively (the angle is defined on the covering space of the circle). The spaces \( H' \times \mathbb{R} \) can be considered as real Hilbert spaces with the inner product

\[
\langle \xi, \eta \rangle_{H', \mathbb{R}} = \langle \xi_1, \eta_1 \rangle_{H'} + \langle \xi_2, \eta_2 \rangle_{\mathbb{R}}.
\]

(4.8)

The norm in \( H' \times \mathbb{R}^2 \) is denoted by \( \| \xi \|_{H', \mathbb{R}} \), and we let \( K'(\rho) = B'(\rho) \times \mathbb{R}^2 \). As above, we use the inner product in \( H' \times \mathbb{R}^2 \) to define the gradient \( \nabla \) for smooth functions on \( H' \times \mathbb{R}^2 \). The symplectic structure on \( H' \times \mathbb{R}^2 \), denoted by \( \mathcal{J} \), will be the tensor product of \( \mathcal{J} \) with the standard symplectic structure in \( \mathbb{R}^2 \), i.e. compare with the Poisson bracket in (2.6). Also, we let \( \mathcal{V} = \mathcal{J} \nabla \). If the dispersion management function \( d(t) \) is \( C^1 \) then the Hamiltonian \( H = H(a, \phi, J) \) of (2.8) is \( C^1 \) in \( H' \times \mathbb{R}^2, \) \( r \geq 1 \), and Hamilton’s equations (2.7) can be written as

\[
\dot{x} = \mathcal{J} H_1(x) = \mathcal{V}_1(x).
\]

(4.9)

Following the previous section, we want to construct a one-parameter family of canonical transformations \( \Phi_{J_1} = T \) by integrating

\[
\frac{d}{dt}T^i(y) = \mathcal{V}_{J_1} (T^i(y)), \quad T^0(y) = y
\]

(4.10)

with \( J_1 \) as in (3.25). The components of \( T^i \) in \( H' \times \mathbb{R} \times \mathbb{R} \) will be denoted by \( T^i, i = 1, 2, 3 \). Since \( J_1 \) does not depend on the action \( J_2 \), the second component of the equation is integrated trivially yielding \( T^2_0(y_3) = y_2 \), for all \( e \), i.e. the angle variable does not change. The existence of solutions to (4.10) for \( |y| \) sufficiently small and appropriate initial conditions is shown in Lemma 4.4 below, where we also show that \( T^1_0(y) \) and \( y_1 \) are \( O(|y|) \) close in \( H' \). To exhibit the normal form equation and the remainder we write the Hamiltonian system of (4.9) as

\[
\dot{z} = \tilde{L} z + \tilde{V}_H(z)
\]

(4.11)

with \( \tilde{L} z = \tilde{V}_{h_0 + J_2}(z) \). Eq. (4.11) is then written in the new variable \( y, x = T^i(y), \) as

\[
\dot{y} = \tilde{L} y + \tilde{V}_g(y), \quad \dot{x} = \tilde{V}_h(y).
\]

(4.12)

where we are using the notation of (3.26) for the Hamiltonian \( H \circ T^{-1} = H \circ \Phi_1 \), and the fact that \( T^i \) is symplectic. The remainder \( R_1 \) of (3.26) is

\[
R_1 = (h_2 \circ T^1 - h_2) + (H_0 \circ T^1 - H_0) + (h_0 \circ T^1 - h_0 - [J_1, h_0]),
\]

(4.13)

and the quartic normal form system is

\[
\dot{z} = \tilde{z} \dot{z} + \tilde{V}_H(z)
\]

(4.14)

An elegant way of comparing solutions of the normal form and the full system is to consider the evolution of the “error” \( \dot{r} = x(\tau) - T^i(\tau(t)) \) (see [3]). In particular, combining (4.11), (4.12) and (4.14) we obtain

\[
\dot{r} = \tilde{L} r + (\tilde{V}_{h_1}(x) - \tilde{V}_{h}(T^i(z))) + [DT^i(z)] \dot{V}_h(z).
\]

(4.15)
Proof. We first consider the equation for the rate of change of Lemma 4.4. Let $\epsilon = T^1(\epsilon(0))$. Note that we are interested in the size of the first component of $T^1(\epsilon(0))$. We will denote the three components of $V_1$ in $H^1 \times \mathbb{R} \times \mathbb{R}$ by $V_{1i}$, $i = 1, 2, 3$. Also, we let $D_{1i}$, $D_2$, $D_3$ denote the Fréchet derivatives along each of the three components in $H^1 \times \mathbb{R} \times \mathbb{R}$. Observe that $R_1(z)$ depends on the first two components of $z$ only. This is clear for the first two terms of $R_1$ in (4.13) since $h_0$ and $H_0$ depend on the first two components of $H^1 \times \mathbb{R} \times \mathbb{R}$ only, and $T^1_2(\epsilon) = T^1_2(\epsilon_1, \epsilon_2), T^1_2(\epsilon) = \tilde{z}_2$. For the third term in (4.13), the observation follows from the cohomology equation $[\chi, h_0] = G_4$, with $G_4 = N_H - H_4$, and

$$h(0 \circ T^1 - h_0)(y) = \int_0^1 \frac{d}{ds} h_0(T^1(y)) \, ds = \int_0^1 G_4(T^1(y)) \, ds. \tag{4.16}$$

The second component of $\tilde{V}_{R_1}$ and $D_2T^1_2(\epsilon)$ therefore vanish, and the first component of (4.15) becomes

$$r_1 = L_{11} + (V_{1i}(y) - V_{1i}(T^1(z))) + [D_2T^1_2(\epsilon)]_{11}V_{1i}(\epsilon) \tag{4.17}$$

with $L_{11} = V_{1i}(r_1(t))$. The vector field $[D_2T^1_2(\epsilon)]_{11}V_{1i}(\epsilon)$ is estimated using Lemmas 4.5 and 4.7 below, and we will conclude the proof by estimating the size of $r_1(t)$ from (4.17). We start by solving (4.10) and showing that the first component of $T^1$ is near-identity.

Lemma 4.4. Let $\rho > 0$, $a_1 \in (0, 1)$, and set $\rho_1 = \rho a_1$. Assume that $\varepsilon > 1/2, |\Omega| > 10||\nu||^2$, and that $\delta = a_1^{-3} \in H^0$. Then for $|\varepsilon| \leq 1, \varepsilon \in (0, \rho_1)$, the initial value problem (4.10) with $y \in K(\rho_1)$ has a unique solution $T^1(y), \varepsilon \in [0, 1]$, and its flow defines a one-parameter family of canonical transformations $T^1 : K(\rho_1) \to K(\rho)$. Moreover, we have

$$\sup_{y \in K(\rho_1)} \|T^1(y) - y\|_1 \leq e C_1|\varepsilon|, \quad \varepsilon \in [0, 1] \tag{4.18}$$

with $C_1 = \rho^3 C_1(\varepsilon, |\Omega|)$ a constant of $O(1)$ for $\rho$ of $O(1)$.

Proof. We first consider the equation for the rate of change of $T^1$. The right hand side is $T^1 = -J\nabla x_1$. Since $x_1$ is independent of the action $y_1$ and the angle $y_2$ does not change, we can consider $x_2$ as a parameter. We will show that $y_{2i}$ is real analytic in $y_1$, uniformly in $y_2$. Denote the Fourier coefficients of $y_1$ by $a_i$, and let $(V_{1i})_k$ be the $k$th Fourier component of the vector field $V_{1i}$. The definition of $x_1$ in the previous section leads readily to explicit expressions, and we have

$$|V_{x_1, \Omega \Omega \Omega \Omega}: | \leq 2|\varepsilon| \sum_{k_1, k_2 \in \Omega \setminus \Omega} |a_i a_2 a_3| |\delta_{k_1+k_2-4} - 4| \frac{F_{\Omega}(\phi)}{12}, \quad k \in U_L \tag{4.19}$$

with

$$F_{\Omega}(\phi) = \sum_{n \in \Omega \setminus \Omega} \frac{f_n(\phi)}{n^3} e^{i n \phi}. \tag{4.20}$$

Note that $|V_{x_1, \Omega \Omega \Omega \Omega}| = 0$ for $k \notin U_L$. Also

$$|V_{x_1, \Omega \Omega \Omega \Omega}| \leq |\varepsilon| \sum_{k_1, k_2 \notin \Omega \setminus \Omega} |a_i a_2 a_3| |\delta_{k_1+k_2-4} - 4| \frac{f_n(0)}{12} + \frac{G_n(\phi, k)}{12}. \tag{4.21}$$

if $k \in U_L$, and

$$|V_{x_1, \Omega \Omega \Omega \Omega}| \leq |\varepsilon| \sum_{k_1, k_2 \notin \Omega \setminus \Omega} |a_i a_2 a_3| |\delta_{k_1+k_2-4} - 4| \frac{f_n(0)}{12} + \frac{G_n(\phi, k)}{12}. \tag{4.22}$$
if \( |k| \in [N, 3N]\), where

\[
G_m(\phi, k) = \mathcal{D} \sum_{n \in \mathbb{Z}} \frac{f_m(n)}{n + k} e^{in\phi},
\]

(4.23)

All other Fourier components of \( V_{x_1, x_2, L} \) vanish. We also have similar estimates for the symplectic gradients of \( x_1, x_2, L, L \), and \( x_1, L, L \). To bound \( F_m(\phi) \) we note that

\[
|F_m(\phi) - F_m(\phi_2)| \leq \sqrt{2\pi} |e^{-iak\phi}| = 2\pi, \tag{4.24}
\]

for all \( \phi, \phi_2 \in [0, 2\pi] \) (with prime the derivative with respect to \( \phi \)). By \( |2\pi| \geq 98N^2 \) and \( |k| \in [N, 3N] \) we see that \( |1 + (k^2/n^2)|^{-1} \geq 2 \), and we similarly obtain \( |G_m(\phi, k) - G_m(\phi, k)| \leq 4\pi, \) for all \( \phi, \phi_2 \in [0, 2\pi] \) and all admissible \( k, m \). Since the functions \( F_m \) and \( G_m \) are continuous in \( \phi \) and have zero average, we therefore have

\[
|F_m(\phi)| \leq 4\pi, \quad |G_m(\phi, k)| \leq 8\pi, \quad \forall \phi \in [0, 2\pi], \tag{4.25}
\]

and all admissible \( k, m \). In (4.21) we also note that \( |\delta| \leq 1/2 \) implies \( |f_m(0)||k^2| \leq 2 \), uniformly in \( m \in \mathbb{Z} \) and \( k \in [N, 3N] \).

Estimating the discrete convolutions in (4.19), (4.21) and (4.22), and in similar expressions for the symplectic gradients of \( x_1, L, L, L, L \), and \( x_1, L, L \), and using similar bounds on the coefficients \( F_m, G_m \), above, we therefore have

\[
\|V_{x_1, y_2}\| \leq C|\gamma|^{1/2}, \quad x > \frac{1}{2} \tag{4.26}
\]

with a constant \( C \) of \( O(1) \) that depends on \( x \) and \( |2\pi| \) and is decreasing in \( |2\pi| \). Complexifying the real and imaginary parts of Fourier coefficients \( a_k \) and \( a_k^* \) (with the obvious abuse of notation), we obtain a similar inequality in \( H^s, x > 1/2 \) with a slightly larger constant \( C \) of \( O(1) \) that depends on \( x \) and \( |2\pi| \). It is also immediate from the convolution estimates that the complexification of \( V_{x_1, y_2} \) is bounded and weakly analytic, and therefore analytic (see [18], Appendix A). The estimates are also uniform in the parameter \( y_2 \). Thus, given \( \rho, a_1 \) as in the statement we can choose \( |\gamma| \) sufficiently small so that the integral curves \( T^\gamma_{\rho}(y_1, y_2) \) exist for all \( y_1 \in B'(\rho), \epsilon \in [0, 1] \) and \( y_2 \in \mathbb{R} \). Note that \( T^\gamma_{\rho}(y_1, y_2) \) is also real analytic in \( y_1 \) and \( \epsilon \), for all \( y_2 \in \mathbb{R} \) (see e.g. [6, Chapter 1]).

The distance between \( T(y_1, y_2) \) and \( y_1 \) can be estimated using

\[
T^\gamma_{\rho}(y_1, y_2) = y_1 + \int_0^\epsilon V_\gamma(\tilde{T}^\gamma_{\rho}(y_1, y_2), y_2) dt, \quad y_2 \in \mathbb{R}. \tag{4.27}
\]

We easily see from (4.27) and (4.26) that \( |\gamma| \leq (1 - \alpha_1)/C\rho^2 \), with \( C \) as in (4.26), implies that if \( y_1 \in B'(\rho) \) then \( T^\gamma_{\rho}(y_1, y_2) \) stays in \( B'(\rho) \) for all \( \epsilon \in (0, 1] \) and \( y_2 \). Using (4.27) again we obtain (4.18).

The third component of (4.10) is

\[
\frac{d}{d\epsilon} T^\gamma_{\rho}(y_1, y_2) = -(D_{x_1, y_2}) T^\gamma_{\rho}(y_1, y_2), \quad T^\gamma_{\rho}(0, y) = y_1, \tag{4.28}
\]

and it suffices to integrate the right hand side with respect to \( \epsilon \). We have

\[
|D_{x_1, y_2} T^\gamma_{\rho}| \leq |\gamma| \sum_{k_1, k_2, k_3, k_4, \epsilon \in \mathbb{Z}} |a_{k_1} a_{k_2} a_{k_3} a_{k_4}| |G_{m}(\phi, k_4)|, \tag{4.29}
\]

\[
|D_{x_1, y_2} T^\gamma_{\rho}| \leq |\gamma| \sum_{k_1, k_2, k_3, k_4, \epsilon \in \mathbb{Z}} |a_{k_1} a_{k_2} a_{k_3} a_{k_4}| |G_{m}(\phi, k_4)| \tag{4.30}
\]

with prime the derivative with respect to \( \phi \). For \( \Lambda(\phi) \in H^s \) with \( s > 1/2 \), the series \( F_m(\phi) \) is absolutely convergent and \( |F_m(\phi)| \) is bounded uniformly in \( m \) since \( m \) takes a finite number of values in (4.29). Similarly, \( |2\pi| > 10|\delta|N^2 \)
Proof. By \( \gamma \) and \( | \gamma \rangle \) and Lemma 4.4 differentiability of the maps \( \tilde{\text{C}} \), \( \text{derivatives of} \quad \tilde{\text{C}} \), \( \epsilon \) the right hand side of (4.28) is therefore real analytic in \( \epsilon \), for all \( y_1 \in B'(\rho_1) \), and \( y_2 \in \mathbb{R} \). 

We now consider the derivative of the canonical transformations \( T' \). We are especially interested in the derivative along the \( H' \) direction. In what follows \( \| A \|_{\epsilon,\rho} \) denotes the operator norm of a linear map \( A : H' \to H' \), and \( I_1 \) is the identity in \( H' \).

**Lemma 4.5.** Let \( \rho, \rho_1 \), and \( s \) as in the Lemma 4.4 and assume that \( A(\phi) \) is in \( H^q, q > 5/2 \), and that \( | \gamma \rangle \leq \gamma \leq 5/2 \). Then, the canonical transformations \( T' : K'(\rho_1) \to H' \) defined above are Fréchet differentiable and we have

\[
\sup_{y \in K'(\rho_1)} \| (D_1 T'_1) y \|_{L_2} \leq 1 + C_2 | | \gamma \rangle |.
\] (4.31)

Also, the maps \( D_1 T'_1(y), y \in K'(\rho_1) \) in (4.31) are invertible and satisfy

\[
\sup_{y \in K'(\rho_1)} \| (D_1 T'_1)^{-1} y \|_{L_2} \leq 1 + C_2 | | \gamma \rangle |.
\] (4.32)

\[
\sup_{y \in K'(\rho_1)} \| (D_1 T'_1)^{-1} y - I_1 \|_{L_2} \leq C_2 | | \gamma \rangle |.
\] (4.33)

The constant \( C_2 = C_2(s, | | \gamma \rangle |, \rho) \) is of \( O(1) \) for \( \rho \) of \( O(1) \).

**Proof.** By

\[
T'_1(y) = y + \int_0^1 V_{-\gamma}(T'(y)) \, d\gamma
\] (4.34)

and Lemma 4.4 differentiability of the maps \( T'(y) \) with \( y \in K'(\rho_1) \), will follow from the differentiability of the vector field \( V_{-\gamma}(y) \) for \( y \in K'(\rho) \). Note that \( D_2 \mathcal{C}^p_{-\gamma} = 1 \), and that the only other nonvanishing partial derivatives, namely \( D_1 \mathcal{C}^p_{-\gamma}, D_2 \mathcal{C}^p_{-\gamma}, D_3 \mathcal{C}^p_{-\gamma} \), and \( D_4 \mathcal{C}^p_{-\gamma} \) of \( V_{-\gamma} \) involve up to two derivatives (in \( y \)) of the functions \( F_{\alpha}(\phi), G_{\alpha}(\phi, k) \) in Lemma 4.4. The condition on \( A(\phi) \) on the other hand implies that the series for the third derivatives of \( F_{\alpha}(\phi), G_{\alpha}(\phi, k) \) are absolutely convergent for all admissible \( m, k \). Thus the above four partial derivatives of \( V_{-\gamma} \) are well-defined and continuous in \( y_2 \). Continuity of the four partial derivatives in \( y_1 \) comes from the real analyticity of \( x_1 \), and the existence of \( \mathcal{D} V_{-\gamma}(y) \) follows immediately. Letting \( | | \gamma \rangle | \) sufficiently small and integrating \( V_{-\gamma} \), as in Lemma 4.4 we also obtain differentiable transformations \( T'' \), \( \epsilon \in [0, 1] \). Moreover, we have \( T''(T'(y)) = T'(T''(y)) = y \), therefore \( (DT')^{-1}(T'(y)) = (DT'')^{-1}(T''(y)) \). Viewing \( (DT')^{-1}(y) \) and its inverse as \( 3 \times 3 \times 2 \) block matrices applied to vectors in \( H' \times \mathbb{R} \times \mathbb{R} \), we see that the 1, 1 entry \( \| (DT')^{-1}(y) \|_{L_2} \) must equal \( \| (D_1 T'_1)^{-1}(y) \|_{L_2} \).

Also note that

\[
T_{-\gamma}'(y) = y + \int_0^1 V_{-\gamma}(y) \, d\gamma
\] (4.35)

and

\[
\| D_1 V_{-\gamma}(y) \|_{L_2} \leq C | | \gamma \rangle | \| V_{-\gamma} \|_{L_2}, \quad s > 0
\] (4.36)

with \( C = C(s, \| D \|) \) of \( O(1) \). Differentiating (4.35) with respect to the first component, taking \( H' \) norms, and using Grönwall’s inequality we then obtain

\[
\sup_{y \in K'(\rho_1)} \| (D_1 T'_1)^{-1}(y) \|_{L_2} \leq (1 + C e^{C t}),
\] (4.37)
Remark 4.6. Note that the size of the partial derivatives $D_i V_{T_1}^{j}$, $D_i V_{T_2}^{j}$, $D_i V_{T_3}^{j}$ may be large.

Next we estimate the Hamiltonian vector field of the remainder $R_1$ in (4.13).

Lemma 4.7. Let $s$, $\rho$, $\alpha$, $\gamma$, and $\Lambda(y)$ as in Lemma 4.5. Consider the transformations $T^s$ above, and the remainder $R_1$ of (4.13). Also, let $s' = s + 2$. Then

$$\sup_{y \in K'(\rho)} \| V_{R_1}(y) \|_1 \leq C(y) \gamma^3$$

(4.40)

with $C_3 = C_3(s, [\Omega], [\Delta], [\alpha], [\gamma])$ a constant of $O(1)$.

Proof. The vector field $V_{R_1}$ has three parts, each corresponding to the three parentheses in (4.13). We first consider the terms involving $h_2$ and $H_4$. Let $f$ be Fréchet $C^1$ in $K'(\rho)$, and consider $y \in K'(\rho)$. By Lemma 4.4 and the fact that the transformations $T^s$ are symplectic we have

$$\tilde{V}_{f, T^s}(y) = [(D_{T^s})^{-1}(y) \tilde{V}_f(T^s(y)), \quad \epsilon \in [0, 1].$$

(4.41)

Observe that $D_{T^s}^{-1}(y)$ vanishes since $\chi_1(y) = \chi_1(y_1, y_2)$. Then, if $\tilde{V}_f$ vanishes, (4.41) becomes

$$\tilde{V}_{f, T^s}(y) = [(D_{T^s})^{-1}(y) \tilde{V}_f(T^s(y)),$$

(4.42)

using $[T(D_{T^s})^{-1}(y)]_{1,1} = (D_{T^s})^{-1}(y)$ (see the proof of Lemma 4.5). We write (4.42) as

$$(V_{f, T^s} - V_f)(y) = [(D_{T^s})^{-1}(y) (V_f(T^s(y)) - V_f(y))] + [(D_{T^s})^{-1}(y) - I] V_f(y).$$

(4.43)

Restricting $y$ to $K'(\rho)$ with $s' \geq s$, taking $H^s$ norms, and using (4.32) and (4.33) in Lemma 4.5, (4.43) yields

$$\| (V_{f, T^s} - V_f)(y) \|_s \leq (1 + C_2) \gamma \sup_{y \in K'(\rho)} \| V_f(T^s(y)) - V_f(y) \|_s + C_2 \gamma \| V_f(y) \|_s,$$

(4.44)

for all $y \in K'(\rho)$, $s' \geq s$.

We now consider the cases where $f = h_2$ and $H_4$, respectively. Note that $V_{h_2}$ and $V_{H_4}$ vanish. The vector field $V_{h_2}$, viewed as a function from $H^s$ to $H^s$, is uniformly Lipschitz, with Lipschitz constant $|\beta|$. The estimate for the distance between $T^s_{1}(y)$ and $\gamma_1$ in Lemma 4.4, (4.44) with $f = h_2$ yield

$$\sup_{y \in K'(\rho)} \| (V_{h_2, T^s} - V_{h_2})(y) \|_s \leq (1 + C_2) \gamma |\beta| \| C_1 \gamma \| | \beta | \rho_1.$$  

(4.45)

Also, $V_{h_2}(y_1, y_2) : B(\rho_1) \times \mathbb{R} \to H^s$, $s > 1/2$, is Lipschitz in $y_2$, uniformly in $y_1$. Moreover, if $\rho_1$ is of $O(1)$ then the Lipschitz constants are bounded by a quantity of $O(\gamma)$. Combining (4.44) for $f = H_4$ with Lemma 3.4, we obtain

$$\sup_{y \in K'(\rho)} \| (V_{h_2} \circ T^s - V_{h_2})(y) \|_s \leq (1 + C_2) \gamma |\beta| \| C_1 \gamma \| | \beta | \rho_1.$$  

(4.46)

with $C(s, \rho_1)$ of $O(1)$. 

\[ \text{sup} \quad \| (D_i T_i^s)_{1,1} - I \|_s \leq \epsilon G (1 + e^G e^{iG}) \quad (4.38) \]

with

$$G \leq \sup_{y \in K'(\rho)} \| V_{L_i}(y) \|_s \leq C(s, [\Omega]) \gamma \| \gamma \|_s \quad (4.39)$$

i.e. (4.32) and (4.33). Estimate (4.31) follows from similar arguments. \[ \square \]
To estimate the third term in $V_R$, let $G_4 = N_4 - H_4$ and recall that $[x_1, h_4] = \tilde{G}_4$. Using (4.16) and (4.41) with $f = G_4$, we have

$$V_{G_4}^2 = \int_0^1 \left( (\partial T^{y = 1}(y)) \tilde{G}_4(T^y(y)) - \tilde{G}_4(y) \right) \ dx.$$  

Since $V_{G_4}^2$ vanishes, we can argue as before that the first component of (4.47) is

$$V_{G_4}^2 = \int_0^1 \left( (\partial T^{y = 1}(y)) \tilde{G}_4(T^y(y)) - \tilde{G}_4(y) \right) \ dx.$$  

Splittting the integrand as in (4.43), and using the derivative estimates from Lemma 4.5, (4.48) leads to

$$\|V_{G_4}^2\|_{L^2} \leq \int_0^1 c_2\|\tilde{G}_4(T^y(y)) - \tilde{G}_4(y)\| \ dx + \int_0^1 c_4\|\tilde{G}_4(y)\| \ dx$$

for all $y \in K_4'$, arguing as for $V_{R_4}$ above, $V_{G_4}(y_1, y_2) : B^0(\rho_4) \times \mathbb{R} \rightarrow H'$ is Lipschitz in $y_1$, uniformly in $y_2$, with Lipschitz constants bounded by a quantity of $O(\rho_4)$. Combining this with Lemma 4.4, (4.49) implies

$$\sup_{y \in K_4'} \|V_{G_4}(y_1, y_2)\| \leq \int_0^1 C(\rho_4)^2 \ dx$$

with $C = C(s, \rho_4)$ of (4.1). Collecting the estimates (4.44), (4.46), (4.50) for the three parts of $V_R$, and using $|\rho| \leq c_2|\rho|_i$ in (4.45), we therefore have the statement. 

**Proof of Theorem 4.2.** Fix $s \geq 1$, and assume that $d(\Omega T^{-1}) \in H^s$, $q > 3/2$. We want to estimate the size of $r_1(t)$ in (4.17). The initial condition will be $r_1(0) = z_1(0) - T_1(z(0))$. Let $s = s + 2$, and consider Proposition 4.1 with $\rho_0 = \beta$ of (4.1). Applying Proposition 4.1 to the first component of the quartic normal form Eq. (4.14), we can choose $|\rho|$ sufficiently small so that $z_1(0) \in B^0(\rho_0)$ implies $z_1(t) \in B^0(2\rho_0)$ for a time interval of $O(|\rho|^{-2})$. We also have existence in $H'$ for the full system (4.1) with the initial condition $s_1(0) = z_1(0) \in H'$, also over a time interval of $O(|\rho|^{-3})$. In particular, the $H'$ norm of $s_1(t)$ remains of $O(1)$ over that time. We set $\rho = 2\rho_0$, $\alpha_1 = 3\rho_0$, and let $|\rho| \leq \rho_0$ so that Lemmas 4.4 and 4.5 apply. Thus $\tilde{T}_1^i(z_1(t), \Omega)$ is well-defined over the time interval of the local existence theory. Define $\mathcal{R}_1(t)$ by $\mathcal{R}_1(t) = e^{\theta t} \mathcal{R}_1(t)$ and note that the operator $U_t = e^{\theta t}$ is an isometry in $H'$ and commutes with $L$ (in a dense subset of $H'$). By (4.17), $\mathcal{R}_1(t)$ then satisfies

$$\mathcal{R}_1 = U_t(\mathcal{R}(z(s)) - H_k(T_1(z))) + U_t[D_1 T_1^i(z)]H_k(z).$$

(5.2)

Note that $\mathcal{R}_1(y_1, y_2) : B^0(\rho_0') \times \mathbb{R} \rightarrow H'$, $\rho' = 1/2$, is Lipschitz continuous in $y_1$, uniformly in $y_2$, and that for $\rho'$ of (4.1) the Lipschitz constants are of $O(\rho_0')$, e.g. see (4.4). From (5.2) we then have that for $t$ in the time interval of the local existence theorem we have

$$\|\mathcal{R}_1(t)\| \leq \|\mathcal{R}_1(0)\| + \int_0^t C|\mathcal{R}_1(t)| \ dx + \int_0^t \|D_1 T_1^i(z(t))\| \ dx$$

(5.4)

with $C = C(s, \rho_0)$ of (4.1). Bounding $\mathcal{R}_1(0)$, $D_1 T_1^i$ and $\mathcal{R}_k$, by Lemmas 4.4, 4.5 and 4.7, respectively, and using Gronwall’s inequality, (5.4) yields

$$|\mathcal{R}_1(t)| \leq C_0|\rho|^{-1}, \quad t \in [0, C_0|\rho|^{-1}]$$

(5.3)

with $C_0, \tilde{C}_0$ that depend on $s, |\rho_0|, \rho_0', c_2$ and are of $O(1)$.

To extend the error estimate to $d(\Omega T^{-1}) \in L^2$, consider two dispersion management functions $d(\Omega T^{-1})$ in $H^s$, $q > 3/2$, $d_1(\Omega T^{-1})$ in $L^2$ with the same average $\delta$ (and period $T = 2\pi/12$). Also let $\Lambda_1(t) = \int_0^t d_1(t) \ dx$. 

(following the notation of (2.4)). The nonlinearities of (4.1) and (4.2) with \( \tilde{A}(t) \) replaced by \( \tilde{A}_1(t) \) are denoted by \( F_1, G_1 \), respectively, and we compute the solutions of \( \dot{a}_1 = L a_1 + F_1(a_1, t) \) and \( \dot{b}_1 = L b_1 + G_1(b_1, t) \), with the initial conditions \( a_1(0) = b_1(0) = a(0) = b(0) \in H^s, s = s + 2, \) as above) to solutions of (4.1) and (4.2). Define \( \Delta \)

\[
\Delta = \sqrt{T} |\dot{a}(t) - \dot{a}_1(t)| \geq |A(t) - \tilde{A}_1(t)|. \tag{4.54}
\]

Choosing \( dt(2^{-1}) \) that is close to \( dt(2^{-1}) \), i.e. letting \( \Delta > 0 \) small, we observe that (4.53) is uniform in \( \Delta \). This is because the estimates for \( \partial' \) and \( D \) only require \( dt(2^{-1}) \in L^2 \), and the extra derivatives were used to make \( DT^* \) well-defined, but do not appear in the quantities involved in (4.53). From

\[
\|a_t(t) - b_1(t)\|_s \leq \|a(t) - a_1(t)\|_s + \|a(t) - b(t)\|_s + \|b(t) - b_1(t)\|_s, \tag{4.55}
\]

it then suffices to check that by choosing \( \Delta > 0 \) sufficiently small the first and third terms of (4.55) can be made of \( O(|\gamma|^1) \) over a time interval of \( O(|\gamma|^1) \). To see this, we let \( U_t = e^{t\tilde{A}}, A(t) = U_{\gamma}a(t), A_1(t) = U_{\gamma}a_1(t) \), and use (4.54) to obtain

\[
|F(U_t A(t), t) - F(U_t A_1(t), t)| \leq C_F (\|A(t) - A_1(t)\|_s + C\Delta |A(t)|_s) \tag{4.56}
\]

with \( C, F \) a constant that is quadratic in \( |A|_s, |A_1|_s, \) and \( C \) of (1). The \( H^s \) norms of \( A \) and \( A_1 \) remain of \( O(1) \) over a time of \( O(|\gamma|^1) \), and by (4.56) we see that

\[
\|a(t) - a_1(t)\|_s \leq O(\Delta) \tag{4.57}
\]

for a time of \( O(|\gamma|^1) \). Similar arguments apply to the third term in (4.55).

As noted in [3], (4.15) allows us to estimate the remainder on the solutions of the normal form equation for which we have more information. On the other hand, in (4.52) we had to use the fact that the solution of the full system remains of \( O(1) \) in \( H^s \) over the time interval of interest. This control of the solutions of the full system for all \( s > 1/2 \) comes from the local existence theory, and may not be available for longer times. It appears that the possible growth of the norms is one of the main problems in extending the formal theory to higher orders.

The extension of Theorem 4.2 to more general dispersion relations is straightforward, with the index \( s \) determined by the number of derivatives in the low and high frequency regions of the dispersion.

We also expect that there is an alternative way of obtaining Theorem 4.2, by following the first normalization procedure of Section 3 in a sufficiently large Galerkin projection of (4.1). Although we do not obtain an invariant subspace this way, we can use the regularity of the solution to see that the error of the Galerkin approximation and the remainder of the finite dimensional normal form are comparable and small. This strategy is developed in [16] for a general dispersive PDEs coupled to a fast oscillator, where it is shown that under appropriate regularity assumptions on the solution one can decouple the PDE and the oscillator up to an exponentially small error. At the same time we can view the second normalization of Section 3 as a generalization of works on averaging a dispersive system coupled to a fast oscillator, in that we have considered the coupling of a “fast” system of oscillators to a “slow” system (that here is finite dimensional). The result is not decoupling but rather an invariant manifold for the slow system. When the size of the oscillations of the fast system remains small, e.g. due to the regularity of solutions in the case where the fast system describes the high spatial modes of a PDE, the decoupling result, and the existence of an invariant manifold are practically the same, i.e. the slow motions are the same up to a small error.

At present, there is a gap between Theorem 4.2 and the periodic orbits of Proposition 3.2 since we have not estimated any higher Sobolev norms of these solutions for fixed \( L^2 \) norm. It is possible that some of these solutions have oscillations, and that we will need to decrease their amplitude to bring them inside the domain where the normal form equation is valid. We expect that this question can be partially addressed with some numerical work in progress.
Acknowledgements

The author is supported by NSF grant DMS-9810751.

References