Near-monochromatic water waves on the sphere

Panayotis Panayotaros

Department of Mathematics

The University of Texas at Austin

Austin, TX 78712-1802

U.S.A.

and

FENOMEC - I.I.M.A.S., U.N.A.M. *

Apdo. Postal 20-726

Depto. Matematicas y Mecanica

01000 Mexico D.F.

Mexico

panos@uxmym1.iimas.unam.mx

panos@math.utexas.edu

* *current address*

Abstract.

We study analytically and numerically a class of traveling and standing waves in a model of weakly non-linear gravity water waves on the sphere. These waves are "near-monochromatic" in space, i.e. their amplitude consists of one spherical harmonic plus small corrections, and we see numerically that they retain this property for long time. A main feature of the model we consider is that it possesses a Hamiltonian structure. This structure is preserved by our numerical implementation, and we use formal and rigorous arguments from classical perturbation theory to understand the numerical observations.

1. Introduction

We present an analytical and numerical study of a class of weakly non-linear traveling and standing waves in a model of surface water waves on the sphere. In earlier work on the same model we have established the existence of families of approximate periodic orbits that are "near-monochromatic" in space, i.e. the wave amplitude consists of a spherical harmonic plus small corrections (see [1]). The present paper concerns the stability of such motions, emphasizing their spatial features and the transfer of energy between the various spherical harmonic modes. Our main numerical observation is that trajectories starting from initial conditions consisting of one spherical harmonic plus small terms exhibit small transfer of energy to other modes for long times. In addition to supporting this claim we also propose a theoretical explanation of the observed behavior.

Our study is based on a Hamiltonian model for free surface potential flow of a fluid layer surrounding a gravitating sphere. The model was developed in [1] and is a generalization to the spherical case of the Hamiltonian theory of water waves of Zakharov [2] (see also [3]). The equations of motion are integro-differential and in our derivation of approximate Hamiltonians for small amplitude oscillations we have incorporated work of [4] and [5] on the non-local operators appearing in the problem. An advantage of the Hamiltonian formalism is that it yields Hamiltonian Galerkin truncations of the equations of motion. Thus we solve numerically a finite dimensional Hamiltonian system, and we also use methods from canonical perturbation theory to understand the numerical observations.

The theoretical tool we will use is a Birkhoff normal form calculation. We see that the quartic normal form Hamiltonian of the system we integrated numerically possesses several quantities ("asymptotic invariants") that evolve slowly and control the transfer of energy between the different spherical harmonic modes for long time. In the case of axisymmetric oscillations the quartic normal form is completely integrable and the existence of asymptotic invariants gives a good explanation of the numerical results. For general initial conditions, the asymptotic invariants leave only a few modes to which a near-monochromatic oscillation may decay. For certain (non-axisymmetric) initial conditions a heuristic argument suggests that decay to even these few modes must be slow.

Note that although our arguments are detailed for the particular truncations we have considered in the numerical simulation, they also apply to larger finite dimensional truncations of the equations of motion, and also to models with higher order non-linearity. However we can not at present extend our theory to the infinite dimensional case.

The Hamiltonian formalism and Birkhoff normal forms for water wave problems have been considered by several authors. Of special interest are limiting cases where quartic normal forms are completely integrable, see e.g. [6], [7] on 2-D deep water waves. Such stronger results rely on detailed information on the non-linearities, and since they concern different geometries, their connection with the spherical problem is not clear. We therefore discuss them briefly in the conclusion of this paper. Also, there is an extensive literature on the drift of the actions in near-integrable Hamiltonian systems. These ideas should be particularly relevant to axisymmetric motions (see especially [8] and references therein), and they suggest that other motions, e.g with many spherical harmonic modes should also be stable for long time. One difficulty in applying such results to our problem directly is that we can not control the distance from the quartic resonances so well. Also there may be regions of phase space where the quartic normal form is degenerate. The situation for general initial conditions is more complicated since the normal form is not integrable, and this is where our numerical results and the theory are more interesting and novel.

The Hamiltonian water wave model and our numerical implementation are described in section 1. In section 2 we present the results of the numerical simulations. In section 4 we give formal and heuristic arguments to explain the numerical observations, and in section 5 we give rigorous versions of some of these arguments. In section 6 we summarize and discuss our results.

2. Hamiltonian formulation and numerical implementation

We consider a fluid layer of thickness h surrounding a gravitating sphere of radius b. Using polar coordinates r = radius, ϑ = polar, φ = azimuth, the outer (or "free") surface of the layer will be the set of points with radius $r(\vartheta,\varphi) = \rho + \eta(\vartheta,\varphi)$, with $\rho = b + h$. We are interested in the oscillations of the free surface around the quiescent state with $\eta(\vartheta,\varphi) = 0$, and we assume that the flow inside the layer is potential. Denoting the hydrodynamic potential by ϕ, Euler's equations for free surface potential flow take the form (see e.g. [9])

$$\frac{\partial \eta}{\partial t} = \frac{\partial \phi}{\partial r} - \frac{1}{r^2}\frac{\partial \phi}{\partial \vartheta}\frac{\partial \eta}{\partial \vartheta} - \frac{1}{r^2 \sin^2 \vartheta}\frac{\partial \phi}{\partial \varphi}\frac{\partial \eta}{\partial \varphi}, \tag{2.1}$$

and

$$\frac{\partial \phi}{\partial t} = -\frac{1}{2}|\nabla \phi|^2 + \frac{1}{\rho + \eta}, \tag{2.2}$$

at the free surface,

$$\Delta \phi = 0 \tag{2.3}$$

in the region occupied by the fluid, and

$$\frac{\partial \phi}{\partial r} = 0 \tag{2.4}$$

at the bottom $r = b$. Note that the wave amplitude $\eta(\vartheta,\varphi)$ and the surface potential $\Phi(\varphi,\vartheta) = \phi(\varphi,\vartheta,\rho + \eta(\vartheta,\varphi))$ determine the hydrodynamic potential ϕ inside the layer uniquely at each instant.

The above equations can also be written as a Hamiltonian system (the Hamiltonian formulation for the sphere was developed in [1], here we only outline the relevant results). The canonical variables are the wave amplitude η, and the hydrodynamic potential at the surface Φ, while the Hamiltonian H is the total energy (kinetic + potential) of the system. The kinetic energy part of the Hamiltonian can be formally expanded in powers of the wave amplitude η, and we may write $H = \sum_{j=0}^{\infty} H_j$, with H_j of order $j+2$ in the canonical variables. To rewrite (2.1)-(2.4) as Hamilton's equations it is convenient to use the variables $\tilde{\eta}$ and $\tilde{\Phi}$, defined by $\tilde{\eta} = \eta\left(1 + \frac{\eta}{2\rho}\right)$, $\tilde{\Phi} = \Phi\left(1 + \frac{\eta}{\rho}\right)$. ¿From now on all quantities will be expressed in terms of $\tilde{\eta}$ and $\tilde{\Phi}$ and we drop the tilde from the notation. Note that the H_j are modified so that they still consist of the terms of order $j+2$ in the new variables. Also, we expand η and Φ in spherical harmonics $Y_\gamma(\vartheta, \varphi)$, with $\gamma = [l, m]$, $l = 1, 2, \ldots$, $m = -l, \ldots, l$ (see e.g. [10]), i.e. $\eta = \sum_\gamma \eta_\gamma Y_\gamma$, $\Phi = \sum_\gamma \Phi_\gamma Y_\gamma$. Then the equations of motion for each mode η_γ, Φ_γ are

$$\dot{\eta}_\gamma = \frac{\partial H}{\partial \Phi^*_\gamma}, \quad \dot{\Phi}_\gamma = -\frac{\partial H}{\partial \eta^*_\gamma}, \tag{2.5}$$

and the quadratic and cubic parts of the Hamiltonian are

$$H_0 = \frac{\rho^2}{2} \sum_\gamma \left(\frac{u_\gamma{}'(\rho)}{u_\gamma(\rho)} \Phi_\gamma \Phi^*_\gamma + \eta_\gamma \eta^*_\gamma \right), \quad \text{with} \quad u_\gamma(r) = (l+1)\left(\frac{r}{b}\right)^l + l\left(\frac{b}{r}\right)^{l+1}, \tag{2.6}$$

$$H_1 = \sum_{\gamma_1, \gamma_2, \gamma_3} I_{\gamma_1, \gamma_2, \gamma_3} \Phi_{\gamma_1} \Phi_{\gamma_2} \eta_{\gamma_3}, \quad \text{with} \tag{2.7}$$

$$I_{\gamma_1,\gamma_2,\gamma_3} = \frac{\rho^2}{2}\left(\frac{u''_{\gamma_2}(\rho)}{u_{\gamma_2}(\rho)} - \frac{u'_{\gamma_1}(\rho)}{u_{\gamma_1}(\rho)} \frac{u'_{\gamma_2}(\rho)}{u_{\gamma_2}(\rho)} \right) \int_{S^2} Y_{\gamma_1} Y_{\gamma_2} Y_{\gamma_3} + \frac{1}{2} \int_{S^2} Y_{\gamma_1} \nabla Y_{\gamma_2} \cdot \nabla Y_{\gamma_3}.$$

In the numerical experiments we have set $b = 1.0$ and $h = 0.2$. By the dimensional analysis of [1], and with our choice of b and h, equations (2.5) with the cubic Hamiltonian $H = H_0 + H_1$ should describe small amplitude oscillations adequately. Writing the Hamiltonian in (2.5) as $H = \sum_{j=0}^\infty \epsilon^j H_j$, with ϵ the ratio of a typical wave amplitude to the depth h (and setting $b = 1$), equations (2.5)-(2.7) may be considered as being in dimensionless form.

To solve the initial value problem numerically, we consider Galerkin truncations of the equations (2.5), i.e. the summations in the Hamiltonian of (2.6) and (2.7) are replaced by summations over a finite set of spherical harmonic mode indices. In our observations we have used a Galerkin truncation that includes all spherical harmonic modes with $l \leq l_{\max} = 8$. Taking into account the real and imaginary parts of η_γ, Φ_γ and the reality condition $\eta_{[l,-m]} = \eta^*_{[l,m]}$, $\Phi_{[l,-m]} = \Phi^*_{[l,m]}$ we then have a Hamiltonian system in $\mathbf{R}^{2N}$, with $N = 80$ ($N = l^2_{\max} + 2l_{\max}$). We may also express the equations of motion and the Hamiltonian (2.5)- (2.7) using the variables a_γ, a^*_γ, defined by

$$\eta_\gamma = \frac{\sqrt{2\omega_\gamma}}{2}(a_\gamma + a^*_\gamma), \quad \Phi_\gamma = -i\frac{\sqrt{2}}{2\sqrt{\omega_\gamma}}(a_\gamma - a^*_\gamma), \quad \text{with} \quad \omega^2_\gamma = \frac{u_\gamma{}'(\rho)}{u_\gamma(\rho)} \tag{2.8}$$

i.e. ω_γ is the dispersion relation. Letting $q_\gamma = \mathrm{Re}\,(a_\gamma)$ and $p_\gamma = \mathrm{Im}\,(a_\gamma)$, Hamilton's equations (2.5) become

$$\dot{q}_\gamma = \frac{1}{2}\frac{\partial H}{\partial p_\gamma}, \quad \dot{p}_\gamma = -\frac{1}{2}\frac{\partial H}{\partial q_\gamma}. \tag{2.9}$$

Equations (2.9), with the Hamiltonian $H = H_0 + \epsilon H_1$, and the truncation and parameters b and h above, are the equations we integrate numerically. To integrate the equations we used the Fortran package LSODE, available through netlib (see [11]). The integration algorithm we chose is a predictor-corrector multi-step method, where the prediction step is performed by an (explicit) Adams-Bashforth formula, and the correction step is done by an (implicit) Adams-Moulton formula (see [12]). The choice of step-size and order is performed automatically during the integration so that the computed values of q_γ, p_γ are within a user-specified local error tolerance. The estimation of the local error, and the procedure for choosing the step-size and the order are described in [13], p.155 (see also [14]). Tests of the accuracy of the numerical integration are reported in [15].

3. Numerical Observations

We are interested in the evolution from initial conditions where one spherical harmonic mode a_Γ has initial amplitude $O(1)$ and all other modes a_γ with $\gamma \neq \Gamma$ have initial amplitude $<< O(1)$. We refer to such initial conditions as "near-monochromatic". Our basic observation is that numerical trajectories starting from near-monochromatic initial conditions remain near-monochromatic for long time.

In presenting the numerical evidence for this behavior we will consider separately two cases, axisymmetric oscillations and general oscillations. Axisymmetric oscillations are trajectories belonging to the vector space $V_A = \mathrm{span}\{a_\gamma : \gamma = [l, 0]\}$. The invariance of $V_A \subset \mathbf{R}^{2N}$ under the flow of Hamilton's equations (2.9) follows from the fact that the coefficients $I_{\gamma_1\gamma_2\gamma_3}$ of the cubic Hamiltonian H_1 in (2.7) vanish if $m_1 + m_2 + m_3 \neq 0$. Although the procedure we followed and the numerical results are very similar for both axisymmetric and general initial conditions, the distinction is useful in the theoretical discussion of the next section. Also note that in all the numerical experiments described below, the motion on the plane a_Γ is (roughly) a sinusoidal oscillation with an amplitude modulation. Thus, in space, the motions are close to those of the linear equations with

$a_\gamma(0) = 0$ for $\gamma \neq \Gamma$, and can be characterized as traveling waves (rotating around the north-south axis) for $\Gamma = [l, m]$ with $m \neq 0$, and standing waves in the axisymmetric case.

In a first set of measurements, summarized below, the goal was to quantify our observation that near-monochromatic solutions do not decay for long time. The emphasis was on trying a large number of initial conditions. The reliability of these numerical observations will be discussed afterwards.

Axisymmetric motions: We have considered a set of axisymmetric initial conditions of the form $|a_\Gamma(0)| = 1.0$, and $|a_\gamma(0)| \leq 5 \times 10^{-2}$ for $\gamma \neq \Gamma$. The index Γ in this set ranges from $\Gamma = [1, 0]$ to $\Gamma = [5, 0]$. The (non-linearity) parameter ϵ was varied in the interval $[10^{-4}, 10^{-2}]$. Note that for ϵ around 5×10^{-2}, depending on the initial condition, trajectories start to become unbounded in finite time. Since we are considering trajectories starting from different initial conditions, we want to avoid these "blow-ups", (which, incidentally, are not due to numerical instability, see [15]). The observation time T is 250 (i.e. $50 - 80$ periods of the linear motion, depending on the l_Γ). The local error tolerance W (see [13] p.155) is 10^{-9}. For such W, energy is conserved to $10^{-5} - 10^{-5}$ over $T = 250$ (for total energy ~ 1.0). For each mode a_γ we observe the amplitude variation $\Delta A_\gamma(T)$, defined by

$$\Delta A_\gamma(T) \equiv \sup_{t \in [0,T]} |A_\gamma(t) - A_\gamma(0)|, \quad A_\gamma(t) = \sqrt{a_\gamma(t) a^*_\gamma(t)}. \tag{3.1}$$

Our main observations are as follows.

(i) The mode a_Γ does not decay. The range of observed $\Delta A_\Gamma(T)$ for $T = 250$ is: for $\epsilon = 10^{-4}$, $\Delta A_\Gamma(T) \in [10^{-5}, 10^{-4}]$; for $\epsilon = 10^{-3}$, $\Delta A_\Gamma(T) \in [10^{-4}, 10^{-3}]$; for $\epsilon = 10^{-2}$, $\Delta A_\Gamma(T) \in [10^{-3}, 10^{-2}]$. For a given initial condition, $\Delta A_\Gamma(T)$ is roughly (but not exactly) proportional to ϵ.

(ii) The transfer of energy (or amplitude) to other modes is limited. If we let $a_{\tilde{\gamma}}$ be the mode with the largest amplitude variation $\Delta A_{\tilde{\gamma}}(T)$ over T, we have seen that for $\epsilon = 10^{-4}$, $\Delta A_{\tilde{\gamma}}(T) \in [10^{-4}, 10^{-3}]$; for $\epsilon = 10^{-3}$, $\Delta A_{\tilde{\gamma}}(T) \in [10^{-3}, 10^{-2}]$; for $\epsilon = 10^{-2}$, $\Delta A_{\tilde{\gamma}}(T) \in [10^{-2}, 1.7 \times 10^{-1}]$. Other modes a_γ, with $\gamma \neq \Gamma, \tilde{\gamma}$, may have $\Delta A_\gamma(T)$ with same order of magnitude as $\Delta A_{\tilde{\gamma}}(T)$. Clearly for $\epsilon = 10^{-2}$ the energy transfer to other modes is appreciable, but still the mode Γ dominates. We generally observe that, for given initial conditions, the $\Delta A_\gamma(T)$ are roughly (but not exactly) proportional to ϵ for the various modes a_γ. Also, in all experiments we see that $A_\gamma(t)$ oscillates so that in general $\Delta A_\gamma(T)$ is attained quite early, by $T \sim 50$. This fact suggests that the measured drift of the amplitudes should give information for much longer intervals.

General motions: We consider initial conditions of the form $|a_\Gamma(0)| = 1.0$ and $|a_\gamma(0)| \leq 10^{-2}$, for $\gamma \neq \Gamma$. The index Γ ranges over the modes with $l \leq 4$. The error tolerance W, the range of ϵ and the observation time T are the ones used for the axisymmetric case. Again, we measure the quantity $\Delta A_\gamma(T)$ defined as in (3.1). Our main observations are summarized below.

(i) The mode Γ does not decay. Overall we have for $\epsilon = 10^{-4}$, $\Delta A_\Gamma(T) \in [10^{-6}, 10^{-5}]$; for $\epsilon = 10^{-3}$, $\Delta A_\Gamma(T) \sim 10^{-4}$; for $\epsilon = 10^{-2}$, $\Delta A_\Gamma(T) \sim 10^{-2}$.

(ii) Transfer of energy to other modes is limited. We generally observe that for $\epsilon = 10^{-4}$, $\Delta A_{\tilde\gamma}(T) \in [10^{-4}, 10^{-3}]$; for $\epsilon = 10^{-3}$, $\Delta A_{\tilde\gamma}(T) \in [10^{-3}, 10^{-2}]$; for $\epsilon = 10^{-2}$, $\Delta A_{\tilde\gamma}(T) \in [10^{-2}, 1.5 \times 10^{-2}]$. Again, the transfer of amplitude to other modes can be appreciable as ϵ increases, but the mode Γ dominates, and also the $\Delta A_\gamma(T)$ is attained by $T \sim 50$.

In a second set of experiments we assess the accuracy and reliability of the above numerical results. Our strategy is to consider a smaller set of initial conditions and integrate numerically for longer time and with smaller local error tolerance W. In particular, we repeated the first set of experiments with the same initial conditions, but for $\epsilon = 0.001$ only. The integration time is now $4T = 1000$ and $W = 10^{-15}$ and 10^{-9}. We observe that:

(i) Generally, the $\Delta A_\gamma(T)$ obtained with $W = 10^{-15}$ and 10^{-9} agree to at least $1{-}3\%$ (usually to higher accuracy). Similar agreement is observed between the $\Delta A_\gamma(4T)$ obtained with $W = 10^{-15}$ and 10^{-9}

(ii) Generally, $\Delta A_\gamma(T)$ and $\Delta A_\gamma(4T)$ are very close, to within $1 - 3\%$, for both $W = 10^{-15}$ and 10^{-9}. Typical data indicating (i) and (ii) are in Tables 1 and 2, where the dominant modes a_Γ are $a_{[3,2]}$ and $a_{[4,3]}$ respectively. The second line on each table shows the largest $\Delta A_\gamma(T)$, i.e. $\Delta A_{\tilde\gamma}(T)$, and the other lines show other modes with the highest amplitude variation.

(iii) The largest discrepancies between $\Delta A_\gamma(T)$ and $\Delta A_\gamma(4T)$ are observed for the mode a_Γ. Such discrepancies may be due to some kind of numerical drift of the amplitudes. The most dramatic discrepancy we observed is shown in Table 3. It was only observed for $W = 10^{-9}$ and was suppressed with $W = 10^{-15}$. Note, that for the times we have integrated such drifts are rare, moreover even in the worst cases they do no alter our quantitative results significantly.

¿From the above comparisons, we may conclude that the results obtained for $T = 250$ and with $W = 10^{-9}$ give reliable information on the change of amplitude of the various modes. Our observations that $\Delta A_\gamma(T)$ is attained early, and that $\Delta A_\gamma(4T)$ is only slightly larger than $\Delta A_\gamma(T)$ suggest non-decay of the mode a_Γ for longer times. Also, the approximate proportionality of the ΔA_γ to ϵ and the results of the second set of experiments suggest that the ΔA_γ we corresponding to different integration times and W should be close for other values of ϵ as well.

4. Normal forms

We now propose an explanation for the stability of the near-monochromatic solutions. We employ formal and heuristic arguments, some of which form the basis of the rigorous statements of the next section. The theory is based on a Birkhoff normal form calculation. In particular, we bring the system to a Birkhoff normal form, which truncated to $O(\epsilon^3)$, has families of near-monochromatic solutions and several constants of motion that give information on the stability of such solutions.

The starting point is the Hamiltonian $H = H_0 + \epsilon H_1$ of (2.6), (2.7). Since we are interested in explaining the numerical simulations, we consider Galerkin truncations to modes with $l \leq l_{\max} = 8$, i.e. the phase space is $\mathbf{R}^{2N}$, $N = 80$. We will also indicate which statements are valid for other truncations.

We will use the following properties of ω_γ and $I_{\gamma_1,\gamma_2,\gamma_3}$:

(i) Let $\gamma = [l, m]$, and ω_γ the dispersion relation of (2.8). Then ω_γ satisfies:

(4.1) $$\omega_\gamma = \omega(l), \quad \text{and} \quad \omega(l) > 0, \quad \omega'(l) < 0, \quad \omega''(l) < 0; \quad \forall l > 0.$$

(ii) Let $\gamma = [l_i, m_i]$, $i = 1, 2, 3$. The coefficients $I_{\gamma_1,\gamma_2,\gamma_3}$ of the cubic Hamiltonian in (2.7) satisfy

(4.2) $$I_{\gamma_1,\gamma_2,\gamma_3} \neq 0 \Rightarrow |l_1 - l_2| \leq l_3 \leq l_1 + l_2 \quad \text{and} \quad m_1 + m_2 + m_3 = 0,$$

$$\forall l_i \in \mathbf{Z}^+, \quad m_i = -l_i, \ldots, l_i, \quad i = 1, 2, 3.$$

(iii) Let ω_γ be as in (2.8). Consider the equation

$$\omega_{\gamma_1} + \omega_{\gamma_2} = \omega_{\gamma_3} + \omega_{\gamma_4}, \tag{4.3}$$

for $l_1, \ldots l_4$ that satisfy

$$\exists l \in \mathbf{Z}^+ \quad \text{s.t.} \quad |l_1 - l_2| \leq l \leq l_1 + l_2 \quad \text{and} \quad |l_3 - l_4| \leq l \leq l_3 + l_4. \tag{4.4}$$

Then the only positive integers $l_1, \ldots l_4 \leq l_{\max} = 8$ that satisfy (4.3) and (4.4) are the trivial solutions $l_1 = k_+$, $l_2 = k_-$, $l_3 = k_\pm$, $l_4 = k_\mp$.

Properties (i) and (ii) have been verified for the particular problem in [1], but they are of wider applicability for waves on the sphere. The fact that the dispersion depends on l alone, and property (ii) are consequences of spherical symmetry. The other properties in (i) make ω_γ analogous to "non-decay" dispersion relations encountered, for instance, in plasma physics. Property (iii) is verified numerically for the particular b, h we considered. We expect that if we fix any $l_{\max}$ and vary $\beta = \frac{h}{b} > 0$, the property will hold for an open (non-empty) set of β. However we can not estimate how far we are from solving the Diophantine equation (4.3). Note that such an estimate would not be uniform in $l_{\max}$. As $l \to \infty$, $\omega(l) \to \sqrt{l}$ and (4.3), (4.4) has non-trivial solutions.

The next propositions are consequences of the above properties, and follow immediately from results of a normal form calculation found in [1], where we refer for the proofs. Here we add some remarks and corollaries that relate to the numerical experiments.

Proposition 4.1 Consider the Galerkin truncations to modes with $l \leq l_{\max} = 8$ of the Hamiltonian system (2.9) with $H = H_0 + \epsilon H_1$, H_0, H_1 as in (2.6), (2.7). Then properties (i) and (ii) above hold and imply that there exists a formal canonical transformation $\Phi^{(\epsilon S_1)} : \mathbf{R}^{2N} \to \mathbf{R}^{2N}$ ($N = 80$) so that $H \circ \Phi^{(\epsilon S_1)} = H_0 + \epsilon^2 G_2 + O(\epsilon^3)$, and G_2 is quartic in the canonical variables (terms $O(j)$ are of order $j + 2$).

The canonical transform is produced by the so called Lie-series method (see e.g. [16]), in particular $H \circ \Phi^{(\epsilon S_1)} = \exp Ad_{\epsilon S_1} H$, where S_1 is an appropriate function on $\mathbf{R}^{2N}$. The transform $\Phi^{(\epsilon S_1)}$ is a formal power series, and its domain will be specified in the next section. The gist of Prop. 4.1 is that properties (i) and (ii) imply that there are no cubic resonances. Moreover we can bound away cubic resonances uniformly in

$l_{\max}$ and therefore the proposition holds for arbitrary Galerkin truncations and for the infinite dimensional system as well.

Proposition 4.2 Consider the Hamiltonian $H \circ \Phi^{(\epsilon S_1)}$ of proposition 4.1, and the corresponding Hamiltonian system in the Galerkin truncation to modes with $l \leq l_{\max} = 8$. Then property (iii) holds and implies that there exists a formal canonical transformation $\Phi^{(\epsilon^2 S_2)} : \mathbf{R}^{2N} \to \mathbf{R}^{2N}$ such that $H \circ \Phi^{(\epsilon S_1)} \circ \Phi^{(\epsilon^2 S_2)} = H_0 + \epsilon^2 N_2 + O(\epsilon^3)$, and the (resonant quartic) term N_2 is a linear combination of monomials $a_{\gamma_1} a_{\gamma_2} a^*_{\gamma_3} a^*_{\gamma_4}$ with $l_3 = l_1$ and $l_4 = l_2$ (or $l_3 = l_1$, $l_4 = l_2$), and $m_1 + m_2 = m_3 + m_4$.

The canonical transform is again produced using the Lie-series and $H \circ \Phi^{(\epsilon S_1)} \circ \Phi^{(\epsilon^2 S_2)} = \exp Ad_{\epsilon^2 S_2} H \circ \Phi^{(\epsilon S_1)}$, with an appropriate function S_2. Proposition 4.2 holds for finite dimensional systems, because of the inapplicability of property (iii) in the limit $l_{\max} \to \infty$.

We now consider properties of the Hamiltonian system (2.9) with the quartic normal form Hamiltonian $\overline{H} = H_0 + \epsilon^2 N_2$ in the phase space of the modes with $l \leq l_{\max} = 8$. We introduce the action variables $J_\gamma = a_\gamma a^*_\gamma$ and the angles $\theta_\gamma = Arg(a_\gamma)$, $\forall \gamma$ with $l \leq l_{\max} = 8$. Away from the hyper-planes $a_\gamma = 0$ the Hamiltonian $\overline{H}$ is real analytic in the J_γ, θ_γ. Hamilton's equations become

$$\dot{J}_\gamma = -\frac{\partial \overline{H}}{\partial \theta_\gamma}, \quad \dot{\theta}_\gamma = \frac{\partial \overline{H}}{\partial J_\gamma}. \tag{4.5}$$

Form Proposition 4.2 we immediately have the following.

Corollary 4.3 The Hamiltonian $\overline{H} = H_0 + \epsilon^2 N_2$, restricted to the subspace of axisymmetric solutions is independent of the angles.

Proof: By proposition 4.2 the Hamiltonian $\overline{H}$, restricted to the axisymmetric solutions, consists of monomials of the type $a_{[l_1,0]} a_{[l_2,0]} a^*_{[l_1,0]} a^*_{[l_2,0]} = J_{[l_1,0]} J_{[l_2,0]}$.

■

Corollary 4.4 The Hamiltonian system corresponding to $\overline{H} = H_0 + \epsilon^2 N_2$ has $l_{\max}$ constants of motion I_l given by

$$I_l = \frac{1}{2l+1} \sum_{m=-l}^{l} J_{[l,m]}, \quad l = 1, 2, \ldots, l_{max}.$$

Proof: We make a canonical transformation to a new set of variables $\phi_{[l,j]}$ and $I_{[l,j]}$, $l = 1, \ldots, l_{\max}$, $j = 1, \ldots, 2l+1$ defined by

$$\phi_{[l,1]} = \sum_{m=-l}^{l} \theta_{[l,m]}, \quad l = 1, \ldots, l_{\max},$$

$$\phi_{[l,j]} = \sum_{m=-l}^{l} k^m_{[l,j]} \theta_{[l,m]}, \quad l = 1, \ldots, l_{\max}, \quad j = 2, 3, \ldots, 2l+1,$$

where the $k^i_{[l,j]}$ are chosen so that $\phi_{[l,1]}$ is perpendicular to all the $\phi_{[l,j]}$, $j = 2, 3, \ldots, 2l+1$ and the $\phi_{[l,j]}$, $j = 2, 3, \ldots, 2l+1$ are orthogonal to each other. Also,

$$I_{[l,1]} = \frac{1}{2l+1} \sum_{m=-l}^{l} J_{[l,m]}, \quad l = 1, \ldots, l_{\max},$$

$$I_{[l,j]} = \frac{1}{\sum_{m=-l}^{l} (k^m_{[l,j]})^2} \sum_{m=-l}^{l} k^m_{[l,j]} J_{[l,m]}, \quad l = 1, \ldots, l_{\max}, \quad j = 2, 3, \ldots, 2l+1.$$

¿From (4.5), Hamilton's equations in the new variables are

$$\dot{I}_{[l,j]} = -\frac{\partial \overline{H}}{\partial \phi_{[l,j]}}, \quad \dot{\phi}_{[l,j]} = \frac{\partial \overline{H}}{\partial I_{[l,j]}}, \tag{4.6}$$

for $l = 1, \ldots, l_{\max}$, $j = 1, \ldots, 2l+1$. By proposition 4.4 we see that N_2 does not depend on the angles $\phi_{[l,1]}$ and therefore by (4.6) the $I_{[l,1]}$ are constant. Setting $I_l = I_{[l,1]}$, $l = 1, \ldots, l_{\max}$ we have the proposition.

■

For axisymmetric solutions, corollary 4.3 implies that the amplitude of the modes should change slowly, to $O(\epsilon)$ over time of $O(\epsilon^{-2})$, so that solutions starting from near-monochromatic initial conditions should stay near-monochromatic for long time. For general solutions, corollary 4.4 implies that a near-monochromatic solution can not decay too fast to modes with different l wave-numbers. However the corollary does not give us information on the transfer of energy to modes with the same l wave-number as that of the initial condition. In some cases we can account for the observed behavior by using a heuristic argument that is based on the following statement we have shown in [1].

Proposition 4.5 Let $\Gamma = [l_\Gamma, m_\Gamma]$ be any mode index with $l_\Gamma \leq l_{\max} = 8$. Then the plane $V_\Gamma = \{a \in \mathbf{C}^N : a_\gamma = 0, \quad \gamma \neq \Gamma\}$ is invariant under the flow of the system corresponding to $\overline{H}$.

The idea is now to show that under the flow of the Hamiltonian $\overline{H} = H_0 + \epsilon^2 N_2$ any linear instabilities of the orbits of proposition 4.5 are $O(\epsilon^3)$. We consider Hamilton's equations for $\overline{H}$. For a given Γ (i.e. $a_\Gamma(0) = O(1)$) we substitute $a_\Gamma(t) = A_\Gamma e^{-i\omega_\Gamma t}$ and keep in the equation only terms that are linear in all the a_γ with $\gamma \neq \Gamma$. The resulting system would be the variational equation around the periodic orbit $a_\Gamma(t) = A_\Gamma e^{-i\omega_\Gamma t}$, $a_\gamma(t) \equiv 0$ for $\gamma \neq \Gamma$, off-course $a_\Gamma(t)$ is only $O(\epsilon^3)$ close to what we prescribe. On the invariant plane V_Γ the system is integrable. Near the origin it looks like an oscillator with amplitude dependent frequency. We thus look for possible parametric instability along the other directions. The equations for $\dot{a}_\gamma$, $\gamma \neq \Gamma$, form a non-autonomous linear system with periodic coefficients, of period ω_Γ. For each $\gamma_i \neq \Gamma$, $\dot{a}_{\gamma_i}$ is proportional to terms of type (a) $a_{\gamma_j} a_\Gamma a_\Gamma^*$, and (b) $a^*_{\gamma_j} a_\Gamma a_\Gamma$. We can eliminate non-resonant terms in the equations for $\dot{a}_\gamma$ by averaging, i.e. by writing the equations for the amplitudes $A_\gamma(t)$ defined by $a_\gamma(t) = A_\gamma(t) e^{-i\omega_\gamma t}$ and time-averaging. The terms of type (a) that remain for each $\dot{A}_{\gamma_i}(t)$ are those with $\omega_{\gamma_j} = \omega_{\gamma_i}$ and $m_{\gamma_j} = m_{\gamma_i}$. Such terms only alter the frequencies of $a_\gamma(t)$ and can not give rise to hyperbolic directions. Resonant terms of type (b) that remain after the averaging must satisfy $\omega_{\gamma_i} + \omega_{\gamma_j} = 2\omega_\Gamma$ and $m_{\gamma_i} + m_{\gamma_j} = 2m_\Gamma$. By property (iii) these conditions require $\omega_{\gamma_i} = \omega_{\gamma_j} = \omega_\Gamma$. Therefore instability can occur only along directions with $l = l_\Gamma$. In addition, if $m_\Gamma = l$ or $-l$, then no $O(\epsilon^2)$ terms can induce parametric instability.

5. Stability estimates

The normal form argument of the previous section gives us a partial explanation of the behavior we observed in the numerical experiments. In this chapter we obtain bounds for the variation of (the square of) the mode amplitudes and estimate the time and the size of initial conditions for which such bounds are valid. With these results the normal form argument of the previous section becomes quantitative and rigorous. We summarize the results in two theorems, for axisymmetric and general solutions respectively.

Since all functions in the problem are given explicitly as formal power series on $\mathbf{R}^{2N}$, it is natural to work with their extension to power series on $\mathbf{C}^{2N}$. Specifically, consider

$\mathbf{C}^{2N}$ with coordinates $z = (z_1, z_2, \ldots, z_{2N})$. Let $\varrho = (\rho_1, \ldots, \rho_{2N})$ and $k = (k_1, \ldots, k_{2N})$ be multi-indices. We define the polydisks D_ϱ by

$$D_\varrho = \{z \in \mathbf{C}^{2N} : \quad |z_i| \leq \rho_i\},$$

and we consider formal power series on D_ϱ, i.e. expressions $\sum_{k \geq 0} f_k z^k$, $f_k \in \mathbf{C}$ (using multi-index notation, and also $k \geq 0$ to denote multi-indices for which $k_i \geq 0$, $i = 1, \ldots, 2N$). Let $|| \sum_{k \geq 0} f_k z^k ||_\varrho = \sum_k |f_k| \varrho^k$ and define the set $\mathcal{A}(D_\varrho)$ of analytic functions on D_ϱ by

$$\mathcal{A}(D_\varrho) = \left\{ \sum_{k \geq 0} f_k z^k : \quad \left|\left| \sum_{k \geq 0} f_k z^k \right|\right|_\varrho < \infty \right\}.$$

The spaces $\mathcal{A}(D_\varrho)$, with the norm $|| \, ||_\varrho$ are Banach algebras (see e.g. [17] for their basic properties). For simplicity, here we will consider special polydisks D_ϱ in which $\varrho_i = \rho$, $\forall i = 1, \ldots, 2N$. We denote these polydiscs by D_ρ. The corresponding spaces of analytic functions and norms are denoted by $\mathcal{A}(D_\rho)$ and $|| \, ||_\rho$ respectively. In the Appendix we state three lemmas that follow directly from the above definitions, and that are used in the proofs of theorems 5.1, 5.2 below.

The notation of theorem 5.1 on axisymmetric solutions is as follows. The phase space is $\mathbf{R}^{2N}$, $N = l_{\max} = 8$, and the coordinates x_i, q_i, p_i are defined by $x_i = q_i = q_{[i,0]}$, $x_{i+N} = p_i = p_{[i,0]}$, $i = 1, \ldots, l_{\max}$ (see section 2 for the $q_{[i,0]}$, $p_{[i,0]}$). Also, $x = [x_1, \ldots, x_{2N}]$. The Hamiltonian will be denoted by h and $h = h_0 + \epsilon h_1$, with h_0, h_1 the restriction of H_0, H_1 respectively to the subspace of axisymmetric solutions V_A. Note that the canonical transformations of the propositions 4.1, 4.2 leave V_A invariant. We denote the restrictions of S_1, S_2 to V_A by s_1 and s_2 respectively (i.e. the notation is that of sections 2 and 4, with functions restricted to V_A denoted by small letters). The theorem will involve the constants of lemmas 1, 2, 3 of the Appendix and also c_1 and c_2, the bounds from the cubic and quartic resonances respectively (for $l \leq l_{\max} = 8$).

Theorem 5.1 Consider axisymmetric solutions of the equations (2.9) with the notation above. Let $1 \geq \rho > 5\delta > 0$ and consider initial conditions $x(0) \in D_{\rho_1}$, with $\rho_1 \leq (\sqrt{2} - 1)(\rho - 5\delta)$. Also assume that $\epsilon > 0$ is such that

$$\lambda_1 = \epsilon c_1 \frac{e^2 c^2}{\delta^2} ||h_1||_\rho < 1, \quad \lambda_2 = (\lambda_1)^2 \frac{c_2}{c_1 e^2} < 1, \tag{5.1}$$

$$\epsilon\mu \leq \frac{2-\sqrt{2}}{2}(\rho - 5\delta), \quad \frac{e^2\delta}{2\kappa_1 ||r||_{\rho-2\delta}} \leq 1,$$

with $\mu = O(1)$, $r = O(\epsilon^3)$ defined in lemmas 5.1, 5.2 below. Then, properties (i)-(iii) of section 4 imply that the actions $J_i = q_i^2 + p_i^2$, $i = 1, \ldots, N$ satisfy

$$|J_i(t) - J_i(0)| \leq 8\epsilon\mu + 4\kappa_1 \frac{||r||_{\rho-2\delta}}{\epsilon^2} = O(\epsilon), \quad \forall t \in [0, \delta\epsilon^{-2}). \tag{5.2}$$

In theorem 5.2 for general solutions of (2.9) below, we consider modes up to $l_{\max} = 8$. The phase space will be $\mathbf{R}^{2N}$ with $N = 80$ and we use the coordinates $y_{l^2+l+m} = q_{[l,m]}$, $y_{N+l^2+l+m} = p_{[l,m]}$ (as in section 2). The notation for the Hamiltonian and the functions generating the canonical transformations are as in sections 2 and 4. The constants c_1, c_2 are as in the axisymmetric case.

Theorem 5.2 Consider solutions of the equations (2.9) with the notation above. Let $1 \geq \rho > 5\delta > 0$ and consider initial conditions $y(0) \in D_{\rho_2}$, with $\rho_2 \leq 3^{-1}(4l_{\max} + 2)^{-\frac{3}{2}}(\rho - 5\delta)$. Moreover, let ϵ be such that

$$\Lambda_1 = \epsilon c_1 \frac{e^2 c^2}{\delta^2} ||H_1||_\rho < 1, \quad \Lambda_2 = (\Lambda_1)^2 \frac{c_2}{c_1 e^2} < 1, \tag{5.3}$$

$$\epsilon M \leq 3^{-1}(4l_{\max} + 2)^{-\frac{3}{2}}(\rho - 5\delta), \quad \frac{e^2\delta}{\sqrt{2}(2l_{\max}+1)\kappa_1 ||R||_{\rho-2\delta}} \leq 1,$$

with $M = O(1)$, $R = O(\epsilon^3)$ defined below. Then, properties (i)-(iii) of section 4 imply that the variables $I_i = (2l+1)^{-1} \sum_{m=-l}^{l} (q_{[l,m]}^2 + p_{[l,m]}^2)$, $i = 1, \ldots, N$ satisfy

$$|I_i(t) - I_i(0)| \leq 4\epsilon M + \frac{\kappa_1 ||R||_{\rho-2\delta}}{2\delta\epsilon^2} = O(\epsilon), \quad \forall t \in [0, \delta\epsilon^{-2}). \tag{5.4}$$

Remark 5.1 The Hamiltonians of the finite dimensional Galerkin systems are positive definite near the origin, and therefore, for sufficiently small initial conditions, solutions exist for all time. This fact is not used in the theorems. Rather, both theorems only require local existence of solutions, and the conclusion of the theorems can be considered as a theorem of existence for time of $O(\epsilon^{-2})$.

Remark 5.2 In the case of axisymmetric solutions, the result of [8] implies that, if N_2 is definite as a quadratic from in the actions J_γ, the properties (i)-(iii) of section 4 guarantee stability of near-monochromatic solutions for exponentially long time.

The proofs of the two theorems are very similar. The machinery is quite standard, and has been developed in the context of Nekhoroshev-type results (see e.g. [18]), averaging, KAM theorems etc. One distinct feature of the present proof is that the canonical transformations are around the origin (not around an invariant torus) and some care is needed in dealing with the singularity of the change to action-angle variables at the hyperplanes $a_\gamma = 0$.

Proof of Theorem 5.1: The proof is broken into four parts. In lemma 5.1 we estimate the domain of the canonical transformation and the remainder of the quartic normal form. lemma 5.2 gives us the distance between the original and the transformed variables. In lemma 5.3 we find times for which the trajectories of the equations stay within the domain of the transformation. In lemma 5.4 we use the normal form equations to estimate the drift of the actions for initial conditions and times suggested by the previous lemmas.

Lemma 5.1 Let $\rho > 5\delta > 0$ and $\lambda_1, \lambda_2 < 1$ as in (5.1). Then properties (i)-(iii) of section 4 imply that there is a canonical transformation $\Phi^{(\epsilon s_1, \epsilon^2 s_2)} : \mathcal{A}(D_{\rho-2\delta}) \to \mathcal{A}(D_\rho)$ so that $h \circ \Phi^{(\epsilon s_1, \epsilon^2 s_2)} = h_0 + \epsilon^2 n_2 + r$, with n_2 the quartic part of the normal form Hamiltonian of proposition 4.2 (restricted to the space axisymmetric solutions V_A), and r the remainder i.e. the terms of $O(\epsilon^3)$. The remainder satisfies

$$||r||_{\rho-2\delta} \leq (\lambda_1)^3 \alpha, \tag{5.5}$$

with α given in the proof.

Proof: Let s_i, $i = 1, 2$ be the functions that generate the canonical transformations on V_A, and let $L_i f = Ad_{\epsilon^i s_i} f$, for any function f. With this notation we have

$$h \circ \Phi^{(\epsilon s_1)} = \exp Ad_{\epsilon s_1} h = h_0 + \epsilon^2 g_2 + r_1^0 + r_1^1,$$

with

$$g_2 = \frac{1}{2} L_1 h, \quad r_1^0 = \sum_{k=3}^{\infty} \epsilon^k \frac{L_1^k}{k!} h_0, \quad r_1^1 = \sum_{k=2}^{\infty} \epsilon^k \frac{L_1^k}{k!} \epsilon h_1.$$

Also,

$$h \circ \Phi^{(\epsilon s_1, \epsilon^2 s_2)} = \exp Ad_{\epsilon^2 s_2} \exp Ad_{\epsilon s_1} h = h_0 + \epsilon^2 n_2 + r,$$

with

$$n_2 = g_2 + L_2 h_0, \quad r = r_1^0 + r_1^1 + r_2^0 + r_2^1,$$

$$r_2^0 = \epsilon^2 L_2(\epsilon^2 g_2 + r_1^0 + r_1^1), \quad r_2^1 = \sum_{k=3}^{\infty} \epsilon^{2k} \frac{L_2^k}{k!} \exp Ad_{\epsilon s_1} h.$$

We use lemmas 2 and 3 of the Appendix and $||s_1||_\rho \leq c_1 ||h_1||_\rho$, $||s_2||_\rho \leq c_2 ||h_2||_\rho$ to bound the different terms of r in $\mathcal{A}(D_{\rho-2\delta})$. For example, we have

$$||g_2||_{\rho-2\delta} \leq \frac{\tilde{c}c_1}{2\delta^2} ||h_1||_\rho^2,$$

$$||r_1^0||_{\rho-\delta} < \frac{(\lambda_1)^3}{1-\lambda_1} ||h_0||_\rho, \quad \text{and} \quad ||r_1^1||_{\rho-\delta} < \frac{(\lambda_1)^3}{1-\lambda_1} \frac{\delta^2}{\tilde{c}c_1}, \quad \text{both for} \quad \lambda_1 < 1.$$

The condition $\lambda_1 < 1$ guarantees the convergence of the first canonical transformation. Similarly we estimate n_2, while for r_2^0 we use

$$||L_2 g_2||_{\rho-2\delta} \leq \frac{2\tilde{c}}{\delta^2} ||s_2||_{\rho-\delta} ||g_2||_{\rho-\delta} \leq \frac{\tilde{c}c_2}{\delta^2} ||L_1 h_1||_{\rho-\delta}^2 \leq$$

$$\leq \frac{\tilde{c}c_2}{\delta^2} \left(\frac{\epsilon\tilde{c}}{\delta^2} ||s_1||_{\rho-\frac{\delta}{2}} ||h_1||_\rho \right)^2 \leq \frac{\tilde{c}c_2}{\delta^2} \left(\frac{2\tilde{c}c_1}{\delta^2} ||h_1||_\rho^2 \right)^2,$$

and likewise for the rest of r_2^0, using the bounds for r_1^0, r_1^1. Collecting the calculations we have

$$||r_2^0||_{\rho-2\delta} \leq (\lambda_1)^4 \frac{4c_2\delta^2}{\tilde{c}c_1^2 e^8} + (\lambda_1)^2 \frac{2c_2}{c_1 e^4} \left(\frac{(\lambda_1)^3}{1-\lambda_1} ||h_0||_\rho + \frac{(\lambda_1)^3}{1-\lambda_1} \frac{\delta^2}{\tilde{c}c_1} \right)$$

and also, using lemma 3 of the Appendix,

$$||r_2^1||_{\rho-2\delta} \leq \frac{(\lambda_2)^2}{(1-\lambda_1)(1-\lambda_2)} ||h||_\rho, \quad \text{for} \quad \lambda_2 < 1.$$

The condition $\lambda_2 < 1$ is needed for the convergence the second canonical transformation. Collecting all the terms of the remainder we obtain (5.5) with

$$\alpha = \left(10 + (\lambda_1)^2 \frac{c_2}{2c_1} \right) \left(||h||_\rho + \frac{1}{8Nc_1} \right) + \frac{100\lambda_1}{e^4} \left(\frac{c_2}{c_1} \right)^2 ||h||_\rho + \frac{\lambda_1 c_2}{2c_1^2 N e^2}.$$

■

Lemma 5.2 Let $\rho > 5\delta > 0$, $\lambda_1, \lambda_2 < 1$ as in lemma 5.1 and consider the restriction $\Phi^{(\epsilon s_1, \epsilon^2 s_2)} : \mathcal{A}(D_{\rho-4\delta}) \to \mathcal{A}(D_{\rho-2\delta})$. Then

(5.6) $$|x_i - x_i \circ \Phi^{(\epsilon S_1, \epsilon^2 S_2)}| < \epsilon\mu, \quad \forall i = 1, \dots 2N,$$

with μ defined below.

Proof: The first canonical transformation $\Phi^{(\epsilon s_1)}$ is the time-1 map of the Hamilton's equations for ϵs_1, i.e. $\frac{dq_i}{d\tau} = -\epsilon \frac{\partial s_1}{\partial p_i}$, $\frac{dp_i}{d\tau} = \epsilon \frac{\partial s_1}{\partial q_i}$. Suppose $x \in D_{\rho-4\delta}$ at $\tau = 0$. Then the smallest time τ_m for the trajectories of Hamilton's equations for ϵs_1 to be outside $D_{\rho-3\delta}$ is $\tau_m = \delta(\sup_{x\in D_{\rho-3\delta}, j=1,\dots 2N} |\frac{\partial S_1}{\partial x_j}|)^{-1}$. We easily see that $\lambda_1 < 1$ implies $1 < \tau_m$. Similarly, if we consider the time-1 map of Hamilton's equations for s_2 we have that if $\lambda_2 < 1$ and $x \in D_{\rho-3\delta}$ at $\tau = 0$ then $x \circ \Phi^{(\epsilon^2 s_2)} \in D_{\rho-2\delta}$. Thus, $\forall i = 1, \dots, 2N$ we have

$$|x_i - x_i \circ \Phi^{(\epsilon s_1, \epsilon^2 s_2)}| \leq \epsilon \max_{j=1,\dots,2N} \left|\left| \frac{\partial s_1}{\partial x_j} \right|\right|_{\rho-3\delta} + \epsilon^2 \max_{j=1,\dots,2N} \left|\left| \frac{\partial s_2}{\partial x_j} \right|\right|_{\rho-2\delta}$$

(5.7)
$$\leq \epsilon \left(\frac{c_1}{3\delta} ||h_1||_\rho + \frac{\epsilon c_2}{\delta} ||g_2||_{\rho-\delta} \right) = \epsilon \mu.$$

■

Lemma 5.3 Let ρ, δ be as in lemmas 5.1, 5.2, and define $\tilde{x}_i = x_i \circ \Phi^{(\epsilon s_1, \epsilon^2 s_2)}$, $i = 1, \dots, 2N$ and $\tilde{x} = [\tilde{x}_1, \dots, \tilde{x}_{2N}]$. Assume that $\tilde{x}(0) \in D_{\frac{\tilde{\rho}}{2}}$, with $\tilde{\rho} \leq \frac{1}{\sqrt{2}}(\rho - 5\delta)$. Then

$$\tilde{x}(t) \in D_{\frac{\tilde{\rho}+\delta}{\sqrt{2}}}, \quad \forall t \in [0, \frac{\delta^2}{2\kappa_1 ||r||_{\rho-2\delta}}).$$

Proof: Let $\tilde{q}_i = q_i \circ \Phi^{(\epsilon s_1, \epsilon^2 s_2)}$, $\tilde{p}_i = p_i \circ \Phi^{(\epsilon s_1, \epsilon^2 s_2)}$ and define $\tilde{J}_i = \tilde{q}_i^2 + \tilde{p}_i^2$, $\tilde{A}_i = \tilde{J}_i^{\frac{1}{2}}$, $i = 1, \dots N$. From $\tilde{q}(0) \in D_{\tilde{\rho}/2}$ we have $\tilde{A}_i(0) < \frac{\tilde{\rho}}{\sqrt{2}}$. Then, if $\tilde{J}_i \neq 0$ we have

$$\frac{d\tilde{A}_i}{dt} = 2\tilde{A}_i^{-1} \frac{d}{dt}\tilde{J}_i = 4\tilde{A}_i^{-1} \left(-\frac{\partial r}{\partial \tilde{q}_i}\tilde{p}_i + \frac{\partial r}{\partial \tilde{p}_i}\tilde{q}_i \right)$$

hence

(5.8)
$$\left| \frac{d\tilde{A}_i}{dt} \right| \leq \frac{1}{4} \left(\left| \frac{\partial R}{\partial \tilde{q}_i} \right| + \left| \frac{\partial R}{\partial \tilde{p}_i} \right| \right).$$

Note that since $\tilde{q}_i(t)$, $\tilde{p}_i(t)$ are analytic in time (5.8) also holds for the times for which $\tilde{J}_i = 0$. Now, let T_0 be such that $\tilde{A}_i(t) \leq \frac{\tilde{\rho}}{\sqrt{2}} + \frac{\delta}{2}$, $\forall t \in [0, T_0)$. Then for $t \in [0, T_0$ we have

$$\left| \frac{d\tilde{A}_i}{dt} \right| \leq \sup_{\tilde{A}_i < \frac{\sqrt{2}\tilde{\rho}+\delta}{2}} \left(\left| \frac{\partial r}{\partial \tilde{q}_i} \right| + \left| \frac{\partial r}{\partial \tilde{p}_i} \right| \right),$$

and using the assumption $\tilde{\rho} < \frac{1}{\sqrt{2}}(\rho - 5\delta)$ we have

$$\left|\frac{d\tilde{A}_i}{dt}\right| \leq \sup_{\tilde{x}\in D_{\rho-4\delta}} \left(\left|\frac{\partial r}{\partial \tilde{q}_i}\right| + \left|\frac{\partial r}{\partial \tilde{p}_i}\right|\right) \leq \frac{\kappa_1}{\delta}||r||_{\rho-2\delta}. \tag{5.9}$$

Thus $T_0 > \delta^2(2\kappa_1||R||_{\rho-2\delta})^{-1}$ and the conclusion follows.

■

Lemma 5.4 Let ρ_1, $\epsilon\mu$, $\tilde{\rho}$ as in lemmas 5.1, 5.3, and assume that

$$\epsilon\mu \leq \frac{2-\sqrt{2}}{2}(\rho - 5\delta), \quad \text{and} \quad \frac{\epsilon^2\delta}{2\kappa_1||r||_{\rho-2\delta}} \geq 1. \tag{5.10}$$

Then axisymmetric solutions of (2.9) with initial condition $x(0) \in D_{\rho_1}$ satisfy

$$|J_i(t) - J_i(0)| \leq 8\epsilon\mu + 4\kappa_1 \frac{||r||_{\rho-2\delta}}{\epsilon^2}, \quad \forall t \in [0, \delta\epsilon^{-2}), \quad \forall \quad i = 1, \ldots, N. \tag{5.11}$$

Proof: By the assumptions on ρ_1, $\epsilon\mu$, we have $x(0) \in D_{\rho_1} \subseteq D_{\frac{\tilde{\rho}}{2}-\epsilon\mu}$ and therefore, by lemma 5.2, lemma 5.3 applies and we have

$$\tilde{x}(t) \in D_{\tilde{\rho}+\frac{\delta}{\sqrt{2}}}, \quad \forall t \in [0, T_0), \quad \text{with} \quad T_0 = \frac{\delta^2}{2\kappa_1||r||_{\rho-2\delta}}.$$

By the assumption on $\epsilon\mu$ in (5.10) and lemma 5.2 we then have that $x(t) \in D_{\rho-4\delta}$, $\forall t \in [0, T_0)$. Thus for $x(0) \in D_{\rho_1}$ trajectories stay within the domain of the canonical transformation, for all times up to T_0. The second inequality of (5.10) implies that $\frac{\delta}{\epsilon^2} < T_0$, so that using (5.9), the proof of lemma 5.3, and the fact that $|q_i(t)|$, $|p_i(t)| < \rho \leq 1$, $\forall t \in [0, \frac{\delta}{\epsilon^2})$ and $i = 1, \ldots, N$, we have

$$\left|\frac{d\tilde{J}_i}{dt}\right| \leq \frac{\kappa_1}{\delta}||R||_{\rho-2\delta}, \quad \forall t \in [0, \delta\epsilon^{-2})$$

hence

$$|\tilde{J}_i(t) - \tilde{J}_i(0)| \leq \kappa_1 \frac{||R||_{\rho-2\delta}}{\epsilon^2}, \quad \forall t \in [0, \delta\epsilon^{-2}),$$

$i = 1, \ldots, N$. We also know that $|x_i(t) - q_i(t)| \leq \epsilon\mu$ and $|x_i(t) + q_i(t)| \leq 2$, $\forall t \in [0, \delta\epsilon^{-2})$, so that from

$$|J_i(t) - J_i(0)| \leq |\tilde{J}_i(t) - J_i(t)| + |\tilde{J}_i(0) - J_i(0)| + |\tilde{J}_i(t) - \tilde{J}_i(0)|$$

we easily obtain

$$|J_i(t) - J_i(0)| \leq 8\epsilon\mu + 4\kappa_1 \frac{||r||_{\rho-2\delta}}{\epsilon^2}, \quad \forall t \in [0, \delta\epsilon^{-2}), \quad \forall i = 1, \ldots, N.$$

$\blacksquare$

Proof of Theorem 5.2: Lemma 5.5 is analogous to lemmas 5.1, 5.2 combined. The proof is exactly the same, with capital letters replacing small letters. In lemma 5.6 we estimate times for which the trajectories stay in the domain of the canonical transformation, and in lemma 5.7 we estimate the drift of the actions.

Lemma 5.5 Let $\rho > 5\delta > 0$ and $\Lambda_1, \Lambda_2 < 1$ as in (5.3). Then properties (i)-(iii) of section 4 imply that there is a canonical transformation $\Phi^{(\epsilon S_1, \epsilon^2 S_2)} : \mathcal{A}(D_{\rho-2\delta}) \to \mathcal{A}(D_\rho)$ so that the transformed Hamiltonian $H \circ \Phi^{(\epsilon S_1, \epsilon^2 S_2)} = H_0 + \epsilon^2 N_2 + R$, with N_2 the quartic part of the normal form Hamiltonian of proposition 4.2 and R the remainder i.e. the terms of $O(\epsilon^3)$. The remainder satisfies

$$||R||_{\rho-2\delta} \leq (\Lambda_1)^3 \tilde{\alpha}, \tag{5.12}$$

with $\tilde{\alpha}$ the same as the α of theorem 5.1, but with capital letters. Moreover consider the restriction $\Phi^{(\epsilon S_1, \epsilon^2 S_2)} : \mathcal{A}(D_{\rho-4\delta}) \to \mathcal{A}(D_{\rho-2\delta})$. Then

$$|y_i - y_i \circ \Phi^{(\epsilon S_1, \epsilon^2 S_2)}| \leq \epsilon M, \quad \forall i = 1, \ldots 2N, \tag{5.13}$$

with M defined below.

Lemma 5.6 Let ρ, δ as in lemma 5.5 and let $\tilde{y}_i = y_i \circ \Phi^{(\epsilon S_1, \epsilon^2 S_2)}$, $i = 1, \ldots, 2N$ and $\tilde{y} = [\tilde{y}_1, \ldots, \tilde{y}_{2N}]$. Assume that $\tilde{y}(0) \in D_{\tilde{\rho}}$, with $\tilde{\rho} \leq (4l_{\max} + 2)^{-\frac{3}{2}}(\rho - 5\delta)$. Then

$$\tilde{y}(t) \in D_{(4l_{\max}+2)^{\frac{3}{2}}\tilde{\rho}+\delta}, \quad \forall t \in [0, \frac{\delta^2}{\sqrt{2}(2l_{\max} + 1)\kappa_1 ||R||_{\rho-2\delta}}).$$

Proof: Let $\tilde{B}_i = \tilde{I}_i^{\frac{1}{2}}$. Initially $\tilde{B}_i \leq 2(2l_{\max} + 1)^{\frac{1}{2}}$. Also, as in lemma 5.3, for $t \in [0, T_0)$ for which $\tilde{B}_i \leq 2(2l_{\max} + 1)^{\frac{1}{2}}\tilde{\rho} + 2^{-\frac{1}{2}}(2l_{\max} + 1)^{-1}\delta$ we have that

$$\left|\frac{d\tilde{B}_i}{dt}\right| \leq \frac{1}{4(2l_{\max} + 1)} \sup_{\tilde{B}_i \leq b_0} \sum_{m=-l}^{l} \left(\left|\frac{\partial R}{\partial \tilde{q}_{[l,m]}}\right| + \left|\frac{\partial R}{\partial \tilde{p}_{[l,m]}}\right| \right),$$

where $b_0 = 2(2l_{\max}+1)^{\frac{1}{2}}\tilde\rho + 2^{-\frac{1}{2}}(2l_{\max}+1)^{-1}\delta$. Hence

$$\left|\frac{d\tilde B_i}{dt}\right| \leq \frac{1}{4}\sup_{\tilde y\in D_{\rho-4\delta}}\left(\left|\frac{\partial R}{\partial \tilde q_{[l,m]}}\right| + \left|\frac{\partial R}{\partial \tilde p_{[l,m]}}\right|\right) \leq \frac{\kappa_1}{2\delta}||R||_{\rho-2\delta}.$$

Then T_0 is at least $\delta^2(\sqrt{2}(2l_{\max}+1)\kappa_1||R||_{\rho-2\delta})^{-1}$.

■

Lemma 5.7 Let ρ, δ, $\tilde\rho$ as in lemmas 5.5, 5.6, and ρ_2 as in theorem 5.2. Let $y(0)\in D_{\rho_2}$ and assume that

$$\epsilon M \leq 3^{-1}(4l_{\max}+2)^{-\frac{3}{2}}(\rho-5\delta), \qquad \frac{e^2\delta}{\sqrt{2}(2l_{\max}+1)\kappa_1||R||_{\rho-2\delta}} \leq 1. \tag{5.14}$$

Then we have that $\forall i = 1,\ldots,N$

$$|I_i(t) - I_i(0)| \leq 4\epsilon M + \frac{\kappa_1||R||_{\rho-2\delta}}{2\delta\epsilon^2}, \quad \forall t\in[0,\delta\epsilon^{-2}).$$

Proof: The hypotheses in (5.14) imply that $\tilde y(t)$ and $y(t)$ stay within $D_{\rho-4\delta}$ for $t\in[0,\delta\epsilon^{-2})$ and we can estimate the drift of $\tilde I_i = (2l+1)^{-1}\sum_{m=-l}^{l}(\tilde q^2_{[l,m]} + \tilde p^2_{[l,m]})$ using

$$\left|\frac{d\tilde I_i}{dt}\right| \leq \frac{\kappa_1}{2\delta}||R||_{\rho-2\delta}, \quad \forall t\in[0,T_0), \quad \text{with} \quad T_0 = \frac{\delta^2}{\sqrt{2}(2l_{\max}+1)\kappa_1||R||_{\rho-2\delta}}.$$

Therefore

$$|\tilde I_i(t) - \tilde I_i(0)| \leq \frac{\kappa_1||R||_{\rho-2\delta}}{2\delta\epsilon^2}, \quad \forall t\in[0,T_0), \quad \forall i - 1,\ldots,N.$$

Using the above and lemma 5.5 we easily obtain the conclusion.

■

6. Discussion

The arguments of section 4 imply that the stability of near-monochromatic solutions (in the asymptotic sense we are considering) should persist for larger Galerkin truncations of Hamilton's equations (2.9) with the Hamiltonian $H = H_0 + \epsilon H_1$, provided that property (iii) of section 4 is verified (numerically) for each truncation. Also, the addition of quartic terms $\epsilon^2 H_2$ or higher order terms to the Hamiltonian will not weaken the arguments of section 4. Similarly, the addition of more modes or higher order non-linearity will not have any qualitative effect on the rigorous normal forms, although the domain of the transformations will shrink. It would also be interesting to investigate near-monochromatic motions from a more geometrical point of view, by calculating Liapunov orbits numerically (in the axisymmetric case), and looking at their stability and bifurcations.

Our argument for the reliability of the numerics was based on the quantitative agreement of results obtained through different local error tolerances, i.e. on the self consistency of the integration method. It would be useful to use symplectic integrators, although we think that for the times we considered the results would be very close to what we obtained. A possible sign of discrepancy between symplectic and non-symplectic numerical integrators is the appearance of monotonic decreases or increases of the energy and actions in the non-symplectic integration (see [19] for an example). Such "numerical drifts" were observed, but were rare and of small size. Moreover we suppressed them by decreasing the error tolerance. Thus we think that, in the present problem, symplectic integration would come on it own in longer time integrations.

As we noted, the conclusion of the present work does not apply to the infinite dimensional case, since we can not exclude the possibility of resonant terms that destroy some or all of the adiabatic invariants of the finite dimensional truncations and produce parametric instabilities of the approximate orbits. For 2-D water waves it was shown in [6], [7] that the coefficients of analogous resonant terms vanish. We do not at present know whether such cancelations can occur on the sphere, e.g. for high frequency terms of axisymmetric solutions. Even in such a case it would be interesting to investigate the Diophantine equation of quartic resonance, maybe some relation between the distance from the resonance and the wavenumber l can be revealed.

7. Acknowledgments

I would like to thank Rafael de la Llave for valuable discussions and comments. The numerical part of this work was completed at the Mathematics Department of the University of Texas at Austin.

8. Appendix

We include three technical lemmas used in section 5. Since the lemmas are variants of well known facts the proofs are omitted. The notation is that of the paragraph defining analytic function and norms in section 5.

Lemma 1 (Cauchy estimate) Let $F \in \mathcal{A}(D_\rho)$ and $0 < \delta < \rho$. Then $\forall i = 1, \ldots, 2N$ we have

$$||\frac{\partial F}{\partial x_i}||_{\rho-\delta} \leq \frac{\kappa_1}{\delta}||F||_\rho, \quad \text{with} \quad \kappa_1 = e^{-1}\left(1 - \frac{\delta}{\rho}\right)^{-1}.$$

Lemma 2 (Poisson bracket estimate) Let $F \in \mathcal{A}(D_\rho)$, $G \in \mathcal{A}(D_{\rho-\delta})$ and $0 < \delta, \delta' < \rho$. Then

$$||[F, G]||_{\rho-\delta-\delta'} \leq \frac{\tilde{c}}{\delta(\delta + \delta')}||F||_{\rho-\delta}||G||_\rho, \quad \text{with} \quad \tilde{c} = 2N(\kappa_1)^2,$$

and κ_1 as in Lemma 1.

Lemma 3 (convergence of the Lie-series) Let $S, H \in \mathcal{A}(D_\rho)$, and define $G_0 = H$, $G_r = \frac{1}{r}[S, G_{r-1}]$ for $r \geq 1$. Then

$$||G_r||_{\rho-\delta} \leq B_r, \quad \text{with} \quad B_r = \left(\frac{\kappa_2}{\delta^2}||S||_\rho\right)^r ||H||_\rho, \quad \kappa_2 = e^2\tilde{c},$$

and $\tilde{c}$ as in Lemma 2. If $\frac{\kappa_2}{\delta^2}||S||_\rho < 1$ then $\exp(Ad_S H) \in \mathcal{A}(\mathcal{D}_{\rho-\delta})$.

9. REFERENCES

[1] R. de la Llave, P. Panayotaros: Water waves on the surface of the sphere, Jour. Nonlinear Sci. 6, 147-167 (1996)

[2] V.E Zakharov: Stability of periodic waves of finite amplitude on the surface of a deep fluid, Jour. Appl. Tech. Phys. 2, 190-194 (1968)

[3] J.W Miles: On Hamilton's principle for surface waves, Jour. Fluid Mech. 83, 153-158 (1977)

[4] W. Craig, M.D Groves: Hamiltonian long-wave approximation to the water wave problem, Wave Motion 19, 267-389 (1994)

[5] W. Craig, C. Sulem: Numerical simulation of gravity waves, Jour. Comp. Phys. 108, 73-83 (1993)

[6] A.I. Dyachenko, V.E. Zakharov: Is free surface hydrodynamics an integrable system ?, Phys. Let. A 190, 114-118 (1994)

[7] W. Craig, P. Worfolk: An integrable normal form for water waves in infinite depth, Physica D 84 513-531 (1995)

[8] F. Fassò, M. Guzzo, G. Benettin: Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems, preprint (1997)

[9] H. Lamb: Hydrodynamics, Dover, New York (1932)

[10] T.M MacRoberts: Spherical harmonics, Pergamon Press, London, (1967)

[11] A. C. Hindmarsh: Odepack, a systematized collection of ode solvers, in Scientific computing, ed. S. S. Stepleman et al., North-Holland, Amsterdam, 55-64 (1983)

[12] A. Iserles: A first course in the Numerical Analysis of differential equations, Cambridge Univ. Press, Cambridge (1996)

[13] W. Gear: Numerical initial value problems in ordinary differential equations, Prentice Hall, Englewood Cliffs, New Jersey (1971)

[14] L.F. Shampine, M. K. Gordon: Computer solution of ordinary differential equations: the initial value problem, Freeman, San Francisco (1975)

[15] P. Panayotaros: Numerical simulation of surface water waves on the sphere, preprint (1997)

[16] A.J Dragt, J.M Finn: Lie series and invariant functions for analytic symplectic maps, Jour. Math. Phys. 17, 2215-2227 (1976)

[17] H. Grauert, K. Fritzche: Several Complex Variables, Springer-Verlag, New York (1976)

[18] F. Fassò, G. Benettin: Composition of Lie transforms with rigorous estimates and applications to Hamiltonian perturbation theory, Jour. Appl. Math. Phys. (ZAMP), 140, 307-329 (1989)

[18] P.J. Channell, C. Scovel: Symplectic integration of Hamiltonian systems, Nonlinearity 3, 231-259 (1990)