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

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.

* *email: panos@colorado.edu*

1. Introduction

The nonlinear Schrödinger equation with dispersion management $u_t = id(t)u_{xx} - 2i\gamma|u|^2u$, where $d(t)$ is a periodic real valued function, and γ 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. [CMK], [YKT]), 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. [SKDBB], [GT]).

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 finding numerically approximate solutions that have variational characterizations. We must however emphasize the possible differences in the dynamics of the periodic and unbounded cases, especially in questions of persistence of approximate solutions and of stability.

We will concentrate on the practically interesting parameter regime where the average δ of the "dispersion management function" $d(t)$, and the nonlinearity parameter γ are comparable and small in absolute value. We also assume that the frequency Ω 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 [GT], [TM], also [AB]). 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. [LYKM], [NDFK]). More recently the existence of periodic solutions for the averaged equation on the line was also proved by variational methods (for $\gamma\delta < 0$, see [ZGJT], [MZJT]), and by a bifurcation argument (for $\gamma\delta \neq 0$, see [Kun1]). The variational characterization of the DM soliton for $\gamma\delta < 0$ also implies nonlinear stability, and there are also error estimates for the averaged equations (see [MZJT]).

In this work we investigate an alternative but related asymptotic theory for the periodic problem. Our first goal is to modify the approach of [GT], [MT] and produce a simpler normal form equation. In particular, to recover the averaging theory of [GT], [TM] 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 Ω). 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 Ω, the quartic normal form equation has a finite dimensional invariant subspace. The dimension of this subspace may be large (of $O(|\delta|^{-\frac{1}{2}})$), 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 [MZJT], [MJG]). 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 $O(|\gamma|^{-1})$. 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. [K]). 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 can not grow too much over the $O(|\gamma|^{-1})$ 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 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 non-autonomous equation

$$u_t = id(t)u_{xx} - 2i\gamma|u|^2u, \tag{2.1}$$

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 γ are real. As remarked in the introduction 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 the signal we send, and is assumed to $2\pi-$periodic. Also, the function $d(t)$ in (2.1) will be $T-$periodic, and we decompose it as

$$d(t) = \delta + \tilde{d}(t), \quad \text{with} \quad \delta = \frac{1}{T}\int_0^T d(s)ds \tag{2.2}$$

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

Remark 2.0.1 The parameters γ, δ, Ω 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 $\frac{|\Omega|}{h}$ of $O(1)$, and (iii) $|\Omega| \sim |\gamma|^{-1}$ with $\frac{|\Omega|}{h} << 1$, where h is the amplitude of $\tilde{d}(t)$. Also of interest is the case where (iv) $h << |\gamma|$.

It is easy to see that equation (2.1) has the structure of a non-autonomous 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 $a_k(t)$, $k \in \mathbf{Z}$ by

$$a_k(t) = u_k(t) e^{i\omega_k \tilde{\Lambda}(t)}, \quad \text{with} \quad \omega_k = k^2, \quad \tilde{\Lambda}(t) = \int_0^t \tilde{d}(s) ds. \tag{2.3}$$

From (2.1), the variables $a(k,t)$ then evolve according to

$$\dot{a}_k = -i\delta\omega_k a_k - 2i\gamma \sum_{k_1,k_2,k_3 \in \mathbf{Z}} a_{k_1} a_{k_2} a^*_{k_3} \delta_{k_1+k_2-k_3-k} e^{-i(\omega_{k_1}+\omega_{k_2}-\omega_{k_3}-\omega_k)\tilde{\Lambda}(t)}, \quad k \in \mathbf{Z}, \tag{2.4}$$

where $\delta_r = 1$ if $r = 0$, and 0 otherwise. The initial condition is $a_k(0) = u_k(0)$, $k \in \mathbf{Z}$. By the definition of $\tilde{\Lambda}$ the right hand side of (2.4) is $T-$periodic. The time dependence is therefore absorbed in the non-linear 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

$$\dot{\phi} = \Omega, \quad \text{with} \quad \phi(0) = \phi_0 = 0. \tag{2.5}$$

We also define the function Λ by $\Lambda(\phi) = \tilde{\Lambda}(t(\phi)) = \tilde{\Lambda}(\frac{\phi}{\Omega})$, and note that Λ is $2\pi-$periodic with zero average. Then, the non-autonomous system (2.4) is equivalent to the autonomous system consisting of (2.4) with $\tilde{\Lambda}(t)$ replaced by $\Lambda(\phi)$, and (2.5). Adding an "action" variable $J \in \mathbf{R}$, we further define the Poisson bracket $[\ ,\ ]$ on pairs of functions F, G of the variables a_k, a^*_k, $k \in \mathbf{Z}$, and ϕ, J by

$$[F,G] = -i \sum_{k \in \mathbf{Z}} \left(\frac{\partial F}{\partial a_k} \frac{\partial G}{\partial a^*_k} - \frac{\partial F}{\partial a^*_k} \frac{\partial G}{\partial a_k} \right) + \frac{\partial F}{\partial J} \frac{\partial G}{\partial \phi} - \frac{\partial F}{\partial \phi} \frac{\partial G}{\partial J}. \tag{2.6}$$

A straightforward calculation then shows that:

Proposition 2.1 The evolution equation for the variables a_k, $k \in \mathbf{Z}$, and ϕ, J above is the Hamiltonian system

$$\dot{a}_k = [a_k, H], \quad k \in \mathbf{Z}, \quad \dot{\phi} = [\phi, H], \quad \dot{J} = [J, H], \tag{2.7}$$

where the Hamiltonian H is

$$H = \delta \sum_{k \in \mathbf{Z}} \omega_k |a_k|^2 - \Omega J + \gamma \sum_{k_1,k_2,k_3,k_4,n \in \mathbf{Z}} e^{in\phi} a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} I(k_1,k_2,k_3,k_4,n), \tag{2.8}$$

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

$$I(k_1,k_2,k_3,k_4,n) = \hat{f}_m(n) \delta_{k_1+k_2-k_3-k_4}, \tag{2.9}$$

$$m = \omega_{k_1} + \omega_{k_2} - \omega_{k_3} - \omega_{k_4}, \quad \hat{f}_m(n) = (2\pi)^{-\frac{1}{2}} \int_0^{2\pi} e^{-im\Lambda(\phi)} e^{-in\phi} d\phi. \tag{2.10}$$

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

Remark 2.1.1 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 $|\gamma| << 1$ will be considered elsewhere. Note that the weakly nonlinear regime poses some interesting problems related to the works of [B], [P], [KP] on weakly perturbed 1-D NLS equations (with Dirichlet boundary conditions). As we see in the next section, the special parameter ranges (iii) and (iv) of Remark 2.0.1 also lead to "near-integrable" systems (see [AB], [YK] respectively for the two limits).

The above setup can be generalized to other dispersion relations ω_k, and to the case where the parameter γ 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 $\hat{f}_m(n)$ in (2.10) can be evaluated in closed form is the piecewise constant $T-$periodic dispersion management function

$$d(t) = \begin{cases} \delta + \tilde{A}, & \text{if } t \in [0, \tau); \\ \delta + \tilde{B}, & \text{if } t \in [\tau, T), \end{cases} \tag{2.11}$$

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

$$d(t) = \delta + \tilde{A} \sin \Omega t \tag{2.12}$$

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 Ω, δ, γ 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 [TM] and [GT] (see also [AB]). 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 [TM] and [GT] 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 ψ_1 such that the canonical transformation obtained by the time-1 map $\Phi^1_{\psi_1}$ of the Hamiltonian flow of ψ_1 simplifies the "perturbation part" $H_2 + H_4$. Specifically, we formally write

$$H \circ \Phi^1_{\psi_1} = \exp Ad_{\psi_1} H = -\Omega J + H_2 + H_4 + [\psi_1, -\Omega J] + Y_1, \tag{3.1}$$

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

$$\gamma I(k_1, k_2, k_3, k_4, n) a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} e^{in\phi} \tag{3.2}$$

in H_4 is eliminated by a monomial

$$i\gamma (n\Omega)^{-1} I(k_1, k_2, k_3, k_4, n) a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} e^{in\phi} \tag{3.3}$$

in ψ_1. Consequently, the resonance condition for the part H_4 is

$$n\Omega = 0, \quad k_1 + k_2 - k_3 - k_4 = 0, \quad \hat{f}_m(n) \neq 0, \quad k_1, \ldots, k_4, n \in \mathbf{Z}. \tag{3.4}$$

with $m = \omega_{k_1} + \omega_{k_2} - \omega_{k_3} - \omega_{k_4}$. We immediately see that the resonant part $\overline{H}_4$ of H_4 is

$$\overline{H}_4 = \gamma \sum_{k_1, k_2, k_3, k_4 \in \mathbf{Z}} a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} \hat{f}_m(0) \delta_{k_1 + k_2 - k_3 - k_4}. \tag{3.5}$$

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

$$\psi_1 = i\gamma \sum_{k_1,k_2,k_3,k_4 \in \mathbf{Z}, n \in \mathbf{Z}^*} a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} \frac{I_1(k_1,k_2,k_3,k_4,n)}{n\Omega}, \tag{3.6}$$

the quartic normal form Hamiltonian is $H_2 + \overline{H}_4$.

Regarding the structure of $H_2 + \overline{H}_4$, let Λ_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 Λ_m can be parameterized explicitly by two integers (we omit this here). The parameterization of Λ_0 becomes especially simple and we can write $\overline{H}_4$ as $\overline{H}_4 = \overline{H}_{4,I} + \overline{H}_{4,NI}$ with

$$H_{4,I} = 2\gamma \hat{f}_0(0) \sum_{k_1,k_2 \in \mathbf{Z}} |a_{k_1}|^2 |a_{k_2}|^2, \tag{3.7}$$

$$H_{4,NI} = \gamma \sum_{m \in \mathbf{Z}^*} \hat{f}_m(0) \sum_{k_1,k_2,k_3,k_4 \in \Lambda_m} a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4}. \tag{3.8}$$

The part $\overline{H}_{4,I}$ is integrable (in the sense of Poincare).

To indicate the structure of the coefficients $\hat{f}_m(0)$, we first consider the piecewise constant dispersion management function of (2.11) with $\tau = \frac{T}{2}$, $\tilde{A} = h$, and $\tilde{B} = -h$. Then,

$$\hat{f}_m(0) = \frac{2i}{\sqrt{2\pi}} \left[e^{-ihm\Omega^{-1}\pi} - 1 \right] \frac{\Omega}{hm}, \quad m \in \mathbf{Z}^*, \tag{3.9}$$

and $\hat{f}_0(0) = \sqrt{2\pi}$. We can see that $N_{4,NI}$ can only vanish for the discrete values of h where $\frac{h}{\Omega} \in \pi \mathbf{Z}$, and also in the limit $h \to +\infty$. In comparison, the coefficients $\hat{f}_m(0)$ for the real analytic dispersion management function of (2.12) with $\tilde{A} = h$ are

$$\hat{f}_m(0) = \sqrt{2\pi} \mathcal{J}_0(|m| h \Omega^{-1}), \quad m \in \mathbf{Z}^* \tag{3.10}$$

with $\mathcal{J}_0$ the Bessel function of order 0. For $h \to +\infty$, the coefficients $\hat{f}_m(0)$, $m \neq 0$, of (3.10) decay as $|hm|^{-\frac{1}{2}}$. For more general integrable dispersion management functions, the definition of the coefficients $\hat{f}_m(0)$ in (2.10) implies that $\hat{f}_0(0) = \sqrt{2\pi}$. 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 non-degenerate critical points we expect a $|hm|^{-\frac{1}{2}}$ decay, while for $\Lambda(\phi)$ Lipschitz but not differentiable we have $|hm|^{-1}$ 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 $\hat{f}_m(0)$ can decay even faster as $|hm| \to \infty$. Thus the non-integrable part of $\overline{H}_4$ decays faster in $|m|$ and h for the more singular dispersion management functions. In the limit $h \to 0$, $H_2 + \overline{H}_4$ 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 [YK], [YKT], [LP]).

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 $|\delta| \sim |\gamma| \sim |\Omega|^{-1}$; 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 $N = [|\delta|^{-\frac{1}{2}}]$ and let

$$H_2 = \tilde{H}_2 + h_2 \tag{3.11}$$

with

$$\tilde{H}_2 = \delta \sum_{|k|>N} \omega_k |a_k|^2, \quad h_2 = \delta \sum_{|k|\leq N} \omega_k |a_k|^2. \tag{3.12}$$

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 non-resonant part of $h_2 + H_4$ by a canonical transformation $\Phi^1_{\chi_1}$ that is the time-1 map of an appropriate function χ_1. We will have

$$H \circ \Phi^1_{\chi_1} = \exp Ad_{\chi_1} H = h_0 + h_2 + H_4 + [\chi_1, h_0] + R_1, \tag{3.13}$$

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

$$\gamma I(k_1, k_2, k_3, k_4, n) a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} e^{in\phi} \tag{3.14}$$

in H_4 is eliminated by a monomial

$$i\gamma[n\Omega - \delta(\tilde{\omega}_{k_1} + \tilde{\omega}_{k_2} - \tilde{\omega}_{k_3} - \tilde{\omega}_{k_4})]^{-1} I(k_1, k_2, k_3, k_4, n) a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} e^{in\phi} \tag{3.15}$$

in χ_1, where

$$\tilde{\omega}_k = \begin{cases} 0, & \text{if } |k| \le N; \\ \omega_k, & \text{if } |k| > N. \end{cases} \tag{3.16}$$

The resonance conditions for the quartic terms are therefore

$$n\Omega - \delta\tilde{m} = 0, \quad k_1 + k_2 - k_3 - k_4 = 0, \quad \hat{f}_m(n) \neq 0, \quad k_1, \ldots, k_4, n \in \mathbf{Z}. \tag{3.17}$$

with $\tilde{m} = \tilde{\omega}_{k_1} + \tilde{\omega}_{k_2} - \tilde{\omega}_{k_3} - \tilde{\omega}_{k_4}$. 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 $\tilde{m}$ in $\mathbf{Z}^4$ subject to $k_1 + k_2 - k_3 - k_4 = 0$, and then examine the level sets of $n\Omega - \delta\tilde{m}$ in $\mathbf{Z}^2$. Instead of this we will consider a partial normalization that eliminates only a subset of the non-resonant 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 $|k| \le N$ can be invariant under the evolution of $h_0 + h_2 + N_4$. To see this we consider "low" and "high" frequency mode index sets $U_L = \{k \in \mathbf{Z} : |k| \le N\}$ and $U_H = \mathbf{Z} \setminus U_L$, and decompose $\mathbf{Z}^4$ into disjoint products of the two sets. We use the notation $U_{LLLH} = U_L \times U_L \times U_L \times U_H$, $U_{LLHL} = U_L \times U_L \times U_H \times U_L$, etc., and subdivide $\mathbf{Z}^4$ into 2^4 disjoint subregions. Accordingly, we also decompose H_4 into 2^4 parts coupling different combinations of quartets of low and high index modes. We now seek to eliminate all quartic non-resonant terms involving only low modes, and quartic non-resonant interactions involving one high frequency mode and three low frequency modes. The corresponding parts of H_4 will be denoted by $H_{4,LLLL}$, and $H_{4,LLLH}$, $H_{4,LLHL}$, $H_{4,LHLL}$, $H_{4,HLLL}$ respectively. To eliminate $H_{4,LLLL}$ we use the function $\chi_{1,LLLL}$ given by the expression of (3.6), with summation over k_1, …, k_4 in U_L. The resonant part of $H_{4,LLLL}$ is

$$N_{4,LLLL} = \gamma \sum_{k_1,k_2,k_3,k_4 \in U_L} a_{k_1} a_{k_2} a^*_{k_3} a^*_{k_4} \hat{f}_m(0) \delta_{k_1+k_2-k_3-k_4}, \tag{3.18}$$

and we denote the non-resonant part of $H_{4,LLLL}$ by $\tilde{H}_{4,LLLL}$ To eliminate $H_{4,LLLH}$, … $H_{4,HLLL}$ we examine the resonance condition (3.17) in U_{LLLH}, …, U_{HLLL} respectively. For instance, in U_{LLLH} (3.17) reduces to

$$n\Omega + \delta k_4^2 = 0, \quad k_1 + k_2 - k_3 - k_4 = 0, \quad k_1, k_2, k_3 \in U_L, \quad k_4 \in U_H, \quad n \in \mathbf{Z}, \tag{3.19}$$

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

$$|\Omega| > 10|\delta|N^2 \tag{3.20}$$

implies $|n\Omega + \delta k_4^2| > 1$, for all $n \in \mathbf{Z}$, i.e. (3.20) implies that the part $H_{4,LLLH}$ can be formally eliminated using an appropriate function $\chi_{1,LLLH}$. The resonance conditions for the other parts of H_4 coupling three low index modes to one high index mode are similar, and condition (3.20) on $|\Omega|$ guarantees the absence of resonances and small divisors. We denote the functions that eliminate $H_{4,LLHL}$, ..., $H_{4,HLLL}$ by $\chi_{1,LLHL}$, ..., $\chi_{1,HLLL}$ respectively. Then letting

$$\chi_1 = \chi_{1,LLLL} + \chi_{1,LLLH} + \chi_{1,LLHL} + \chi_{1,LHLL} + \chi_{1,HLLL} \tag{3.21}$$

(note that χ_1 is real valued) we consider the transformed Hamiltonian

$$H \circ \Phi^1_{\chi_1} = \exp Ad_{\chi_1} H = h_0 + h_2 + N_4 + R_1, \tag{3.22}$$

with

$$N_4 = H_4 - (\tilde{H}_{4,LLLL} + H_{4,LLLH} + H_{4,LLHL} + H_{4,LHLL} + H_{4,HLLL}). \tag{3.23}$$

The remainder R_1 contains higher order terms, of $O(\delta\gamma)$, and $O(\gamma^2)$. (The part $h_0+h_2+N_4$ in (3.22) will be referred to as the normal form part, although it is more accurately a "partial normal form".) Denoting the (complex) span of the modes a_k with $|k| \leq N$ by M_L we have the following observation.

Proposition 3.1 Let $|\Omega| > 10|\delta|N^2$, and define $H \circ \Phi^1_{\chi_1}$ as above. Then the subspace M_L is invariant under the flow of Hamilton's equations for the normal form part Hamiltonian $h_0 + h_2 + N_4$ of (3.22). The restriction of the flow to M_L is the Hamiltonian system corresponding to the Hamiltonian $h_2 + N_{4,LLLL}$.

Proof: It is enough to check that that the equations for $\dot{a}_k$ with $|k| > N$ do not include any monomials of the form $a_{l_1} a_{l_2} a^*_{l_3}$ with l_2, l_2, $l_3 \in U_L$. Such terms can only come from $H_{4,LLLH}$ and $H_{4,LLHL}$, which, however have been eliminated. The equations on M_L are clearly Hamilton's equations for $-\Omega J + h_2 + N_{4,LLLL}$, and since h_2 and $N_{4,LLLL}$ are independent of ϕ, i.e. see (3.5), the term $-\Omega J$ can be omitted form the Hamiltonian on M_L.

■

Note that the quartic normal form parts of $H \circ \Phi^1_{\psi_1}$ and $H \circ \Phi^1_{\chi_1}$, restricted to the subspace M_L, are the same, the difference being that M_L is an invariant subspace for the quartic normal form Hamiltonian system obtained by the second procedure. The dynamics on this subspace can be complicated and interesting. Some special solutions can be obtained by simple arguments.

Proposition 3.2 Let $|\Omega| > 10|\delta|N^2$. Then the Hamiltonian flow of the quartic normal Hamiltonian $h_0 + h_2 + N_4$ of (3.22) has at least $2N+2$ periodic orbits on each level set $\sum_{k\in\mathbf{Z}} |a_k|^2 = C$, $C > 0$.

Proof: We examine Hamilton's equations for $h_0 + h_2 + N_4$ on the invariant subspace M_L. We have the autonomous system

$$\dot{a}_k = -i\frac{\partial}{\partial a^*}(h_2 + N_{4,LLLL}), \quad k \in U_L, \tag{3.24}$$

and we look for solutions of the form $a_k = e^{i\lambda t}A_k$, $k \in U_L$, with λ real. The equation for the $A_k = q_k + ip_k$ has the structure of the constrained variational problem

$$\lambda\frac{\partial \mathcal{I}}{\partial q_k} = \frac{\partial \mathcal{V}}{\partial q_k}, \quad \lambda\frac{\partial \mathcal{I}}{\partial p_k} = \frac{\partial \mathcal{V}}{\partial p_k}, \quad k \in U_L, \tag{3.25}$$

with

$$\mathcal{I} = \sum_{k\in U_L}(q_k^2 + p_k^2), \quad \mathcal{V} = h_2 + N_{4,LLLL}. \tag{3.26}$$

Critical points of $\mathcal{V}$ on $(2N+1)$-spheres of radius $C > 0$ in $\mathbf{R}^{2N+2}$ thus yield solutions of (3.25), and the statement follows from the fact that a smooth function on the $n-$sphere has at least $n+1$ critical points.

■

Remark 3.2.1 In the argument above we can also look for solutions of (3.24) that have the more general form $a_k = e^{ig_k t}A_k$, $k \in U_L$, with g_k satisfying $g_{k_1} + g_{k_2} - g_{k_3} - g_{k_4} = 0$ for all k_1, k_2, k_3, k_4 in U_L. The equations for the A_k will then have the form of (3.25) with $\mathcal{I} = \sum_{k\in U_L} g_k|A_k|^2$. For instance, for $g_k = N+1-k$ the level sets $\mathcal{I} = C$, $C > 0$ are "asymmetric" ellipsoids and we similarly have at least $2N+2$ periodic orbits. Some other choices of g_k, for instance $g_k = k$, do not lead to compact level sets $\mathcal{I} = C$, and the existence of solutions of (3.25) is not guaranteed.

Some of the periodic orbits in Proposition 3.2 can be thought of as approximations of periodic orbits for the quartic normal form obtained by the first version of the normal form argument. Of special interest is the DM soliton solutions discussed by many authors. Proofs of existence of these solutions concern the analogue of the quartic normal form system with Hamiltonian $H_2 + \overline{H}_4$ in (3.5) for the line (i.e. with summation over the k_i replaced by integration). In the case $\delta\gamma < 0$ the DM soliton is characterized as a minimum of $H_2 + \overline{H}_4$ over functions with fixed L_2 norm (see [TM]). An existence proof based on the variational characterization of the DM soliton is in [ZGJG]. The functional is minimized in the Sobolev space $H^1(\mathbf{R}, \mathbf{C})$ and the DM soliton solution decays at infinity. The analogous solution here is the minimum of $\mathcal{V}$ over the spheres $\mathcal{I} = C$ in (3.26). The sign of $\delta\gamma$ is irrelevant in Proposition 5.2, and we thus obtain solutions that may be also related to the ones shown in [Kun1] for $\delta\gamma \neq 0$ by a bifurcation argument. Note that the assumption $\delta\gamma < 0$ is crucial in the variational argument of [ZGJT]. (We have also recently seen a variational argument of [Kun2] for $\delta = 0$, obtained from $\delta\gamma \to 0^-$, with γ fixed.) Note that the second normal form construction we presented is inapplicable for $\delta = 0$, while the first averaging procedure is still meaningful.

Remark 3.2.2 The comparisons with a above works are not strictly appropriate since we are here considering a different problem, and we do not have an analogue of the second normalization procedure for the NLS+DM system on the line. The proofs on the line also involve additional technical questions.

Conservation of the quartic normal form Hamiltonian $h_2 + N_{4,LLLL}$ and the L^2 norm on the invariant subspace M_L implies that the minima and maxima of $\mathcal{V}$ in Proposition 3.2 are nonlinearly stable on the invariant subspace M_L. 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 [TM] implies nonlinear stability (see [ZGJG], [MZJG]). Note however that in (3.11), (3.12), we can split H_2 into $\tilde{H}_2 + h_2$ differently, by making N larger. Assuming $|\Omega| > 10|\delta|N^2$, 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 ω_k (a cubic dispersion is considered in [MZJG], [MGJ]). For instance, let N_1, $M_1 > 0$ with δM_1 of $O(1)$, and assume that $\omega_k : \mathbf{R}^+ \to \mathbf{R}$ is strictly increasing for $k > N_1$, diverges as $k \to +\infty$, and is bounded by M_1 for $k \leq N_1$. Also extend ω_k to $\mathbf{R}$ 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 $a_{k_1} \ldots a_{k_n} a^*_{k_{n+1}} \ldots a^*_{k_{2n}}$ with $k_1 + \ldots + k_n - k_{n+1} - \ldots - k_{2n} = 0$. A resonant monomial with indices $k_1, \ldots, k_{2n-1} \in U_L$, $k_{2n} \in U_H$ must satisfy $\delta\Omega - \delta k_{2n}^2 = 0$, and by $|k_{2n}| \leq (2n-1)N$, we can avoid resonances and small divisors by requiring that $|\Omega| > \delta(2n-1)^2 N^2 + 1$. 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 $\frac{\Omega}{\delta}$ 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

$$\dot{a} = La + F(a,t), \quad a(0) = a_0 \tag{4.1}$$

with $L = i\delta\partial_{xx}$, and $F(a,t)$ the nonlinearity. We also write Hamilton's equation for the quartic normal form Hamiltonian $h_0 + h_2 + N_4$ in (3.22) as

$$\dot{b} = Lb + G(b,t), \quad b(0) = b_0. \tag{4.2}$$

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

$$||u||_s^2 = \sum_{k \in \mathbf{Z}} (1 + |k|^2)^s |u_k|^2, \tag{4.3}$$

with u_k the Fourier coefficients of u. First, we have the following basic local existence theorem.

Proposition 4.1 Let $s > \frac{1}{2}$, $\beta > 0$, with $\beta \sim O(1)$, and assume that $d(t)$ is locally absolutely integrable. Consider the initial value problems of (4.1), (4.2) with initial conditions satisfying $||a_0||_s \sim ||b_0||_s \sim O(1)$. Then for $|\gamma|$ sufficiently small there exists a positive constant $C = C(||a_0||_s, ||b_0||_s, \beta) \sim O(1)$ and a time $t_1 \geq C|\gamma|^{-1}$ for which (4.1), (4.2) have unique solutions $a(t)$, $b(t) \in R_s(t_1, y_0, \beta)$, where $R_s(t_1, y_0, \beta) = \{y(t) \in C^0([0, t_1], H^s) : ||y(t) - y_0||_s \leq \beta\}$.

Notation: A quantity Q will be of $O(1)$ if $|\gamma| << |Q| << |\gamma|^{-1}$. Recall that we are interested in $|\gamma| << 1$.

Thus, assuming initial conditions of $O(1)$, solutions exist for a "long" time of $O(|\gamma|^{-1})$, and their size remains of $O(1)$ during that interval. Proposition 4.1 follows from a standard fixed point argument. The operator relating the Fourier coefficients $u_k(t)$ and $a_k(t)$ in (2.4) is an isometry in H^s, and the maps $F(u, t)$ are Lipschitz in u, uniformly in t. In particular, we have

$$||F(u, t) - F(v, t)||_s \leq L_F(||u||_s, ||v||_s)||u - v||_s, \quad s > \frac{1}{2}, \tag{4.4}$$

with $L_F(||u||_s, ||v||_s) = |\gamma| C_s^2 (||u||_s^2 + ||u||_s ||v||_s + ||v||_s^2)$, and C_s a constant satisfying $||uv||_s \leq C_s ||u||_s ||v||_s$ for $s > \frac{1}{2}$. The Lipschitz constant L_F is precisely the one for the map $u \mapsto i\gamma |u|^2 u$ 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 $\Phi_{\chi_1}^1$ are obtained readily following the arguments of Lemma 4.3 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, β. We now state our estimate for the distance between the solutions of the full system (4.1) and and the quartic normal form equations (4.2).

Theorem 4.2 Let $s \geq 1$, and let $0 < |\delta| \leq c_\delta |\gamma|$ for some $c_\delta \sim O(1)$. Also assume that $|\Omega| > 10|\delta| N^2$ 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 $||b_0||_{s'} \leq \rho_0$, with $s' = s + 2$, $\rho_0 \sim O(1)$, 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 γ_0 for which $|\gamma| \leq \gamma_0$ implies

$$||a(t) - b(t)||_s < C_0|\gamma|, \quad \forall t \in [0, \tilde{C}_0|\gamma|^{-1}]. \tag{4.5}$$

The constants C_0, C_1, and γ_0 depend on s, $|\Omega|$, c_δ, and ρ_0.

Remark 4.2.1 The constants C_0, $\tilde{C}_0$, γ_0 do not change significantly as $|\Omega|$ diverges. This is seen in Lemma 4.3 below. Also, the constants are independent of the amplitude h of the dispersion management function $d(t)$. Thus the estimate applies to all the physically interesting parameter ranges of $|\Omega|$ and h discussed in Remark 2.0.1, 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 $\Phi^1_{\chi_1}$ 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

$$\langle u, v \rangle_s = \mathrm{Re} \sum_{k \in \mathbf{Z}} (1 + |k|^2)^s u_k v_k^*, \tag{4.6}$$

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 ρ around the origin is denoted by $B^s(\rho)$. We also let $\langle u, v \rangle = \langle u, v \rangle_0$. Functions on H^s can be extended to the complexification H^s_c of H^s by letting the real and imaginary parts of the Fourier coefficients u_k of $u \in H_s$ become complex.

For $f : B^s(\rho) \to \mathbf{R}$ Fréchet C^1 smooth in $B^s(\rho)$, we define the gradient map $\nabla f : B^s(\rho) \to H^{-s}$ by $\langle \nabla f(u), v \rangle = Df(u)v$, with D the Fréchet derivative, and $v \in H^s$. Also, $\mathcal{J} = -i$ defines a symplectic structure in H^s, and we denote the Hamiltonian vector field $\mathcal{J}\nabla f$ of a C^1 function f by V_f. The time-dependent Hamiltonian for the system of (2.4) is $h(a, t) = H(a, \Omega t, J) + \Omega J$, with H as in (2.8). If $d(t)$ is integrable, $h : H^s \times \mathbf{R} \to \mathbf{R}$, $s \geq 1$, is C^1 in the variable a, for all $t \in \mathbf{R}$, and (4.1) can be written as the non-autonomous Hamiltonian system

$$\dot{a} = \mathcal{J}\nabla h(a, t) = V_h(a, t). \tag{4.7}$$

The extended phase space will be the set of $x = [x_1, x_2, x_3] \in H^s \times \mathbf{R} \times \mathbf{R}$, where the second and third components correspond to the angle ϕ and action J respectively (the angle is defined on the covering space of the circle). The spaces $H^s \times \mathbf{R}^2$ can be considered as real Hilbert spaces with the inner product

$$\langle x, y \rangle_{(s,2)} = \langle x_1, y_1 \rangle_s + x_2 y_2 + x_3 y_3. \tag{4.8}$$

The norm in $H^s \times \mathbf{R}^2$ is denoted by $|| \, ||_{(s,2)}$, and we let $K^s(\rho) = B^s(\rho) \times \mathbf{R}^2$. As above, we use the inner product in $H^0 \times \mathbf{R}^2$ to define the gradient $\tilde{\nabla}$ for smooth functions on $H^s \times \mathbf{R}^2$. The symplectic structure on $H^s \times \mathbf{R}^2$, denoted by $\tilde{\mathcal{J}}$, will be the tensor product of $\mathcal{J}$ with the standard symplectic structure in $\mathbf{R}^2$, i.e. compare with the Poisson bracket in (2.6). Also, we let $\tilde{V}_g = \tilde{\mathcal{J}} \tilde{\nabla} g$. 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 $H^s \times \mathbf{R}^2$, $s \geq 1$, and Hamilton's equations (2.7) can be written as

$$\dot{x} = \tilde{\mathcal{J}} \tilde{\nabla} H(x) = \tilde{V}_H(x). \tag{4.9}$$

Following the previous section, we want to construct a one-parameter family of canonical transformations $\Phi^\epsilon_{\chi_1} = \mathcal{T}^\epsilon$ by integrating

$$\frac{d}{d\epsilon} \mathcal{T}^\epsilon(y) = \tilde{V}_{-\chi_1}(\mathcal{T}^\epsilon(y)), \quad \mathcal{T}^0(y) = y, \tag{4.10}$$

with χ_1 as in (3.25). The components of $\mathcal{T}^\epsilon$ in $H^s \times \mathbf{R} \times \mathbf{R}$ will be denoted by $\mathcal{T}^\epsilon_i$, $i = 1$, 2, 3. Since χ_1 does not depend on the action y_3, the second component of the equation is integrated trivially yielding $\mathcal{T}^\epsilon_2(y) = \mathcal{T}^\epsilon_2(y_2) = y_2$, for all ϵ, i.e. the angle variable does not change. The existence of solutions to (4.10) for $|\gamma|$ sufficiently small and appropriate initial conditions is shown in Lemma 4.3 below, where we also show that $\mathcal{T}^1_1(y)$ and y_1 are $O(|\gamma|)$ close in H^s. To exhibit the normal form equation and the remainder we write the Hamiltonian system of (4.9) as

$$\dot{x} = \tilde{L} x + \tilde{V}_{H_4}(x), \tag{4.11}$$

with $\tilde{L} x = \tilde{V}_{h_0 + h_2}(x)$. Equation (4.11) is then written in the new variable y, $x = \mathcal{T}^1(y)$, as

$$\dot{y} = \tilde{L} y + \tilde{V}_{N_4}(y) + \tilde{V}_{R_1}(y), \tag{4.12}$$

where we are using the notation of (3.26) for the Hamiltonian $H \circ \mathcal{T}^1 = H \circ \Phi^1_{\chi_1}$, and the fact that $\mathcal{T}^1$ is symplectic. The remainder R_1 of (3.26) is

$$R_1 = (h_2 \circ \mathcal{T}^1 - h_2) + (H_4 \circ \mathcal{T}^1 - H_4) + (h_0 \circ \mathcal{T}^1 - h_0 - [\chi_1, h_0]), \tag{4.13}$$

and the quartic normal form system is

$$\dot{z} = \tilde{L}z + \tilde{V}_{N_4}(z). \tag{4.14}$$

An elegant way of comparing solutions of the normal form and the full system is to consider the evolution of the "error" $r(t) = x(t) - \mathcal{T}^1(z(t))$ (see [BCP]). In particular, combining (4.11), (4.12), and (4.14) we obtain

$$\dot{r} = \tilde{L}r + (\tilde{V}_{H_4}(x) - \tilde{V}_{H_4}(\mathcal{T}^1(z))) + [D\mathcal{T}^1(z)]\tilde{V}_{R_1}(z). \tag{4.15}$$

The existence of $D\mathcal{T}^1$ is shown in Lemma 6.4. We will consider (4.15) with the initial condition $r(0) = z(0) - \mathcal{T}^1(z(0))$. Note that we are interested in the size of the first component $r_1(t)$ of $r(t)$. We will denote the three components of $\tilde{V}_f$ in $H^s \times \mathbf{R} \times \mathbf{R}$ by $\tilde{V}_f^i$, $i = 1,\ 2,\ 3$. Also, we let D_1, D_2, D_3 denote the Fréchet derivatives along each the three components in $H^s \times \mathbf{R} \times \mathbf{R}$. 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_4 depend on the first two components of $H^s \times \mathbf{R} \times \mathbf{R}$ only, and $T_1^1(y) = T_1^1(y_1, y_2)$, $T_2^1(y) = y_2$. For the third term in (4.13), the observation follows from the cohomology equation $[\chi_1, h_0] = G_4$, with $G_4 = N_4 - H_4$, and

$$(h_0 \circ \mathcal{T}^1 - h_0)(y) = \int_0^1 \frac{d}{d\epsilon} h_0(\mathcal{T}^\epsilon(y))d\epsilon = \int_0^1 [\chi_1, h_0](\mathcal{T}^\epsilon(y))d\epsilon = \int_0^1 G_4(\mathcal{T}^\epsilon(y))d\epsilon. \tag{4.16}$$

The second component of $\tilde{V}_{R_1}$, and $D_3\mathcal{T}_1^1(z)$ therefore vanish, and the first component of (4.15) becomes

$$\dot{r}_1 = Lr_1 + (V_{H_4}(x) - V_{H_4}(\mathcal{T}^1(z))) + [D_1\mathcal{T}_1^1(z)]V_{R_1}(z), \tag{4.17}$$

with $Lr_1 = V_{h_0+h_2}(r_1)$. The vector field $[D_1\mathcal{T}_1^1(z)]V_{R_1}(z)$ is estimated using Lemmas 4.4 and 4.5 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 $\mathcal{T}^1$ is near-identity.

Lemma 4.3 Let $\rho > 0$, $\alpha_1 \in (0,1)$, and set $\rho_1 = \alpha_1\rho$. Assume that $s > \frac{1}{2}$, $|\Omega| > 10|\delta|N^2$, and that $d(t\Omega^{-1}) \in H^0$. Then for $|\gamma| \leq \gamma_1(s, |\Omega|, \rho, \alpha_0,)$, the initial value problem (4.10) with $y \in K^s(\rho_1)$ has a unique solution $\mathcal{T}^\epsilon(y)$, $\epsilon \in [0,1]$, and its flow defines a one-parameter family of canonical transformations $\mathcal{T}^\epsilon : K^s(\rho_1) \to K^s(\rho)$. Moreover, we have

$$\sup_{y \in K^s(\rho_1)} ||\mathcal{T}_1^\epsilon(y) - y_1||_s \leq \epsilon C_1|\gamma|, \quad \epsilon \in [0,1], \tag{4.18}$$

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

Proof: We first consider the equation for the rate of change of $\mathcal{T}_1^\epsilon$. The right hand side is $V_{-\chi_1} = -\mathcal{J}\nabla\chi_1$. Since χ_1 is independent of the action y_3 and the angle y_2 does not change, we can consider $\phi = y_2$ as a parameter. We will show that V_{χ_1} is real analytic in in y_1, uniformly in y_2. Denote the Fourier coefficients of y_1 by a_k, and let $(V_f)_k$ be the $k-$th Fourier component of the vector field V_f. The definition of χ_1 in the previous section leads readily to explicit expressions, and we have

$$|(V_{-\chi_{1,LLLL}})_k| \leq 2|\gamma| \sum_{k_1,k_2,k_3 \in U_L} |a_{k_1} a_{k_2} a^*_{k_3}| \delta_{k_1+k_2-k_3-k} \left| \frac{F_m(\phi)}{\Omega} \right|, \quad k \in U_L, \tag{4.19}$$

with

$$F_m(\phi) = \sum_{n \in \mathbf{Z}^*} \frac{\hat{f}_m(n)}{in} e^{in\phi}. \tag{4.20}$$

Note that $(V_{-\chi_{1,LLLL}})_k = 0$ for $k \notin U_L$. Also,
(4.21)

$$|(V_{-\chi_{1,LLLH}})_k| \leq |\gamma| \sum_{k_1,k_2 \in U_L, |k_4| \in [N,3N]} |a_{k_1} a_{k_2} a^*_{k_3}| \delta_{k_1+k_2-k-k_4} \left| \frac{\hat{f}_m(0)}{i\delta k_4^2} + \frac{G_m(\phi, k_4)}{\Omega} \right|,$$

if $k \in U_L$, and

$$|(V_{-\chi_{1,LLLH}})_k| \leq |\gamma| \sum_{k_1,k_2,k_3 \in U_L} |a_{k_1} a_{k_2} a^*_{k_3}| \delta_{k_1+k_2-k_3-k} \left| \frac{\hat{f}_m(0)}{i\delta k^2} + \frac{G_m(\phi, k)}{\Omega} \right|, \tag{4.22}$$

if $|k| \in [N, 3N]$, where

$$G_m(\phi, k) = \Omega \sum_{n \in \mathbf{Z}^*} \frac{\hat{f}_m(n)}{i(n\Omega + \delta k^2)} e^{in\phi}. \tag{4.23}$$

All other Fourier components of $V_{-\chi_{1,LLLH}}$ vanish. We also have similar estimates for the symplectic gradients of $\chi_{1,LLHL}$, $\chi_{1,LHLL}$, and $\chi_{1,HLLL}$. To bound $F_m(\phi)$ we note that

$$|F_m(\phi_1) - F_m(\phi_2)| \leq \sqrt{2\pi} ||F'_m||_0 \leq \sqrt{2\pi} ||e^{-im\Lambda}||_0 = 2\pi, \tag{4.24}$$

ϕ_1, $\phi_2 \in [0, 2\pi)$ (with prime the derivative with respect to ϕ). By $|\Omega| \geq 9\delta N^2$ and $|k| \in [N, 3N]$ we see that $|1 + \delta k^2/(n\Omega)|^{-1} \geq 2$, and we similarly obtain $|G_m(\phi_1, k) -$

$G_m(\phi_2, k)| \leq 4\pi$, for all ϕ_1, $\phi_2 \in [0, 2\pi)$ and all admissible k, m. Since the functions F_m and G_m are continuous in ϕ 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 $|\delta| \leq \frac{1}{2}$ implies $|\hat{f}_m(0)||\delta k^2| \leq 2$, uniformly in $m \in \mathbf{Z}$ and $k \in [N, 3N]$.

Estimating the discrete convolutions in (4.19), (4.21), (4.22), and in similar expressions for the symplectic gradients of $\chi_{1,LLHL}$, $\chi_{1,LHLL}$, and $\chi_{1,HLLL}$, and using similar bounds on the coefficients F_m, G_m above, we therefore have

$$||V_{-\chi_1}||_s \leq C|\gamma|||y_1||_s^3, \quad s > \frac{1}{2}, \tag{4.26}$$

with a constant C of $O(1)$ that depends on s and $|\Omega|$ and is decreasing in $|\Omega|$. 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_c^s, $s > \frac{1}{2}$, with a slightly larger constant C of $O(1)$ that depends on s and $|\Omega|$. It is also immediate from the convolution estimates that the complexification of $V_{-\chi_1}$ is bounded and weakly analytic, and therefore analytic (see [PT], Appendix A). The estimates are also uniform in the parameter y_2. Thus, given ρ, α_1 as in the statement we can choose $|\gamma|$ sufficiently small so that the integral curves $\mathcal{T}_1^\epsilon(y_1, y_2)$ exist for all $y_1 \in B^s(\rho_1)$, $\epsilon \in [0, 1]$, and $y_2 \in \mathbf{R}$. Note that $\mathcal{T}_1^\epsilon(y_1, y_2)$ is also real analytic in y_1 and ϵ, for all $y_2 \in \mathbf{R}$ (see e.g. [K], ch.1).

The distance between $\mathcal{T}_1(y_1, y_2)$ and y_1 can be estimated using

$$\mathcal{T}_1^\epsilon(y_1, y_2) = y_1 + \int_0^\epsilon V_{\chi_1}(\mathcal{T}_1^\sigma(y_1, y_2), y_2) d\sigma, \quad y_2 \in \mathbf{R}. \tag{4.27}$$

We easily see from (4.27) and (4.26) that $|\gamma| \leq \frac{(1-\alpha_1)}{C\rho^2}$, with C as in (4.26), implies that if $y_1 \in B^s(\rho_1)$ then $\mathcal{T}_1^\epsilon(y_1, y_2)$ stays in $B^s(\rho)$ for all $\epsilon \in [0, 1]$ and y_2. Using (4.27) again we obtain (4.18).

The third component of (4.10) is

$$\frac{d}{d\epsilon}\mathcal{T}_3^\epsilon(y) = -(D_2\chi_1)(\mathcal{T}^\epsilon(y_1, y_2), y_2), \quad \mathcal{T}_3^0(y) = y_3, \tag{4.28}$$

and it suffices to integrate the right hand side with respect to ϵ. We have

$$|D_2\chi_{1,LLLL}| \leq |\gamma| \sum_{k_1,k_2,k_3,k_4 \in U_L} |a_{k_1} a_{k_2} a_{k_3}^* a_{k_4}^* |\delta_{k_1+k_2-k_3-k_4}|F_m'(\phi)|, \tag{4.29}$$

$$(4.30) \qquad |D_2\chi_{1,LLLH}| \leq |\gamma| \sum_{k_1,k_2,k_3\in U_L,k_4\in U_H} |a_{k_1}a_{k_2}a^*_{k_3}a^*_{k_4}|\delta_{k_1+k_2-k_3-k_4}|G'_m(\phi,k_4)|,$$

with prime the derivative with respect to ϕ. For $\Lambda(\phi) \in H^s$ with $s > \frac{1}{2}$, 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, $|\Omega| > 10|\delta|N^2$ and $|k_4| \in [N, 3N]$ imply that $G'_m(\phi, k_4)$ is bounded uniformly in m, k_4. Analogous statements hold for the other terms of $D_2\chi_1$. By the real analyticity of $\mathcal{T}^\epsilon_1(y_1, y_2)$ in ϵ, the right hand side of (4.28) is therefore real analytic in ϵ, for all $y_1 \in B^s(\rho_1)$, and $y_2 \in \mathbf{R}$.

■

We now consider the derivative of the canonical transformations $\mathcal{T}^1$. We are especially interested in the derivative along the H^s direction. In what follows $||A||_{s',s}$ denotes the operator norm of a linear map $A : H^{s'} \to H^s$, and I_1 is the identity in H^s.

Lemma 4.4 Let ρ, ρ_1, and s as in the Lemma 4.3 and assume that $\Lambda(\phi)$ is in H^q, $q > \frac{5}{2}$, and that $|\gamma| \leq \gamma_2(s, |\Omega|, \rho)$ Then, the canonical transformations $T^\epsilon : K^s(\rho_1) \to H^s$ defined above are Fréchet differentiable and we have

$$(4.31) \qquad \sup_{y\in K^s(\rho_1)} ||(D_1\mathcal{T}^\epsilon_1)(y)||_{s,s} \leq 1 + \epsilon C_2|\gamma|.$$

Also, the maps $D_1\mathcal{T}^\epsilon_1(y)$, $y \in K^s(\rho_1)$ in (4.31) are invertible and satisfy

$$(4.32) \qquad \sup_{y\in K^s(\rho_1)} ||(D_1\mathcal{T}^\epsilon_1)^{-1}(y)||_{s,s} \leq 1 + \epsilon C_2|\gamma|.$$

$$(4.33) \qquad \sup_{y\in K^s(\rho_1)} ||(D_1\mathcal{T}^\epsilon_1)^{-1}(y) - I_1||_{s,s} \leq \epsilon C_2|\gamma|.$$

The constant $C_2 = C_2(s, |\Omega|, \rho)$ is of $O(1)$ for ρ of $O(1)$.

Proof: By

$$(4.34) \qquad \mathcal{T}^\epsilon(y) = y + \int_0^\epsilon \tilde{V}_{-\chi_1}(\mathcal{T}^\sigma(y))d\sigma$$

and Lemma 4.3 differentiability of the maps $\mathcal{T}^\epsilon(y)$ with $y \in K^s(\rho_1)$ will follow from the differentiability of the vector field $\tilde{V}_{\chi_1}(y)$ for $y \in K^s(\rho)$. Note that $D_2\tilde{V}^2_{-\chi_1} = 1$, and that the only other nonvanishing partial derivatives, namely $D_1\tilde{V}^1_{\chi_1}$, $D_2\tilde{V}^1_{\chi_1}$, $D_1\tilde{V}^3_{\chi_1}$, and

$D_3\tilde{V}^3_{\chi_1}$ of $\tilde{V}_{\chi_1}$ involve up to two derivatives (in ϕ) of the functions $F_m(\phi)$, $G_m(\phi,k)$ in Lemma 4.3. The condition on $\Lambda(\phi)$ on the other hand implies that the series for the third derivatives of $F_m(\phi)$, $G_m(\phi,k)$ are absolutely convergent for all admissible m, k. Thus the above four partial derivatives of $\tilde{V}_{\chi_1}$ are well-defined and continuous in y_2. Continuity of the four partial derivatives in y_1 comes from the real analyticity of χ_1, and the existence of $D\tilde{V}_{-\chi_1}(y)$ follows immediately. Letting $|\gamma|$ sufficiently small and integrating $\tilde{V}_{\chi_1}$ as in Lemma 4.3 we also obtain differentiable transformations $\mathcal{T}^{-\epsilon}$, $\epsilon \in [0,1]$. Moreover, we have $\mathcal{T}^{-\epsilon}(\mathcal{T}^{\epsilon}(y)) = \mathcal{T}^{\epsilon}(\mathcal{T}^{-\epsilon}(y)) = y$, therefore $(D\mathcal{T}^{\epsilon})^{-1}(y) = D\mathcal{T}^{-\epsilon}(\mathcal{T}^{\epsilon}(y))$. Viewing $(D\mathcal{T}^{\epsilon})(y)$ and its inverse as 3×3 block matrices applied to vectors in $H^s \times \mathbf{R} \times \mathbf{R}$, we see that the 1, 1 entry $[(D\mathcal{T}^{\epsilon})^{-1}(y)]_{1,1}$ must equal $(D_1\mathcal{T}_1^{\epsilon})^{-1}(y) = D_1\mathcal{T}_1^{-\epsilon}(\mathcal{T}^{\epsilon}(y))$. Also note that

$$\mathcal{T}_1^{-\epsilon}(y) = y + \int_0^{\epsilon} V_{\chi_1}(\mathcal{T}_1^{\sigma}(y_1,y_2),y_2)d\sigma, \tag{4.35}$$

and

$$||D_1V_{\chi_1}(y)g||_s \le C|\gamma|||y_1||_s^2||g||_s, \quad s > \frac{1}{2}, \tag{4.36}$$

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

$$\sup_{y\in K^s(\rho_1)} ||(D_1\mathcal{T}_1^{\epsilon})^{-1}(y)||_{s,s} \le (1+\epsilon G e^{\epsilon G}), \tag{4.37}$$

$$\sup_{y\in K^s(\rho_1)} ||(D_1\mathcal{T}_1^{\epsilon})^{-1}(y) - I_1||_{s,s} \le \epsilon G(1+\epsilon G e^{\epsilon G}), \tag{4.38}$$

with

$$G \le \sup_{y\in K^s(\rho)} ||V_{\chi_1}(y)||_s \le C(s,|\Omega|)\rho^3|\gamma|, \tag{4.39}$$

i.e. (4.32) and (4.33). Estimate (4.31) follows from similar arguments.

■

Remark 4.4.1 Note that the size of the partial derivatives $D_2\tilde{V}^1_{\chi_1}$, $D_1\tilde{V}^3_{\chi_1}$, $D_3\tilde{V}^3_{\chi_1}$ may be large.

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

Lemma 4.5 Let s, ρ, ρ_1, γ, and $\Lambda(\phi)$ as in Lemma 4.4. Consider the transformations $\mathcal{T}^\epsilon$ above, and the remainder R_1 of (4.13). Also, let $s' = s + 2$. Then,

$$\sup_{y \in K^{s'}(\rho_1)} ||V_{R_1}(y)||_s \leq C_3 |\gamma|^2, \tag{4.40}$$

with $C_3 = C_3(s, |\Omega|, \rho, \alpha_1, c_\delta)$ a constant of $O(1)$.

Proof: The vector field $\tilde{V}_{R_1}$ 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 $K^s(\rho)$, and consider $y \in K^s(\rho_1)$. By Lemma 4.3 and the fact that the transformations $\mathcal{T}^\epsilon$ are symplectic we have

$$\tilde{V}_{f \circ \mathcal{T}^\epsilon}(y) = [(D\mathcal{T}^\epsilon)^{-1}(y)] \tilde{V}_f(\mathcal{T}^\epsilon(y)) \quad \epsilon \in [0, 1]. \tag{4.41}$$

Observe that $D_3 \mathcal{T}_1^{-\epsilon}(y)$ vanishes since $\chi_1(y) = \chi_1(y_1, y_2)$. Then, if $\tilde{V}_f^2$ vanishes, (4.41) becomes

$$\tilde{V}^1_{f \circ \mathcal{T}^\epsilon}(y) = [(D_1 \mathcal{T}_1^\epsilon)^{-1}(y)] \tilde{V}^1_f(\mathcal{T}^\epsilon(y)), \tag{4.42}$$

using $[(D\mathcal{T}^\epsilon)^{-1}(y)]_{1,1} = (D_1 \mathcal{T}_1^\epsilon)^{-1}(y)$ (see the proof of Lemma 4.4). We write (4.42) as

$$(V_{f \circ \mathcal{T}^\epsilon} - V_f)(y) = [(D_1 \mathcal{T}_1^\epsilon)^{-1}(y)](V_f(\mathcal{T}^\epsilon(y)) - V_f(y)) + [(D_1 \mathcal{T}_1^\epsilon)^{-1}(y) - I] V_f(y). \tag{4.43}$$

Restricting y to $K^{s'}(\rho_1)$ with $s' \geq s$, taking H^s norms, and using (4.32) and (4.33) in Lemma 4.4, (4.43) yields

$$||(V_{f \circ \mathcal{T}^1} - V_f)(y)||_s \leq (1 + C_2 |\gamma|) \sup_{y \in K^{s'}(\rho_1)} ||V_f(\mathcal{T}^1(y)) - V_f(y)||_s + C_2 |\gamma| ||V_f(y)||_s, \tag{4.44}$$

for all $y \in K^{s'}(\rho_1)$, $s' \geq s$.

We now consider the cases where f is h_2 and H_4 respectively. Note that $\tilde{V}^2_{h_2}$ and $\tilde{V}^2_{H_4}$ vanish. The vector field V_{h_2}, viewed as a function from $H^{s'}$ to H^s, $s' = s + 2$, is uniformly Lipschitz, with Lipschitz constant $|\delta|$. The estimate for the distance between $\mathcal{T}_1^1(y)$ and y_1 in Lemma 4.3, (4.44) with $f = h_2$ yield

$$\sup_{y \in K^{s'}(\rho_1)} ||(V_{h_2 \circ \mathcal{T}^1} - V_{h_2})(y)||_s \leq (1 + C_2 |\gamma|) |\delta| C_1 |\gamma| + C_2 |\gamma| |\delta| \rho_1. \tag{4.45}$$

Also, $V_{H_4}(y_1, y_2) : B^s(\rho_1) \times \mathbf{R} \to H^s$, $s > \frac{1}{2}$, is Lipschitz in y_1, uniformly in y_2. Moreover, if ρ_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

$$\sup_{y \in K^{s'}(\rho_1)} ||(V_{H_4} \circ \mathcal{T}^1 - V_{H_4})(y)||_s \leq (1 + C_2 |\gamma|) C |\gamma|^2 + C |\gamma|^2, \tag{4.46}$$

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

To estimate the third term in $\tilde{V}_{R_1}$, let $G_4 = N_4 - H_4$ and recall that $[\chi_1, h_0] = \tilde{G}_4$. Using (4.16), and (4.41) with $f = G_4$, we have

$$\tilde{V}_{h_0\circ\mathcal{T}^1 - h_0 - [\chi_1,h_0]}(y) = \int_0^1 \Big([(D\mathcal{T}^\epsilon)^{-1}(y)]\tilde{V}_{G_4}(\mathcal{T}^\epsilon(y)) - \tilde{V}_{G_4}(y)\Big)\, d\epsilon. \tag{4.47}$$

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

$$V_{h_0\circ\mathcal{T}^1 - h_0 - [\chi_1,h_0]}(y) = \int_0^1 \big([(D_1\mathcal{T}_1^\epsilon)^{-1}(y)]V_{G_4}(\mathcal{T}^\epsilon(y)) - V_{G_4}(y)\big)\, d\epsilon. \tag{4.48}$$

Splitting the integrand as in (4.43), and using the derivative estimates from Lemma 4.4, (4.48) leads to

(4.49)
$$||V_{h_0\circ\mathcal{T}^1 - h_0 - [\chi_1,h_0]}(y)||_s \le \int_0^1 \epsilon C_2||(V_{G_4}(\mathcal{T}^\epsilon(y)) - V_{G_4}(y)||_s d\epsilon + \int_0^1 \epsilon C_2|\gamma|||V_{\tilde{G}_4}(y)||_s d\epsilon,$$

for all $y \in K^s(\rho_1)$. Arguing as for V_{H_4} above, $V_{G_4}(y_1,y_2) : B^s(\rho_1)\times\mathbf{R} \to H^s$ is Lipschitz in y_1, uniformly in y_2, with Lipschitz constants bounded by a quantity of $O(|\gamma|)$. Combining this with Lemma 4.3, (4.49) implies

$$\sup_{y\in K^{s'}(\rho_1)} ||V_{h_0\circ\mathcal{T}^1 - h_0 - [\chi_1,h_0]}(y)||_s \le \int_0^1 C\epsilon|\gamma|^2 d\epsilon, \tag{4.50}$$

with $C = C(s,\rho_1)$ of $O(1)$. Collecting the estimates (4.45), (4.46), (4.50) for the three parts of V_{R_1} and using $|\delta| \le c_\delta|\gamma|$ in (4.45), we therefore have the statement.

■

Proof of Theorem 4.2: Fix $s \ge 1$, and assume that $d(t\Omega^{-1}) \in H^q$, $q > \frac{3}{2}$. We want to estimate the size of $r_1(t)$ in (4.17). The initial condition will be $r_1(0) = z_1(0) - \mathcal{T}_1^1(z(0))$. Let $s' = s+2$, and consider Proposition 4.1 with $\rho_0 = \beta$ of $O(1)$. Applying Proposition 4.1 to the first component of the quartic normal form equation (4.14), we can choose $|\gamma|$ sufficiently small so that $z_1(0) \in B^{s'}(\rho_0)$ implies $z(t) \in B^{s'}(2\rho_0)$ for a time interval of $O(|\gamma|^{-1})$. We also have existence in $H^{s'}$ for the full system (4.1) with the initial condition $x_1(0) = z_1(0) \in H^{s'}$, also over a time interval of $O(|\gamma|^{-1})$. In particular, the H^s norm of $x_1(t)$ remains of $O(1)$ over that time. We set $\rho_1 = 2\rho_0$, $\alpha_1 = \frac{3}{2}$, and let $|\gamma| \le \gamma_0$ so that Lemmas 4.3, 4.4 apply. Thus $\mathcal{T}_1^1(z_1(t), \Omega t)$ is well defined over the time interval of the local existence theory. Define $\mathcal{R}_1(t)$ by $r_1(t) = e^{tL}\mathcal{R}_1(t)$ and note that the operator $U_t = e^{tL}$ is an isometry in H^s and commutes with L (in a dense subset of H^s). By (4.17), $\mathcal{R}_1(t)$ then satisfies

$$\dot{\mathcal{R}}_1 = U_{-t}(V_{H_4}(x) - V_{H_4}(\mathcal{T}_1(z))) + U_{-t}[D_1\mathcal{T}_1^1(z)]V_{R_1}(z). \tag{4.51}$$

Note that $V_{H_4}(y_1, y_2) : B^s(\rho') \times \mathbf{R} \to H^s$, $s' > \frac{1}{2}$, is Lipschitz continuous in y_1, uniformly in y_2, and that for ρ' of $O(1)$ the Lipschitz constants are of $O(|\gamma|)$, e.g see (4.4). From (4.51) we then have that for t in the time interval of the local existence theorem we have

$$||\mathcal{R}_1(t)||_s \leq ||\mathcal{R}_1(0)||_s + \int_0^t C|\gamma| ||\mathcal{R}_1(t)||_s d\tau + \int_0^t ||D_1\mathcal{T}_1^1(z(\tau))||_{s,s} ||V_{R_1}(z(\tau))||_s d\tau, \tag{4.52}$$

with $C = C(s, \rho_1)$ of $O(1)$. Bounding $\mathcal{R}_1(0)$, $D_1\mathcal{T}_1^1$ and V_{R_1} by Lemmas 4.3, 4.4 and 4.5 respectively, and using Gronwall's inequality, (4.52) yields

$$||r_1(t)||_s = ||\mathcal{R}_1(t)||_s \leq C_0|\gamma|, \quad t \in [0, \tilde{C}_0|\gamma|^{-1}], \tag{4.53}$$

with C_0, $\tilde{C}_0$ that depend on s, $|\Omega|$, ρ_0, c_δ and are of $O(1)$.

To extend the error estimate to $d(t\Omega^{-1}) \in L^2$, consider two dispersion management functions $d(\Omega^{-1}t)$ in H^q, $q > \frac{3}{2}$, $d_1(\Omega^{-1}t)$ in L^2 with the same average δ (and period $T = \frac{2\pi}{\Omega}$). Also let $\tilde{\Lambda}_1(t) = \int_0^t d_1(\sigma)d\sigma$ (following the notation of (2.4)). The nonlinearities of (4.1), (4.2) with $\tilde{\Lambda}(t)$ replaced by $\tilde{\Lambda}_1(t)$ are denoted by F_1, G_1 respectively, and we compare the solutions of $\dot{a}_1 = La_1 + F_1(a_1, t)$ and $\dot{b}_1 = Lb_1 + G_1(b_1, t)$, 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), (4.2). Define Δ by

$$\Delta = \sqrt{T}||\tilde{d}(t) - \tilde{d}_1(t)||_0 \geq |\tilde{\Lambda}(t) - \tilde{\Lambda}_1(t)|. \tag{4.54}$$

Choosing $d(t\Omega^{-1})$ that is close to $d_1(t\Omega^{-1})$, i.e. letting $\Delta > 0$ small, we observe that (4.53) is uniform in Δ. This is because the estimates for $\mathcal{T}_1^\epsilon$ and $D_1\mathcal{T}_1^\epsilon$ only require $d(\Omega^-1t) \in L^2$, and the extra derivatives were used to make $D\mathcal{T}^\epsilon$ well-defined, but do not appear in the quantities involved in (4.53). From

$$||a_1(t) - b_1(t)||_s \leq ||a_1(t) - a(t)||_s + ||a(t) - b(t)||_s + ||b(t) - b_1(t)||_s \tag{4.55}$$

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

$$||F(U_\tau A(\tau), \tau) - F_1(U_\tau A_1(\tau), \tau)||_s \leq C_F(||A(\tau) - A_1(\tau)||_s + C\Delta||A(\tau)||_{s'}), \tag{4.56}$$

with L_F a constant that is quadratic in $||A||_s, ||A_1||_s$, and C of $O(1)$. 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 [BCP], (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 $s > \frac{1}{2}$ comes from the local existence theory, and may not be available for longer times. It appears that 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.

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 likely 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 believe that this question can be partially addressed with some numerical work in the future.

5. Acknowledgments

The author is supported by NSF grant DMS-9810751.

6. REFERENCES

[AB] M.R. Ablowitz, G. Biondini: Multiscale pulse dynamics in communication systems with strong dispersion management, Opt. Lett. 23, 1668-1670 (1998)

[B] D. Bambusi: Nekhoroshev estimates for small amplitude solutions in nonlinear Schrödinger equations, Math. Z. 230, 345-387 (1999)

[BCP] D. Bambusi, A. Carati, A. Ponno: The nonlinear Schrödinger equation as a resonant normal form, Disc. Cont. Dyn. Syst.-Ser. B 2, 109-128 (2002)

[CMK] V. Cautaerts, A. Maruta, Y. Kodama: On the dispersion managed soliton, Chaos 10 515-528 (2003)

[GT] I.R. Gabitov, S.K. Turytsin: Averaged pulse dynamics in a cascaded transmission line with passive dispersion compensation, Opt. Lett. 21, 327-329 (1996)

[K] S.B. Kuksin: Analysis of Hamiltonian PDEs, Oxford Univ. Press, Oxford (2000)

[KP] S.B. Kuksin, J. Pöschel: Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. 143 , 149-179 (1996)

[Kun1] M. Kunze: Bifurcation from the essential spectrum without sign condition on the nonlinearity, Proc. Roy. Soc. Edinburgh Sect. A 131, 927-943 (2001)

[Kun2] M. Kunze: The singular perturbation limit of a variational problem from nonlinear fiber optics, Physica D, article in press

[Kur] C. Kurtzke: Suppression of fiber nonlinearities by appropriate dispersion management, IEEE Photon. Tech. Lett. 5, 1250-1253 (1993)

[LP] T.I. Lakoba, D.E. Pelinovsky: Persistent oscillations of scalar and vector dispersion managed solitons, Chaos 10, 539-550 (2000)

[LYKM] T.I. Lakoba, J.Yang, D.J. Kaup, B.A. Malomed: Conditions for stationary pulse propagation in the strong dispersion management regime, Opt. Comm. 149, 366-375 (1998)

[MJZG] J. Moeser, C.K.R.T. Jones, V. Zharnitsky, I.R. Gabitov: Ground states for the nonlinear Schrödinger equation with higher order dispersion management, preprint (2002)

[MGJ] J. Moeser, I. Gabitov, C.K.R.T. Jones: Pulse stabilization by high-order dispersion management, Opt. Lett. 27, 2206-2208 (2002)

[NDFK] J.H.B. Nijhof, N.J. Doran, W. Forysiak, F.M. Knox: Stable soliton-like propagation in dispersion managed systems with net anomalous, zero, and normal dispersion, Electr. Lett 33, 1726-1728 (1997)

[PT] J. Pöschel, E. Trubowitz: Inverse spectral theory, Academic Press, Boston (1987)

[P] J. Pöschel: On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem of Bambusi, Nonlinearity 12 1587-1600 (1999)

[SKDBB] N.J. Smith, F.M. Knox, N.J. Doran, K.J. Blow, I. Bennion: Enhanced power solitons in optical fibres with periodic dispersion management, Electr. Lett. 32, 54 -55 (1996)

[TM] S. K. Turytsin, V.K. Mezentsev: On the theory of chirped optical solitons in fiber lines with varying dispersion, JETP Lett. 68, 830-836 (1998)

[YK] T.S. Yang, W.L. Kath: Radiation loss of dispersion-managed solitons in optical fibers, Physica D 149, 80-94 (2001)

[YKT] T.S. Yang, W.L. Kath, S. K. Turytsin: The multiple scale averaging and dynamics of dispersion-managed optical solitons, J. Engin. Math. 36, 136-184 (1999)

[ZGJT] V. Zharnitsky, E. Grenier, C.K.R.T. Jones, S.K. Turytsin: Stabilizing effect of dispersion management, Physica D 152-3, 794-817 (2001)