Nekhoroshev stability of non-linear normal modes near an elliptic fixed point of a Hamiltonian system with symmetry

Panayotis Panayotaros*

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

Apdo. Postal 20-726

Depto. Matemáticas y Mecánica

01000 México D.F.

México

Abstract. We consider Hamiltonian systems of resonantly coupled harmonic oscillators with a physically motivated $U(1)$ symmetry, and identify a class of small amplitude periodic orbits (the non-linear normal modes) of approximating systems. We show that under some additional conditions on the quartic part of the Birkhoff normal form Hamiltonian that are in many cases weaker than integrability and convexity, the non-linear normal modes are Nekhoroshev stable.

* *e-mail: panos@uxmym1.iimas.unam.mx*

1. Introduction

We present sufficient conditions for the long time stability of quasi-monochromatic oscillations in an analytic Hamiltonian system of resonant weakly coupled harmonic oscillators with a $U(1)$ symmetry. By quasi-monochromatic oscillations we here denote motions in which the amplitude of all but one of the oscillators is very small. Using the resonance and symmetry properties of the Hamiltonian we will see that such motions are in the neighborhood of families of periodic orbits (the "non-linear normal modes") of Hamiltonian systems that approximate the original one. We will show that, under some additional conditions, trajectories of the original system that start near such periodic orbits stay nearby for a very long time. The frequencies of the oscillators are assumed be equal (the symmetry is best described in the next section), and the stability conditions involve the quartic part of the Birkhoff normal form Hamiltonian.

Although our results concern finite dimensional systems, the motivation and the symmetry condition we impose come from a problem of non-linear water waves in spherical geometry, where analogous motions were observed numerically to be stable (see [P]). The application of the present work to the water wave problem and other wave equations on the sphere is outlined in the last section, where we also point out some heuristic connections with other problems in non-linear wave propagation.

The stability of periodic and quasi-periodic orbits near an elliptic fixed point of a Hamiltonian system has been studied by several authors. Of particular relevance here are results of [FGB] and [N] on the Nekhoroshev stability of tori, and especially of [B] on the Nekhoroshev stability of finite dimensional tori in perturbed non-linear Schrödinger equations. Note that following the work of [L] we can study the stability of tori by looking at neighboring periodic orbits. In the above works the key properties for stability are integrability and convexity (possibly on an invariant set) of the quartic part of the Birkhoff normal form around the origin. In the problem we consider here, the assumed resonance and symmetry conditions lead formally to normal forms that posses near-monochromatic periodic orbits to all orders, but are too weak to guarantee integrability or convexity of the quartic normal form around the origin. Nevertheless, we can still use a strategy similar to the one used by [B] to show that, under two additional conditions on the quartic terms of the normal form, trajectories starting in a neighborhood of a near-monochromatic periodic orbit of a suitably defined approximating system will stay nearby for an exponentially long time. The first condition is that certain non-integrable quartic terms that we will call Benjamin-Feir terms are absent. The second condition is that certain integrable quartic terms be absent. This second condition is neither stronger nor weaker than convexity of the (integrable) quartic terms, while both conditions allow for the presence of non-integrable terms. Some slightly weaker versions of the above assumptions, as well as some special cases will also be pointed out.

The paper is organized as follows: In section 2 we describe the resonance and symmetry properties of the Hamiltonian, and use Birkhoff normal form arguments to show the existence of periodic orbits in approximating systems. This section serves mainly as a set-up and motivation for the stability result of the next chapter. In section 3 we state a stability theorem for near-monochromatic motions , and sketch the plan of the proof. The proof is given in the next three sections. In sections 4 and 5 we perform normal forms around the origin and periodic orbits respectively, and in section 6 we conclude the proof. In section 7 we discuss applications and some heuristic connections of the present work with Stokes waves in non-linear wave equations.

2. Resonance and symmetry properties of the Hamiltonian

We consider $\mathbf{R}^{2n}$ with the standard symplectic structure $\Omega = \sum_{\gamma} dq_{\gamma} \wedge dp_{\gamma}$. The coordinates q_{γ}, p_{γ} are labeled by the indices $\gamma \in \mathcal{J}$, where $\mathcal{J}$ is either the set of integers between $-\frac{n-1}{2}$ and $\frac{n-1}{2}$, for n odd; or the set of non-zero integers between $-\frac{n}{2}$ and $\frac{n}{2}$, for n even. For any smooth functions f, g, the Poisson bracket induced by Ω will be denoted by $[f, g]$.

Also, we consider Hamiltonians h that have an elliptic fixed point at the origin. For the purposes of this section it will suffice to consider Hamiltonians that are given as formal power series in the variables q_{γ}, p_{γ}. Thus, defining the variables a_{γ}, a_{γ}^* by $a_{\gamma} = q_{\gamma} + ip_{\gamma}$, $a_{\gamma}^* = q_{\gamma} - ip_{\gamma}$, we may assume that the Hamiltonians h have the form

$$h = h_0 + \sum_{j=1}^{\infty} \epsilon^j h_j, \quad \text{with}$$

$$(2.1) \quad h_0 = \sum_{\gamma \in \mathcal{J}} \omega_{\gamma} |a_{\gamma}|^2, \qquad h_j = \sum_{\kappa + \overline{\kappa} = j} \sum_{\gamma_1, \ldots, \gamma_{\kappa + \overline{\kappa}} \in \mathcal{J}} I^{(\kappa, \overline{\kappa})}_{\gamma_1, \ldots, \gamma_{\kappa + \overline{\kappa}}} a_{\gamma_1} \ldots a_{\gamma_{\kappa}} a^*_{\gamma_{\kappa+1}} \ldots a^*_{\gamma_{\kappa + \overline{\kappa}}},$$

where $\kappa, \overline{\kappa} \in \mathbf{Z}^+ \cup \{0\}$.

We are interested in Hamiltonians h given by (2.1) that also satisfy three additional properties:

Property (i) The frequencies ω_{γ} in (2.1) satisfy the resonance condition

$$(2.2) \qquad \omega_{\gamma} = \omega \neq 0, \quad \forall \gamma \in \mathcal{J}.$$

Property (ii) Defining the linear circle action $\varrho : S^1 \times \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ by $\varrho_{\phi} a_{\gamma} = e^{i\gamma\phi} a_{\gamma}$ and $\varrho_{\phi} a_{\gamma}^* = e^{-i\gamma\phi} a_{\gamma}^*$, the Hamiltonian h of (2.1) satisfies

$$(2.3) \qquad h(\varrho_{\phi} x) = h(x), \quad \forall \phi \in S^1, \quad \forall x \in \mathbf{R}^{2n}.$$

Note that requirement (2.3) is equivalent to

$$(2.4) \quad I^{(\kappa, \overline{\kappa})}_{\gamma_1, \ldots, \gamma_{\kappa + \overline{\kappa}}} \neq 0 \Rightarrow \sum_{\mu = \gamma_1}^{\gamma_{\kappa}} \mu - \sum_{\mu = \gamma_{\kappa+1}}^{\gamma_{\kappa + \overline{\kappa}}} \mu = 0, \quad \forall \kappa, \overline{\kappa} \in \mathbf{Z}^+ \cup \{0\}, \quad \forall \gamma_1, \ldots, \gamma_{\kappa + \overline{\kappa}} \in \mathcal{J}.$$

Before stating the third property we note some of the implications of properties (i) and (ii). To see these we bring the Hamiltonian h to a Birkhoff normal form by performing formal canonical transformations that eliminate non-resonant terms order by order. We use the normalization procedure described by Dragt and Finn [DF], in which the transformed Hamiltonian $h^{(r)}$ obtained after r steps is given by $h^{(r)} = \exp(Ad_{\epsilon^r \chi_r}) \ldots \exp(Ad_{\epsilon\chi_1})h = h_0 + N_r + R_r$, where N_r is in normal form and R_r is the remainder, i.e. the $O(\epsilon^{r+1})$ yet unnormalized terms of the transformed Hamiltonian. The functions χ_i are computed so as to eliminate the resonant terms of order ϵ^i at the i-th stage of the computation (see e.g. [FB] for an example of a detailed calculation using this method). Other ways of performing the normal form, e.g. the method of Deprit [D], lead to the same order r normal form part N_r. The calculations lead to the following formal result:

Proposition 2.1 Consider a Birkhoff normal form Hamiltonian $h^{(r)} = h_0 + N_r + R_r$ with $r \in \mathbf{Z}^+$ and $[N_r, h_0] = 0$, obtained as described above from a Hamiltonian h given by (2.1) and satisfying (i) and (ii). Also let $\overline{h}^{(r)} = h_0 + N_r$ and let Γ be any index in $\mathcal{J}$. Then the planes $M_\Gamma^{(r)}$ of points with $a_\gamma = 0$, $\forall \gamma \neq \Gamma$ are invariant under the Hamiltonian flow of $\overline{h}^{(r)}$.

Proof: The proof will be sketchy, as details can be filled-in easily using material from the next sections. We start with two observations. First, N_r can not have terms of odd order in ϵ, since by property (i) the resonance condition for odd terms is $\omega = 0$. Second, all monomials of $h^{(r)}$, and hence $h^{(r)}$ itself, satisfy the symmetry condition (ii). We can see this fact recursively. The orbits of the circle action ϱ coincide with the the orbits of the Hamiltonian flow of the function $L = \sum_{\gamma\in\mathcal{J}} \gamma |a_\gamma|^2$. The function L Poisson-commutes with all the monomials of the h_j of (2.1). Then $[\chi_1, L] = 0$ so that by the linearity of $Ad_{\epsilon\chi_1}$ and the Jacobi identity the monomials of $h^{(1)} = \exp(Ad_{\epsilon\chi_1})h$ will also Poisson-commute with L. The argument works for all the stages of the normalization, up to $h^{(r)}$.

To show the proposition, it is enough to show that the rate of change $\dot{a}_\gamma$, $\gamma \neq \Gamma$, does not contain monomials of the form $(a_\Gamma)^n (a_\Gamma^*)^m$ for any n, $m > 0$. Since N_r is in normal form, property (i) implies that such monomials must be of the form $(a_\Gamma)^m (a_\Gamma^*)^{m-1}$. Moreover, such terms can only arise from the monomials $(a_\Gamma)^m (a_\Gamma^*)^{m-1} a_\gamma^*$ of the Hamiltonian. However, the symmetry property (ii), and the fact that it also holds for the monomials of $h^{(r)}$ imply that the coefficients of the monomials $(a_\Gamma)^m (a_\Gamma^*)^{m-1} a_\gamma^*$ of $h^{(r)}$ vanish, unless $m\Gamma - (m-1)\Gamma - \gamma = 0$, i.e. unless $\gamma = \Gamma$.

■

The above formal argument works for any order r, however the remainder R_r of the Birkhoff normal form generally diverges as r increases and we can not conclude that the planes $M_\Gamma^{(r)}$ (i.e. as $r \to \infty$) are invariant under the Hamiltonian flow of h. On the other hand we can make the argument rigorous for analytic Hamiltonians by defining the consecutive canonical transformations in suitable neighborhoods of the origin, and we can also see that for given ϵ sufficiently small we can chose r so that R_r is minimized. Thus the calculation of $\overline{h}^{(r)}$ should be thought of as an asymptotic procedure for approximating periodic orbits of the Hamiltonian flow of h. Also, since the transformations are near-identity, the planes $M_\Gamma^{(r)}$ are near the planes $M_\Gamma^{(0)}$, i.e. the sets with $a_\gamma = 0$, $\forall \gamma \neq \Gamma$ for the original variables.

Remark 2.1 The question of whether the periodic orbits of the $\overline{h}^{(r)}$ approximate periodic orbits of h will not play any role in our analysis. We point out however work on a similar problem by [MRS], who use a variational method to show that periodic orbits of the normal form do approximate periodic orbits of the system.

Thus it is reasonable to study the behavior of trajectories starting near the $M_\Gamma^{(r)}$, and in the next section we give sufficient conditions trajectories of the Hamiltonian flow of h to stay near periodic orbits of the $\overline{h}^{(r)}$ on the $M_\Gamma^{(r)}$ for a long time. Note that the existence of two constants of motion, h and L, does not give any stability information in the particular problem and more detailed knowledge of h is required.

To state the stability conditions we examine more closely the second and third order normal form Hamiltonians $h^{(2)}$ and $h^{(3)}$. We apply the Dragt-Finn algorithm three times, first to eliminate non-resonant cubic terms, then to eliminate the non-resonant, and finally to eliminate non-resonant quintic terms. We arrive at the Hamiltonian $H = h^{(3)}$, which by (i) and (ii) has the general form

$$H = h_0 + \epsilon^2 \hat{h}_2 + f_2, \quad \text{with} \tag{2.5}$$

$$\hat{h}_2 = \sum_{\gamma_1,\gamma_2,\gamma_3,\gamma_4 \in \mathcal{J}} I_{\gamma_1,\gamma_2,\gamma_3,\gamma_4} a_{\gamma_1} a_{\gamma_2} a^*_{\gamma_3} a^*_{\gamma_4}, \tag{2.6}$$

and f_2 the $O(\epsilon^4)$ remainder. The coefficients $I_{\gamma_1,\gamma_2,\gamma_3,\gamma_4}$ can be computed from the coefficients of the cubic and quartic parts of h. Note that the Hamiltonian H also satisfies properties (i) and (ii).

In the next section we will discuss the stability of periodic orbits on the planes $M_\Gamma^{(3)}$ that are invariant under the Hamiltonian flow of H above. Note that our results can be easily modified to deal with periodic orbits of $h^{(2)}$ on the planes $M_\Gamma^{(2)}$. The only difference is that $h^{(3)}$ has a smaller remainder, and this leads to slightly weaker stability conditions that can also be checked without much additional effort.

Fixing some $\Gamma \in \mathcal{J}$ we now assume that the Hamiltonian H of (2.5) (and hence h) satisfies:

Property (iii) $\hat{h}_2$ contains a term $C_\Gamma |a_\Gamma|^4$ for some $C_\Gamma \in \mathbf{R}$, with $C_\Gamma \neq 0$.

Property (iii) is specific to a given Γ. Also note that by (i) and (ii) the monomial $|a_\Gamma|^4$ is the only monomial of the form $(a_\Gamma)^n (a_\Gamma^*)^m$ in $\hat{h}_2$.

We now decompose $\hat{h}_2$ as

$$\hat{h}_2 = \hat{h}_{2,\Gamma} + \hat{h}_{2,(\gamma,\Gamma)} + \hat{h}_{2,\gamma}, \tag{2.7}$$

where $\hat{h}_{2,\Gamma}$ contains all monomials with products of a_Γ and a_Γ^*, $\hat{h}_{2,\gamma}$ contains all monomials with products of a_γ and a_γ^* with $\gamma \neq \Gamma$, and $\hat{h}_{2,(\gamma,\Gamma)} = \hat{h}_2 - \hat{h}_{2,\Gamma} - \hat{h}_{2,\gamma}$.

In $\hat{h}_{2,(\gamma,\Gamma)}$ we identify two further types of terms. We let $\hat{h}_{BF}$ be the sum of monomials proportional to $a_{\gamma_1} a_{\gamma_2} a_\Gamma^* a_\Gamma^*$ with γ_1, $\gamma_2 \neq \Gamma$, and complex conjugates. Also, we let $\hat{h}_I$ be the sum of monomials proportional to $|a_\gamma|^2 |a_\Gamma|^2$ with $\gamma \neq \Gamma$. Finally, we let $h_N = \hat{h}_{2,(\gamma,\Gamma)} - \hat{h}_I - \hat{h}_{BF}$, and note that h_N is in general non-zero, containing terms $a_\Gamma a_{\gamma_1} a_{\gamma_2}^* a_{\gamma_3}^*$ with $\Gamma + \gamma_1 - \gamma_2 - \gamma_3 = 0$, and complex conjugates. The terms $\hat{h}_{BF}$ are the"Benjamin-Feir" terms (see the discussion in section 7 for this terminology).

3. A stability theorem

To state and prove the main theorem we assume that $H = h_0 + \epsilon^2 \hat{h}_2 + f_2$ as in (2.5), (2.6) has already been computed from a Hamiltonian h satisfying properties (i)-(iii) of the previous section. Moreover we assume that h is real analytic in a neighborhood of the origin, so that H is also real analytic in a smaller neighborhood.

We work in the following domains: consider the coordinates q_γ, p_γ, $\gamma \in \mathcal{J}$ as complex and define the domain $D_\rho \subset \mathbf{C}^{2n}$ to be the set of points satisfying $|q_\gamma| \leq \rho$, $|p_\gamma| \leq \rho$, $\forall \gamma \in \mathcal{J}$. Also, for functions $g : D_\rho \to \mathbf{C}$ we denote the supremum norm in D_ρ by $||g||_{D_\rho}$. By the assumption of real analyticity, the Hamiltonian H, defined as a power series in q_γ, p_γ in (2.1) and extended to $\mathbf{C}^{2n}$ by the complexification of q_γ, p_γ, is analytic in some domain $D_R \subset \mathbf{C}^{2n}$. A second type of domain we will need is defined as follows: consider $\mathbf{R}^{2n}$ with the coordinates q_γ, p_γ, $\gamma \in \mathcal{J}$. Fix $\nu > 0$ and some $\Gamma \in \mathcal{J}$, and let $J_\Gamma = |a_\Gamma|^2 - \nu$ (the dependence of J_Γ on ν will be suppressed from the notation). We complexify q_γ, p_γ, $\forall \gamma \neq \Gamma$ and the angle $\theta_\Gamma = \arg(a_\Gamma)$, and we let D_ρ^{ν,R_1} be the set of points of $\mathbf{C}^{2n}$ satisfying $|q_\gamma| \leq \rho$, $|p_\gamma| \leq \rho$, $\forall \gamma \neq \Gamma$, and $\frac{|J_\Gamma|}{R_1} + R_1|Im(\theta_\Gamma)| \leq \rho$. For functions $g : D_\rho^{\nu,R_1} \to \mathbf{C}$ we denote the supremum norm in D_ρ^{ν,R_1} by $||g||_{D_\rho^{\nu,R_1}}$. By the definition of D_ρ, D_ρ^{ν,R_1} the meaning of $\mathbf{R}^{2n} \cap D_\rho$, $\mathbf{R}^{2n} \cap D_\rho^\nu$ is clear.

The theorem below involves $O(1)$ constants α, β, C_{NBF}, $\tilde{C}_I$, $\tilde{C}_{BF}$, $\tilde{H}_1$ and $\tilde{F}_1$. Possible values of these constants are given immediately after the statement. Note that C_{NBF}, $\tilde{C}_I$, $\tilde{C}_{BF}$, $\tilde{H}_1$ and $\tilde{F}_1$ are specific to a particular Γ since they involve the $\hat{h}_I$, $\hat{h}_N$, $\hat{h}_{BF}$, $\hat{h}_{BF}$ of the previous section that are specific to a given Γ.

Theorem 3.1 Consider the real analytic Hamiltonian $H : \mathbf{R}^{2n} \cap D_R \to \mathbf{C}$, $R > 0$, satisfying properties (i)-(iii) of section 2. Consider a mode Γ with $|a_\Gamma|^2 = \nu'$, fix $\sigma \in (0, \frac{1}{2})$, and let $\epsilon > 0$ satisfy

$$\frac{8e^2\tilde{c}\pi}{\omega R^2}\epsilon^2(2\alpha + 10\epsilon^2\beta) < 1, \tag{3.1}$$

$$\frac{4e^{\frac{2}{5e^2}}}{e^2}\frac{\beta}{\alpha}R\epsilon^{2-\sigma} \leq \frac{\xi}{16}, \qquad \frac{8e^{\frac{2}{5e^2}}}{e^2}\frac{\beta}{\alpha}R^2\epsilon^{2-2\sigma} \leq \frac{\xi}{16}, \tag{3.2}$$

$$\frac{2e^{\frac{2}{5e^2}}}{e^2}\frac{\tilde{F}_1}{\tilde{H}_1}\epsilon^{1-2\sigma} + \frac{32\tilde{c}_1\pi}{C_\Gamma\sqrt{\nu'}}C_{NBF}\epsilon^\sigma \leq \frac{\xi}{16}, \quad \text{with} \quad \xi = \frac{4}{n}\left(\sqrt{1+\frac{n}{64}} - 1\right), \tag{3.3}$$

and

$$\epsilon^\sigma\frac{16e^2\tilde{c}_1}{C_\Gamma\nu'}(2\tilde{H}_1 + 10\epsilon^\sigma\tilde{F}_1) < 1. \tag{3.4}$$

Also, let $\rho_2 = \epsilon^\sigma$ and assume that

$$\nu' + \frac{3}{2}(\rho_2)^2 \leq \frac{R^2}{9}, \quad \text{and} \quad \nu' \geq \frac{3}{2}(\rho_2)^2. \tag{3.5}$$

Then if $x(t)$ is a trajectory of the Hamiltonian flow of H with initial condition $x(0) \in \mathbf{R}^{2n} \cap D_{\rho_0}^{\nu',\rho_2}$, where $0 < \rho_0 \leq \frac{\xi\rho_2}{32}$, we have that

$$x(t) \in \mathbf{R}^{2n} \cap D_{\frac{\rho_2}{2}}^{\nu',\rho_2}, \quad \forall t \in [-T_m, T_m], \tag{3.6}$$

where

$$T_m = \frac{1}{2^9 \tilde{c}_1 \epsilon^2}(M_2 + \tilde{M}_1 + \tilde{C}_{BF} + \epsilon^\sigma \tilde{C}_I)^{-1}, \tag{3.7}$$

and

$$M_2 = e^2 \epsilon^{2+4\sigma} \tilde{F}_1 \exp\left(-\sqrt{\frac{3\nu' C_\Gamma}{8e^2 \tilde{c}_1}}(2\tilde{H}_1 + 10\epsilon^\sigma \tilde{F}_1)^{-\frac{1}{2}} \epsilon^{-\frac{\sigma}{2}}\right), \tag{3.8}$$

$$\tilde{M}_1 = 3e^2(2n+2)\epsilon^4 \beta \exp\left(-\sqrt{\frac{2R^2\omega}{e^2 \pi \tilde{c}}}(2\alpha + 10\epsilon\beta)^{-\frac{1}{2}} \epsilon^{-1}\right). \tag{3.9}$$

Constants: The constants α, β are such that

$$||\epsilon^2 \hat{h}_2||_{D_R} \leq \epsilon^2 \alpha, \quad ||f_2||_{D_R} \leq \epsilon^4 \beta \tag{3.10}$$

and C_{NBF}, $\tilde{H}_1$, $\tilde{F}_1$, $\tilde{C}_I$ and $\tilde{C}_{BF}$ are defined by

$$C_{NBF} = 2^5 e^2 \sqrt{\frac{3}{2}} ||\hat{h}_N||_1, \tag{3.11}$$

$$\tilde{H}_1 = 2^5 \sqrt{\frac{3}{2}} \sqrt{\nu'} C_{NBF} + \epsilon^{2\sigma} C_\Gamma + 2^4 \epsilon^{2\sigma} ||\hat{h}_{2,\gamma}||_1, \tag{3.12}$$

$$\tilde{F}_1 = 2\epsilon^{2-4\sigma}\beta + 3\dot{2}^5 \sqrt{\frac{3}{2}} \frac{\pi \tilde{c}_1}{\nu' C_\Gamma} C_{NBF} (2^6 \sqrt{\frac{3}{2}} C_{NBF} + \tilde{H}_1), \tag{3.13}$$

$$\tilde{C}_I = 6\nu' \left(\frac{\tilde{F}_1}{\tilde{H}_1} + \epsilon^\sigma \frac{\pi \tilde{c}_1}{4C_\Gamma \nu'}(1 + \tilde{F}_1 \tilde{H}_1^{-1})\right) ||\hat{h}_I||_1 \tag{3.14}$$

$$\tilde{C}_{BF} = 18\nu' e^{2\epsilon^\sigma} \left(1 + \epsilon^\sigma \frac{\pi \tilde{c}_1}{4C_\Gamma \nu'} C_{NBF}\right) ||\hat{h}_{BF}||_1. \tag{3.15}$$

Here, for any quartic polynomial $P = \sum_{(j_1,j_2,j_3,j_4)\in\mathcal{P}} I_{j_1,j_2,j_3,j_4} a_{j_1} a_{j_2} a^*_{j_3} a^*_{j_4}$ with $\mathcal{P}$ a set of index quartets we let $||P||_1 = \sum_{(j_1,j_2,j_3,j_4)\in\mathcal{P}} |I_{j_1,j_2,j_3,j_4}|$.

Remark 3.1 In the case where the quartic terms h_I and h_{BF} both vanish we have from (3.14) and (3.15) that $\tilde{C}_I = \tilde{C}_{BF} = 0$. Then, (3.7) implies that the time of stability is exponentially long.

If either of h_I and h_{BF} does not vanish our statement is not optimal. If h_{BF} does not vanish, we can not in general obtain an exponentially long stability time. In the special case where the size of h_{BF} is comparable to ϵ^σ an exponentially long stability time can be recovered by slightly modifying the present proof. If h_{BF} vanishes and h_I does not vanish several improvements can be made in special cases. If $\hat{h}_2$ is integrable and convex we can invoke the more general theorems of [N], [FGB] to prove stability for exponentially long time. In the case where the terms of h_I can be absorbed in a quartic term that can be expressed as a function of h_0, stability for exponentially long time can be shown with a slight modification of the proof we give here. Note that convexity of the integrable part of $\hat{h}_2$ neither implies the condition $h_I \equiv 0$ nor is implied by it. Also, even if h_I and h_{BF} vanish, $\hat{h}_2$ can still contain non-integrable terms.

Remark 3.2 The theorem is local, in the sense that the assumptions are not uniform in ν'. On the other hand, here we are mainly interested in the case where ν' is of $O(1)$.

The plan of the proof is to estimate the rates of change of J_Γ and h_0, showing that they are slow, approximately $\sim T_m^{-1}$. To see this we perform a sequence of analytic canonical transformations to a new set of coordinates in which the slow evolution of J_Γ and h_0 becomes manifest. In the first sequence of transforms we apply the Birkhoff normal form algorithm around the origin κ times, with κ chosen so that the unnormalized remainder is minimized. Denoting the composition of these transforms by τ_1, the transformed Hamiltonian has the form $H \circ \tau_1 = h_0 + N_1 + R_1$, with N_1 in normal form (i.e. $[h_0, N_1] = 0$) and R_1 the remainder. Thus in the new coordinates, $\dot{h}_0 = [h_0, R_1]$, which will be small. The domain of the transform τ_1 will be $D_{\frac{R}{2}}$, and this step is described by Lemma 4.1.

Next, we consider normal forms around monochromatic periodic orbits of $\overline{H \circ \tau_1} = h_0 + N_1$. We fix a mode Γ and consider sets with $a_\Gamma a_\Gamma^* = \nu$ and $a_\gamma = 0$, $\forall \gamma \neq \Gamma$. Expressing the Hamiltonian $H \circ \tau_1$ in terms of the coordinates J_Γ, θ_Γ and a_γ, a_γ^* $\forall \gamma \neq \Gamma$ we may write

$$H \circ \tau_1 = h_\nu + (\hat{h}_2 - h_\nu) + (N_1 - \hat{h}_2) + R_1 + h_0,$$

where $h_\nu = 2C_\Gamma J_\Gamma$, i.e. h_ν is the linear (in J_Γ) part of the quartic self-interaction terms of the mode Γ. Our goal is to perform a second sequence of normal form transformations, in which the role of the unperturbed Hamiltonian is played by h_ν. First we consider the case where the quartic terms h_{BF} and h_I vanish. Then, the "non-linear part" (i.e. the part to be averaged) of $H \circ \tau_1$, specifically $H \circ \tau_1 - h_\nu - h_0$, is of size $O(\epsilon^{2+3\sigma})$ while h_ν is of size $O(\epsilon^{2+2\sigma})$, and we can perform the Birkhoff normal form algorithm obtaining an exponentially small remainder. An important feature of the argument is that $h_0 = \sum_{\gamma \neq \Gamma} \omega |a_\gamma|^2 + \omega J_\Gamma$ is invariant under this second sequence of transformations. We therefore have that both h_ν (and hence J_Γ), and h_0 evolve at an exponentially small rate.

When the quartic terms h_{BF} and h_I are present the time of stability is considerably shorter. This is because $\epsilon^2 h_{BF}$ and $\epsilon^2 h_I$ are in general comparable to h_ν and we can not obtain an exponentially small remainder by averaging them. Thus we subtract them from $H \circ \tau_1 - h_\nu - h_0$, add them at the end of $H \circ \tau_1$, and simply apply the transformations on them. Their size is almost unchanged by the transformations and since they don't commute with h_ν they give the dominant contribution to the rate of change of J_ν.

The normal form around the periodic orbits is described by Lemmas 5.1 and 5.2, and the result is used to estimate the drift of J_Γ and the a_γ, for $\gamma \neq \Gamma$ in Lemma 6.1. The results are translated to the original variables in the proof of the theorem, in section 6.

4. Normal form around the origin

The first step of the proof is a Birkhoff normal form around the origin, described by the next lemma.

Lemma 4.1 Consider the real analytic Hamiltonian $H = h_0 + \epsilon^2 \hat{h}_2 + f_2$ defined in the previous sections, and its analytic extension $H : D_R \to \mathbf{C}$. Let A_0, B_0 be such that

$$||\epsilon^2 \hat{h}_2||_{D_R} \leq A_0, \quad ||f_2||_{D_R} \leq B_0 \tag{4.1}$$

and assume that for some $\Delta \in (0, R)$ we have

$$\frac{e^2 \tilde{C} T}{\Delta^2}(2A_0 + 10B_0) < 1, \tag{4.2}$$

where $T = \frac{2\pi}{\omega}$. Then there exists an analytic canonical transformation $\tau_1 : D_{R-\Delta} \to D_R$ for which we have

$$H \circ \tau_1 = h_0 + N_1 + R_1, \tag{4.3}$$

where N_1 is in normal form, i.e. $[N_1, h_0] = 0$, and R_1 is the remainder, satisfying

$$||R_1||_{D_{R-\Delta}} \leq e^2 B_0 \exp\left(- \frac{2}{(e^2 \frac{\tilde{c}T}{\Delta^2})^{\frac{1}{2}}} (2A_0 + 10B_0)^{-\frac{1}{2}} \right). \tag{4.4}$$

Moreover, for any analytic function $g : D_R \to \mathbf{C}$ we have

$$||g \circ \tau_1||_{D_{R-\Delta}} \leq ||g||_{D_R} \exp\left(\frac{4B_0}{e^2}(2A_0 + 10B_0)^{-1}\right) \tag{4.5}$$

and

$$||g \circ \tau_1 - g||_{D_{R-\Delta}} \leq ||g||_{D_R} \frac{4B_0}{e^2(2A_0 + 10B_0)} \exp\left(\frac{4B_0}{e^2}(2A_0 + 10B_0)^{-1}\right). \tag{4.6}$$

Proof of Lemma 4.1: The transformation τ_1 will be the composition of κ canonical transformations $\tau_{(j)}$, $j = 1, 2, \ldots, \kappa$, with κ determined in the course of the proof.

We start by formally constructing normalizing canonical transformations $\tau_{(n+1)}$ and Hamiltonians $H^{(n+1)} = H^{(n)} \circ \tau_{(n+1)}$ recursively for $n \in \mathbf{Z}$, $n \geq 0$. The operations we describe below can be thought of as operations on formal powers series, and we want that at each step the Hamiltonian have the form

$$H^{(n)} = h_0 + \hat{h}_{(n)} + f_{(n)}. \tag{4.7}$$

Setting $\hat{h}_{(0)} = \epsilon^2 \hat{h}_2$, $f_{(0)} = f_2$, and $H^{(0)} = H$, we want to define $\hat{h}_{(n+1)}$, $f_{(n+1)}$ and $H^{(n+1)}$ from $\hat{h}_{(n)}$, $f_{(n)}$ and $H^{(n)}$.

To do this we define the time-average of functions $g : M \to \mathbf{R}$ along the Hamiltonian flow Φ_0^t of h_0 by

$$< g >_0 = \frac{1}{T} \int_0^T g(\Phi_0^t(x)) dt.$$

Also we want functions $\chi_{(n+1)}$ satisfying

$$[\chi_{(n+1)}, h_0] + f_{(n)} = < f_{(n)} >_0, \tag{4.8}$$

$\forall n \in \mathbf{Z}, n > 0$, and we check that the solution of (4.8) is

$$\chi_{(n+1)} = \frac{1}{T} \int_0^T t \left(f_{(n)}(\Phi_0^t(x)) - < f_{(n)} >_0 (\Phi_0^t(x)) \right) dt. \tag{4.9}$$

With these ingredients we define

$$H^{(n+1)} = H^{(n)} \circ \tau_{(n+1)} = \exp Ad_{\chi_{(n+1)}} H^{(n+1)}, \tag{4.10}$$

where $Ad_f g = [f, g]$ for f, g smooth functions on $\mathbf{R}^{2n}$, and $\exp Ad_f g = g + \sum_{j=1}^{\infty} (Ad_f)^j g$. The transformations $\tau_{(n+1)}$ thus formally defined are canonical. ¿From the definition of $H^{(n+1)}$ and $\exp Ad_f g$ we have

$$H^{(n+1)} = h_0 + \hat{h}_{(n)} + f_{(n)} + [\chi_{(n+1)}, h_0] + \tag{4.11}$$

$$+ \Big(h_0 \circ \tau_{(n+1)} - h_0 - [\chi_{(n+1)}, h_0] \Big) + \Big(\hat{h}_{(n)} \circ \tau_{(n+1)} - \hat{h}_{(n)} \Big) + \Big(f_{(n)} \circ \tau_{(n+1)} - f_{(n)} \Big),$$

thus, defining $\hat{h}_{(n+1)}$ and $f_{(n+1)}$ by

$$\hat{h}_{(n+1)} = \hat{h}_{(n)} + < f_{(n)} >_0 \tag{4.12}$$

and

$$f_{(n+1)} = \Big(h_0 \circ \tau_{(n+1)} - h_0 - [\chi_{(n+1)}, h_0] \Big) + \tag{4.13}$$

$$\Big(\hat{h}_{(n)} \circ \tau_{(n+1)} - \hat{h}_{(n)} \Big) + \Big(f_{(n)} \circ \tau_{(n+1)} - f_{(n)} \Big),$$

the Hamiltonian $H^{(n+1)}$ also has the desired decomposition of (4.7).

The fact that $\hat{h}_{(0)}$ is in normal form implies that the $\hat{h}_{(n)}$, $n > 0$ are also in normal from. For suppose that $[h_0, \hat{h}_{(n)}] = 0$. The average $< g >_0$ of any smooth function g Poisson-commutes with h_0 and therefore, by the definition of $\hat{h}_{(n+1)}$ in (4.12), we also have $[h_0, \hat{h}_{(n+1)}] = 0$.

Thus we have constructed formal Birkhoff normal form Hamiltonians $H^{(n)}$ and we now proceed to define the transforms $\tau_{(n)}$ for analytic functions in subsets of D_R. The goal is to have well defined analytic canonical transformations $\tau_{(n+1)} : D_{R-(n+1)\delta} \to D_{R-n\delta}$ for $n+1 \leq \kappa$, with appropriate κ and δ. To achieve this we need to estimate $\chi_{(n+1)}$ and hence $\hat{h}_{(n+1)}$ and $f_{(n+1)}$ recursively. From now on we will consider extension of these functions to $\mathbf{C}^{2n}$.

Let A_n, B_n satisfy

$$||\hat{h}_{(n)}||_{D_{R-n\delta}} \leq A_n, \quad ||f_{(n)}||_{D_{R-n\delta}} \leq B_n, \quad \forall n > 0.$$

For A_0, B_0 we have (4.1). From the definition of $\hat{h}_{(n)}$, $< f_{(n)} >_0$ and $\chi_{(n+1)}$ we have that

$$||f_{(n)}||_{D_{R-(n+1)\delta}} \leq B_n, \quad \text{and} \tag{4.14}$$

$$\text{(4.15)} \qquad ||\chi_{(n+1)}||_{D_{R-(n+1)\delta}} \leq 2TB_n.$$

In (4.15) we have used the fact that $\Phi_0^t(x) \in D_\rho$ for $x \in D_\rho$. ¿From the definition of $\hat{h}_{(n+1)}$ we also have

$$||\hat{h}_{(n+1)}||_{D_{R-(n+1)\delta}} \leq A_n + B_n$$

and we can therefore choose the A_n, B_n so that

$$\text{(4.16)} \qquad A_{n+1} \leq A_n + B_n, \quad \forall n \geq 0.$$

To estimate $f_{(n+1)}$ we use (4.13), Lemma A1 and the bound on $\chi_{(n+1)}$ in (4.15). For the second and third parentheses in (4.12) we have

$$\text{(4.17)} \qquad ||f_{(n)} \circ \tau_{(n+1)} - f_{(n)}||_{D_{R-(n+1)\delta}} \leq \int_0^1 \left|\left|[\chi_{(n+1)}, f_{(n)}]\right|\right|_{D_{R-(n+1)\delta}} \leq 2\frac{\tilde{c}}{\delta^2}TB_n^2,$$

and

$$\text{(4.18)} \qquad ||\hat{h}_{(n)} \circ \tau_{(n+1)} - \hat{h}_{(n)}||_{D_{R-(n+1)\delta}} \leq 2\frac{\tilde{c}}{\delta^2}TA_nB_n.$$

To estimate the first parenthesis in (4.13) we note that, using the definition of $\chi_{(n+1)}$ and letting $g \circ \tau_{n+1}^s = \exp sAd_{\chi_{(n+1)}}g$, we have

$$l = h_0 \circ \tau_{(n+1)} - h_0 - [\chi_{(n+1)}, h_0] = h_0 \circ \tau_{(n+1)} - h_0 + f_{(n)} - < f_{(n)} >_0$$

$$= \int_0^1 \Big([\chi_{(n+1)}, h_0](\tau_{(n+1)}^s(x)) + (f_{(n)} - < f_{(n)} >_0)(x)\Big)\, ds =$$

$$= -\int_0^1 \Big((f_{(n)} - < f_{(n)} >_0)(\tau_{(n+1)}^s(x)) - (f_{(n)} - < f_{(n)} >_0)(x)\Big)\, ds,$$

hence

$$\text{(4.19)} \quad ||\hat{h}_{(n)} \circ \tau_{(n+1)} - \hat{h}_{(n)}||_{D_{R-(n+1)\delta}} \leq \left|\left|[\chi_{(n+1)}, f_{(n)} - < f_{(n)} >_0]\right|\right|_{D_{R-(n+1)\delta}} \leq 4\frac{\tilde{c}}{\delta^2}TB_n^2$$

by Lemma A1. Then, from (4.13) and by using (4.17), (4.18), (4.19), and (4.16) for A_n we obtain

$$\text{(4.20)} \qquad ||f_{(n+1)}||_{D_{R-(n+1)\delta}} \leq \frac{\tilde{c}T}{\delta^2}(2A_n + 6B_n) \leq \frac{\tilde{c}T}{\delta^2}\left(2A_0 + 6B_n + 2\sum_{j=0}^{n-1} B_j\right) B_n.$$

To simplify (4.20), we let

$$\kappa = \left[\frac{1}{\lambda}\right], \quad \text{with} \quad \lambda^2 = e^2 \frac{\tilde{c}T}{\Delta^2}(2A_0 + 10B_0) \tag{4.21}$$

for some $\Delta \in (0, R)$, and choose $\delta = \frac{\Delta}{\kappa}$. ¿From (4.2) we have that $\kappa \geq 1$, and we can use this δ to perform at least κ transformations. Also let

$$E = \frac{\tilde{c}T}{\delta^2}(2A_0 + 10B_0) < \frac{1}{2}. \tag{4.22}$$

The inequality is satisfied by the hypothesis (4.2). Also, from (4.20) we have that $B_1 \leq EB_0$. Suppose now that for some $n < \kappa$ we have $B_j \leq E^j B_0$, for all $j = 1, 2, \ldots, n$. Then, from (4.20) we obtain

$$B_{n+1} \leq \frac{\tilde{c}T}{\delta^2}\left(2A_0 + 6E^n B_0 + 2B_0 \sum_{j=0}^{n-1} E^j\right) B_n \leq \frac{\tilde{c}T}{\delta^2}(2A_0 + 10B_0)E^n B_0, \tag{4.23}$$

and therefore by (4.22) we have $B_{n+1} \leq E^{n+1} B_0$.

With this information we see that the transformations $\tau_{(n+1)} : D_{R-(n+1)\delta} \to D_{R-n\delta}$, with $n+1 \leq \kappa$ are well defined, since from (4.15), and using the δ and κ of (4.21) we have

$$\frac{e^2\tilde{c}}{\delta^2}||\chi_{(n+1)}||_{D_{R-(n+1)\delta}} \leq \frac{2e^2\tilde{c}T}{\delta^2}E^n B_0 \leq \frac{2B_0}{2A_0 + 10B_0} < 1,$$

which, by Lemma A3 is the condition for the power series defining the transformations $\tau_{(n)}$ to converge.

Performing the normal form κ times, we define the remainder R_1 by $R_1 = f_{(\kappa)}$. Then combining $B_\kappa \leq E^\kappa B_0$ with (4.21) and (4.22) we obtain

$$B_\kappa \leq \left(\frac{e^2\tilde{c}\tilde{T}}{\Delta^2}(2A_0 + 10B_0)\frac{\kappa^2}{e^2}\right)^\kappa B_0,$$

or by (4.21)

$$B_\kappa \leq \left(\frac{\lambda\kappa}{e}\right)^{2\kappa} B_0,$$

and therefore

$$||R_1||_{D_{R-\Delta}} \leq B_\kappa \leq e^2 e^{-\frac{2}{\lambda}} B_0,$$

which by the definition of λ gives us conclusion (4.4).

To show conclusion (4.5) we use

$$g \circ \tau_{(n+1)}(x) = g \circ \tau_{(n)}(x) + \int_0^1 [\chi_{(n+1)}, g \circ \tau_{(n)}](x(s))ds, \tag{4.24}$$

which implies

$$||g \circ \tau_{(n+1)}||_{D_{R-(n+1)\delta}} \leq ||g \circ \tau_{(n)}||_{D_{R-(n+1)\delta}} \left(1 + \frac{\tilde{c}}{\delta^2}||\chi_{(n+1)}||_{D_{R-(n)\delta}}\right) \tag{4.25}$$

for $n+1 \leq \kappa$. Then (4.15), and $B_n \leq E^n B_0$, for $n \leq \kappa$, with $E < \frac{1}{2}$ imply

$$||g \circ \tau_{(n+1)}||_{D_{R-(n+1)\delta}} \leq ||g||_{D_R} \prod_{j=0}^{n} \left(1 + \frac{2\tilde{c}}{\delta^2} T E^j B_0\right) \leq ||g||_{D_R} \exp\left(\frac{4\tilde{c}T}{\delta^2} B_0\right). \tag{4.26}$$

¿From $\delta = \frac{\Delta}{\kappa}$ and κ as in (4.21), we see that (4.26) implies conclusion (4.5) in the statement.

Finally, to show conclusion (4.6) we use

$$\begin{aligned} ||g \circ \tau_{(\kappa)} - g||_{D_{R-\Delta}} \leq{}& ||g \circ \tau_{(\kappa)} - g \circ \tau_{(\kappa-1)}||_{D_{R-\Delta}} + ||g \circ \tau_{(\kappa-1)} - g \circ \tau_{(\kappa-2)}||_{D_{R-(\kappa-1)\delta}} \\ & + \ldots + ||g \circ \tau_{(1)} - g||_{D_{R-\delta}}, \end{aligned} \tag{4.27}$$

with $\delta = \frac{\Delta}{\kappa}$, so that by (4.24) and (4.14) we have

$$||g \circ \tau_{(\kappa)} - g||_{D_{R-\Delta}} \leq \frac{2\tilde{c}T}{\Delta^2}\kappa^2 (B_{\kappa-1}||g \circ \tau_{(\kappa-1)}||_{D_{R-(\kappa-1)\delta}} + \ldots B_0||g||_{D_R}). \tag{4.28}$$

Combining(4.28) with (4.26) we therefore have

$$||g \circ \tau_{(\kappa)} - g||_{D_{R-\Delta}} \leq \frac{2\tilde{c}T}{\Delta^2}\kappa^2 \left(\sum_{j=0}^{\infty} E^j\right) B_0||g||_{D_R} \exp\left(\frac{4\tilde{c}T}{\delta^2} B_0\right), \tag{4.29}$$

and from $\delta = \frac{\Delta}{\kappa}$ and $E < \frac{1}{2}$, we see that (4.29) implies (4.6) in the statement.

■

Remark 4.1 Note that for given ϵ the choice to perform $n = \kappa$ transformations minimizes the remainder estimate, in the sense that if we consider E in (4.22) as a function of n, i.e. with $\delta = \frac{\Delta}{n}$, the minimum of E^n is achieved for $n = \frac{1}{\lambda}$.

5. Normal form around periodic orbits

We now consider the Hamiltonian $H \circ \tau_1 = h_0 + N_1 + R_1$ obtained in Lemma 4.1. We want to compute a new Birkhoff normal form, this time around sets of points with $|a_\Gamma|^2 = \nu$ and $a_\gamma = 0$, $\gamma \neq \Gamma$. (By Proposition 2.1 these sets correspond to periodic orbits of the Hamiltonian flow of $h_0 + N_1$.)

Consider $\nu > 0$ and define the action $J_\Gamma = |a_\Gamma|^2 - \nu$, and the angle $\theta_\Gamma = \arg(a_\Gamma)$. (The dependence of J_Γ on ν will be suppressed from the notation.) The Poisson bracket between two functions f, g is

$$(5.1) \qquad [f,g] = 2\left(\frac{\partial f}{\partial J_\Gamma}\frac{\partial g}{\partial \theta_\Gamma} - \frac{\partial f}{\partial \theta_\Gamma}\frac{\partial g}{\partial J_\Gamma}\right) + \sum_{\gamma\neq\Gamma}\left(\frac{\partial f}{\partial q_\gamma}\frac{\partial g}{\partial p_\gamma} - \frac{\partial g}{\partial q_\gamma}\frac{\partial f}{\partial p_\gamma}\right).$$

The Hamiltonian $H \circ \tau_1$ of the previous lemma may be written as

$$(5.2) \qquad H \circ \tau_1 = \epsilon^2 \hat{h}_2 + (N_1 - \epsilon^2 \hat{h}_2) + R_1 + h_0,$$

with $\epsilon^2 \hat{h}_2$ the quartic part of N_1 (it coincides with the quartic part of H so that the notation is consistent). Using the decomposition of $\hat{h}_2$ introduced in (2.7), and the variables J_Γ and θ_Γ, we may write $\hat{h}_{2,\Gamma}$ as

$$(5.3) \qquad \hat{h}_{2,\Gamma} = 2C_\Gamma \nu J_\Gamma + C_\Gamma J_\Gamma^2 + \text{const.},$$

and we may also write the Hamiltonian $H \circ \tau_1$ as

$$(5.4) \qquad H \circ \tau_1 = h_\nu + \tilde{h}_{(0)} + \epsilon^2 \hat{h}_N + \tilde{f}_{(0)} + R_1 + h_0 + \epsilon^2 \hat{h}_{BF} + \epsilon^2 \hat{h}_I,$$

with

$$(5.5) \qquad h_\nu = 2C_\Gamma \nu J_\Gamma, \quad \tilde{h}_{(0)} = \epsilon^2 C_\Gamma J_\Gamma^2 + \epsilon^2 \hat{h}_{2,\gamma},$$
$$\hat{h}_N = \hat{h}_{2,(\gamma,\Gamma)} - \hat{h}_{BF} - \hat{h}_I, \quad \tilde{f}_{(0)} = N_1 - \epsilon^2 \hat{h}_2.$$

We now perform a Birkhoff normal form in some domain D_ρ^{ν,R_1} (see section 3 for this notation). The role of the unperturbed Hamiltonian will be played by h_ν.

Lemma 5.1 Fix ν, ρ_2 satisfying $\rho_2^2 < \nu$ and $8(\nu^2 + 2\nu\rho_2 + 2\rho_2)^{\frac{1}{2}} < R^2$, and let D_ρ^ν denote D_ρ^{ν,ρ_2}. Consider the real analytic Hamiltonian $H \circ \tau_1$ defined in Lemma 4.1, and its complex extension $H \circ \tau_1 : D_{\rho_2}^\nu \to \mathbf{C}$ f. Let $\delta_2 \in (0, \rho_2)$ and suppose that

$$(5.6) \qquad \frac{2e^2 \tilde{c}_1 \tilde{T}}{\delta_2^2} ||\epsilon^2 \hat{h}_N||_{D_{\rho_2}^\nu} < 1,$$

with $\tilde{T} = \frac{\pi}{\epsilon^2 C_\Gamma \nu}$. Then there exists an analytic canonical transformation $\tau_2 : D^\nu_{\rho_2 - \delta_2} \to D^\nu_{\rho_2}$ satisfying

$$h_0 \circ \tau_2 = h_0 \tag{5.7}$$

and for which $H \circ \tau_1 \circ \tau_2$ takes the form

$$H \circ \tau_1 \circ \tau_2 = h_\nu + \tilde{h}_{(1)} + \tilde{f}_{(1)} + R_1^{(1)} + h_0 + \epsilon^2 \hat{h}_{BF} \circ \tau_2 + \epsilon^2 \hat{h}_I \circ \tau_2 \tag{5.8}$$

with $\tilde{h}_{(1)}$ in normal form, i.e. $[h_\nu, \tilde{h}_{(1)}] = 0$, and $\tilde{h}_{(1)}$, $\tilde{f}_{(1)}$ satisfying

$$[h_0, \tilde{h}_{(1)}] = [h_0, \tilde{f}_{(1)}] = 0, \tag{5.9}$$

and

$$||\tilde{h}_{(1)}||_{D^\nu_{\rho_2 - \delta_2}} \leq ||\tilde{h}_{(0)} + \hat{h}_N||_{D^\nu_{\rho_2}}, \tag{5.10}$$

$$\begin{aligned} ||\tilde{f}_{(1)}||_{D^\nu_{\rho_2 - \delta_2}} &\leq 2||\tilde{f}_{(0)}||_{D^\nu_{\rho_2}} + \\ &+ \frac{2\tilde{c}\tilde{T}}{\delta_2^2} ||\epsilon^2 \hat{h}_N||_{D^\nu_{\rho_2}} (3||\epsilon^2 \hat{h}_N||_{D^\nu_{\rho_2}} + ||\tilde{h}_{(0)}||_{D^\nu_{\rho_2}}). \end{aligned} \tag{5.11}$$

Moreover, for any analytic function $g : D^\nu_{\rho_2} \to \mathbf{C}$ we have

$$||g \circ \tau_2||_{D^\nu_{\rho_2 - \delta_2}} \leq \left(1 + \frac{2\tilde{c}\tilde{T}}{\delta_2^2} ||\epsilon^2 \hat{h}_N||_{D^\nu_{\rho_2}}\right) ||g||_{D^\nu_{\rho_2}} \tag{5.12}$$

$$||g \circ \tau_2 - g||_{D^\nu_{\rho_2 - \delta_2}} \leq \frac{2\tilde{c}_1\tilde{T}}{\delta_2^2} ||\epsilon^2 \hat{h}_N||_{D^\nu_{\rho_2}} ||g||_{D^\nu_{\rho_2}}. \tag{5.13}$$

Remark 5.1 The condition on ν and ρ_2 guarantees that $D^\nu_{\rho_2} \subset D_{\frac{R}{2}}$.

Proof of Lemma 5.1: Let Φ^t_ν be the time-t map of the Hamiltonian flow of h_ν, $\tilde{T}$ the period, and $< g >_\nu (x) = \frac{1}{\tilde{T}} \int_0^{\tilde{T}} g(\Phi^t_\nu(x)) dt$, for any function g. The canonical transformation τ_2 will be the time-1 map of the Hamiltonian flow of a function $\tilde{\chi}_{(1)}$ satisfying

$$\epsilon^2 \hat{h}_N + [\tilde{\chi}_{(1)}, \epsilon^2 \hat{h}_N] = < \epsilon^2 \hat{h}_N >_\nu . \tag{5.14}$$

The solution of (5.14) is given by an expression analogous to (4.9), and as in (4.15) we obtain

$$||\tilde{\chi}_{(1)}||_{D^\nu_{\rho_2 - \delta_2}} \leq 2\tilde{T} ||\epsilon^2 \hat{h}_N||_{D^\nu_{\rho_2}}. \tag{5.15}$$

¿From Lemma A3 and (5.15), condition (5.6) guarantees the convergence of the series for $\exp Ad_{\tilde{\chi}_{(1)}}$, and the transform τ_2 is well defined on $D^{\nu}_{\rho_2-\delta_2}$. Also, the estimate on $\tilde{\chi}_{(1)}$ in (5.15) gives us conclusions (5.12) and (5.13) of the statement.

The transformed Hamiltonian $H \circ \tau_1 \circ \tau_2$ has the form

$$H \circ \tau_1 \circ \tau_2 = h_\nu + \tilde{h}_{(0)} \hat{h}_N + \epsilon^2 + [\tilde{\chi}_{(1)}, \epsilon^2 \hat{h}_N] + \tilde{f}_{(0)} \circ \tau_2 + (h_\nu \circ \tau_2 - h_\nu - [\tilde{\chi}_{(1)}, h_\nu]) +$$

$$+ (\tilde{h}_{(0)} \circ \tau_2 - \tilde{h}_{(0)}) + (\epsilon^2 \hat{h}_N \circ \tau_2 - \epsilon^2 \hat{h}_N) + R_1 \circ \tau_2 + h_0 \circ \tau_2 + \epsilon^2 \hat{h}_{BF} \circ \tau_2 + \epsilon^2 \hat{h}_I \circ \tau_2,$$

and we set

$$\tilde{h}_{(1)} = \tilde{h}_{(0)} + < \epsilon^2 \hat{h}_N >_\nu, \quad R_1^{(1)} = R_1 \circ \tau_2, \tag{5.16}$$

and

$$\tilde{f}_{(1)} = \tilde{f}_{(0)} \circ \tau_2 + \Big(h_\nu \circ \tau_2 - h_\nu - [\tilde{\chi}_{(1)}, h_\nu] \Big) + \tag{5.17}$$

$$+ \Big(\tilde{h}_{(0)} \circ \tau_2 - \tilde{h}_{(0)} \Big) + \Big(\epsilon^2 \hat{h}_N \circ \tau_2 - \epsilon^2 \hat{h}_N \Big) .$$

Then (5.10) follows immediately from the definitions of the time-average and $\tilde{h}_{(1)}$. To obtain (5.11) we estimate the parentheses in (5.17) as in Lemma 4.1, and calculate

$$||\tilde{f}_{(1)}||_{D^\nu_{\rho_2-\delta_2}} \leq 2||\tilde{f}_{(0)}||_{D^\nu_{\rho_2}} + \frac{2\tilde{c}_1}{\delta_2^2} ||\tilde{\chi}_{(1)}||_{D^\nu_{\rho_2}} \Big(3||\epsilon^2 \hat{h}_N||_{D^\nu_{\rho_2}} + ||\tilde{h}_{(0)}||_{D^\nu_{\rho_2}} \Big) , \tag{5.18}$$

which together with (5.15) is what we want.

To see that $h_0 \circ \tau_2 = h_0$ we note that $[\hat{h}_N, h_0] = 0$ and $h_0(\Phi^t_\nu(x)) = h_0(x)$, and hence $[< \hat{h}_N >_\nu, h_0] = 0$. ¿From the definition of $\tilde{\chi}_{(1)}$ we then have that $[\tilde{\chi}_{(1)}, h_0] = 0$, and therefore $\exp Ad_{\tilde{\chi}_{(1)}} h_0 = h_0$. ¿From the definition of $\tilde{h}_{(1)}$ and the fact that $[h_0, \hat{h}_N] = 0$ we also see that $[h_0, \tilde{h}_{(1)}] = 0$. For $\tilde{f}_{(1)}$, we examine (5.16), and since $\tilde{f}_{(0)}$, $\tilde{\chi}_{(1)}$, h_ν and $\hat{h}_N$ Poisson-commute with h_0, and τ_2 is canonical we must also have $[h_0, \tilde{f}_{(1)}] = 0$.

■

In the next lemma the Birkhoff normal form around the set of points with $J_\Gamma = 0$ and $a_\gamma = 0$ for $\gamma \neq \Gamma$ is iterated. The goal is to minimize the remainder (unnormalized) part of the Hamiltonian, and the construction is very much along the lines of Lemma 4.1. An important feature of the new sequence of transforms is that they leave h_0 invariant.

Lemma 5.2 Consider the analytic Hamiltonian $H \circ \tau_1 \circ \tau_2 : D^{\nu}_{\rho_3} \to \mathbf{C}$ defined as in Lemmas 4.1, 5.1, with $\rho_3 \in (0, \rho_2 - \delta_2]$. Let $\tilde{A}_1$, $\tilde{B}_1$ be such that

$$||\tilde{h}_{(1)}||_{D^{\nu}_{\rho_3}} \leq \tilde{A}_1, \quad ||\tilde{f}_{(1)}||_{D^{\nu}_{\rho_3}} \leq \tilde{B}_1, \tag{5.19}$$

and assume that for some $\delta_3 \in (0, \rho_3)$ we have

$$\frac{e^2 \tilde{c}_1 \tilde{T}}{\delta_3^2}(2\tilde{A}_1 + 10\tilde{B}_1) < 1. \tag{5.20}$$

Then there exists an analytic canonical transformation $\tau_3 : D_{\rho_3 - \delta_3} \to D_{\rho_3}$, satisfying

$$h_0 \circ \tau_3 = h_0 \tag{5.21}$$

for which we have

$$H \circ \tau_1 \circ \tau_2 \circ \tau_3 = h_\nu + N_2 + R_2 + \tilde{R}_1 + h_0 + \epsilon^2 h_{BF} \circ \tau_2 \circ \tau_3 + \epsilon^2 h_I \circ \tau_2 \circ \tau_3 \tag{5.22}$$

with N_2 in normal form, i.e. $[N_2, h_\nu] = 0$. The N_2 and R_2 also satisfy

$$[N_2, h_0] = [R_2, h_0] = 0, \tag{5.23}$$

and we have

$$||R_2||_{D^{\nu}_{\rho_3 - \delta_3}} \leq e^2 \tilde{B}_1 \exp\left(- \frac{2}{(e^2 \frac{\tilde{c}_1 \tilde{T}}{\delta_3^2})^{\frac{1}{2}}} (2\tilde{A}_1 + 10\tilde{B}_1)^{-\frac{1}{2}} \right), \tag{5.24}$$

$$||\tilde{R}_1||_{D^{\nu}_{\rho_3 - \delta_3}} \leq ||R_1^{(1)}||_{D^{\nu}_{\rho_3}} \exp\left(\frac{4\tilde{B}_1}{e^2} (2\tilde{A}_1 + 10\tilde{B}_1)^{-1} \right). \tag{5.25}$$

Moreover, for any analytic function $g : D_{\rho_3} \to \mathbf{C}$ we have

$$||g \circ \tau_3||_{D^{\nu}_{\rho_3 - \delta_3}} \leq ||g||_{D^{\nu}_{\rho_3}} \exp\left(\frac{4\tilde{B}_1}{e^2(2\tilde{A}_1 + 10\tilde{B}_1)} \right), \tag{5.26}$$

and

$$||g \circ \tau_3 - g||_{D^{\nu}_{\rho_3 - \delta_3}} \leq ||g||_{D^{\nu}_{\rho_3}} \frac{4\tilde{B}_1}{e^2(2\tilde{A}_1 + 10\tilde{B}_1)} \exp\left(\frac{4\tilde{B}_1}{e^2(2\tilde{A}_1 + 10\tilde{B}_1)} \right). \tag{5.27}$$

Proof of Lemma 5.2: The method of proof is that of Lemma 4.1. Letting $H_2 = H \circ \tau_1 \circ \tau_2$, we will define a sequence of canonical transformations $\tilde{\tau}_{(n)}$, $n \geq 2$, and Hamiltonians $H^{(n+1)} = H_2 \circ \tilde{\tau}_{(n+1)}$, $n \geq 1$, and let τ_3 be the composition of $\tilde{\kappa}$ such transforms for appropriate $\tilde{\kappa}$. At the $n-$th iteration the Hamiltonian will have the decomposition

$$H_2^{(n)} = h_\nu + \tilde{h}_{(n)} + \tilde{f}_{(n)} + R_1^{(n)} + {h_0}^{(n)} + \epsilon^2 \hat{h}_{BF}^{(n)} + \epsilon^2 \hat{h}_I^{(n)}, \tag{5.28}$$

The $\tilde{h}_{(1)}$, $\tilde{f}_{(1)}$ and $R_1^{(1)}$ are the ones of Lemma 5.1 (i.e. the notation is consistent), and $h_0^{(1)} = h_0$, $H_2^{(1)} = H_2$, $\hat{h}_I^{(1)} = \hat{h}_I \circ \tau_2$, $\tilde{h}_{BF}^{(1)} = \tilde{h}_{BF} \circ \tau_2$. Also we define functions $\tilde{\chi}_{(n+1)}$ satisfying

$$[\tilde{\chi}_{(n+1)}, h_0] + \tilde{f}_{(n)} =< \tilde{f}_{(n)} >_\nu, \quad n \geq 1. \tag{5.29}$$

The $\tilde{\chi}_{(n+1)}$ are given by an expression analogous to (4.9). To describe the $\tilde{h}_{(n)}$, $\tilde{f}_{(n)}$ recursively, we let $H_2^{(n+1)} = H_2^{(n)} \circ \tilde{\tau}_{(n+1)} = \exp Ad_{\tilde{\chi}_{(n+1)}} H_1^{(n)}$. We have

$$\begin{aligned} H_2^{(n+1)} = h_\nu + \tilde{h}_{(n)} + \tilde{f}_{(n)} + [\tilde{\chi}_{(n+1)}, h_\nu] + (h_\nu \circ \tilde{\tau}_{(n+1)} - h_\nu - [\tilde{\chi}_{(n+1)}, h_\nu])+ \\ +(\tilde{h}_{(n)} \circ \tilde{\tau}_{(n+1)} - \tilde{h}_{(n)}) + (\tilde{f}_{(n)} \circ \tilde{\tau}_{(n+1)} - \tilde{f}_{(n)}) + R_1^{(n)} \circ \tilde{\tau}_{(n+1)} + \\ +h_0 \circ \tilde{\tau}_{(n+1)} + \epsilon^2 \hat{h}_{BF}^{(n+1)} + \epsilon^2 \hat{h}_I^{(n+1)}. \end{aligned} \tag{5.30}$$

Defining $\tilde{h}_{(n+1)}$, $R_1^{(n+1)}$, ${h_0}^{(n+1)}$ and $\tilde{f}_{(n+1)}$ by

$$\tilde{h}_{(n+1)} = \tilde{h}_{(n)} + < \tilde{f}_{(n)} >_\nu, \quad R_1^{(n+1)} = R_1^{(n)} \circ \tilde{\tau}_{(n+1)}, \quad {h_0}^{(n+1)} = {h_0}^{(n)} \circ \tilde{\tau}_{(n+1)}, \tag{5.31}$$

$$\hat{h}_{BF}^{(n+1)} = \hat{h}_{BF}^{(n)} \circ \tilde{\tau}_{(n+1)}, \quad \hat{h}_I^{(n+1)} = \hat{h}_I^{(n)} \circ \tilde{\tau}_{(n+1)}, \tag{5.32}$$

$$\tilde{f}_{(n+1)} = H_2^{(n+1)} - h_\nu - \tilde{h}_{(n+1)} - R_1^{(n+1)} - {h_0}^{(n+1)} \tag{5.33}$$

the Hamiltonian $H_2^{(n+1)}$ has the decomposition of (5.28).

The $\tilde{h}_{(n+1)}$, $\tilde{f}_{(n+1)}$ and $\tilde{\chi}_{(n+1)}$ are analogous to the respective $\hat{h}_{(n)}$, $\hat{f}_{(n)}$ and $\chi_{(n)}$ of Lemma 4.1, and we can similarly define bounds $\tilde{A}_n$, $\tilde{B}_n$ on the the supremum norm of the $\hat{h}_{(n)}$, $\hat{f}_{(n)}$ respectively in the domains $D^\nu_{\rho_3-(n-1)\tilde{\delta}}$, for $n \geq 1$. The iterative estimates for $\tilde{h}_{(n+1)}$, $\tilde{f}_{(n+1)}$ and $\tilde{\chi}_{(n+1)}$ readily follow from the arguments of Lemma 4.1: $\tilde{\delta}$, δ_3, $\tilde{\kappa}$ and $\tilde{\lambda}$ are analogous to δ, Δ, κ and λ respectively in Lemma 4.1. Thus $\tilde{A}_1$, $\tilde{B}_1$ are analogous to A_0, B_0 in Lemma 4.1, so that by assumption (5.20) we can apply the iteration $\tilde{\kappa} \geq 1$ times, where

$$\tilde{\kappa} = \left[\frac{1}{\tilde{\lambda}}\right], \quad \text{with} \quad \tilde{\lambda} = e^2 \frac{\tilde{c}_1 \tilde{T}}{\delta_3^2}(2\tilde{A}_1 + 10\tilde{B}_1). \tag{5.34}$$

Following the argument of Lemma 4.1, by letting $\tilde{\delta} = \frac{\delta_3}{\tilde{\kappa}}$ we have well defined analytic canonical transformations $\tilde{\tau}_{(n)} : D^{\nu}_{\rho_3-(n-1)\tilde{\delta}} \to D^{\nu}_{\rho_3-(n-2)\tilde{\delta}}$, for $n \geq 2$.

The iterative estimates for $\tilde{h}_{(n+1)}$, $\tilde{f}_{(n+1)}$ and $\tilde{\chi}_{(n+1)}$ readily follow from the arguments of Lemma 4.1: $\tilde{\delta}$, δ_3, $\tilde{\kappa}$ and $\tilde{\lambda}$ are analogous to δ, Δ, κ and λ respectively in Lemma 4.1.

Setting $N_2 = \tilde{h}_{\tilde{\kappa}}$ and $R_2 = \tilde{f}_{\tilde{\kappa}}$ and comparing with Lemma 4.1 we thus immediately obtain conclusions (5.24), (5.26) and (5.27), while conclusion (5.25) is (5.26) with g replaced by $R_1^{(1)}$.

Finally, from $[\tilde{f}_{(1)}, h_0] = 0$ shown in Lemma 5.1, the inductive step $[\tilde{f}_{(n)}, h_0] = 0$ for some $n > 1$, and the definition of the $\tilde{\chi}_{(n+1)}$, we have that $[\tilde{\chi}_{(n+1)}, h_0] = 0$. Therefore $h_0 \circ \tilde{\tau}_{(n+1)} = h_0$, $2 \leq n \leq \tilde{\kappa}$, and $h_0 \circ \tau_3 = h_0$. To show (5.23) we first note that $[\tilde{h}_{(1)}, h_0] = [\tilde{f}_{(1)}, h_0] = 0$ by Lemma 5.1. Supposing that $[\tilde{h}_{(n)}, h_0] = [\tilde{f}_{(n)}, h_0] = 0$ for some $n > 1$, we easily see from the definition of $\tilde{h}_{(n+1)}$ that it also Poisson-commutes with h_0. Also, examining the individual terms of $\tilde{f}_{(n+1)}$, $n > 1$, i.e. (5.30) and (5.32), we see that they Poisson-commute with h_0. Hence both $\tilde{h}_{(\tilde{\kappa})}$ and $\tilde{f}_{(\tilde{\kappa})}$ Poisson-commute with h_0.

■

6. Proof of the stability theorem

We now use the transformed Hamiltonian $H \circ \tau_1 \circ \tau_2 \circ \tau_3$ defined by the previous lemma to study the rate of change of J_Γ and h_0 in $\mathbf{R}^{2n} \cap D^{\nu}_{\rho_3-\delta_3}$.

Lemma 6.1 Consider the real analytic Hamiltonian $H \circ \tau_1 \circ \tau_2 \circ \tau_3 : \mathbf{R}^{2n} \cap D^{\nu}_{\rho_3-\delta_3} \to \mathbf{R}$ as in Lemma 5.1, and let ρ_4, δ_4, ρ_5 be positive reals satisfying $\rho_4 \leq \rho_3 - \delta_3$, $\delta_4 < \rho_4$, and $\rho_5 \leq \rho_4 - \delta_4$. Also suppose that $\rho_2 \leq 8\rho_5$, with ρ_2 as in Lemmas 4.1, 5.1. Then for initial conditions $x(0)$ such that

$$x(0) \in \mathbf{R}^{2n} \cap D^{\nu,\rho_2}_{\xi\rho_5}, \quad \text{with} \quad \xi = \frac{4}{n}\left(\sqrt{1 + \frac{n}{64}} - 1\right), \tag{6.1}$$

the trajectories $x(t)$ of the Hamiltonian flow of $H \circ \tau_1 \circ \tau_2 \circ \tau_3$ satisfy

$$x(t) \in \mathbf{R}^{2n} \cap D^{\nu,\rho_2}_{\rho_5} \tag{6.2}$$

for all times $t \in [-T_D, T_D]$, where

$$T_D \leq \frac{\delta_4^2}{2\tilde{c}_1} \left(\frac{\rho_5}{\rho_4} \right)^2 (m_2 + (2n+2)\tilde{m}_1 + m_I + m_{BF})^{-1}, \tag{6.3}$$

and m_2, $\tilde{m}_1$, m_I, m_{BF} are numbers satisfying

$$||R_2||_{D^\nu_{\rho_4}} \leq m_2, \quad ||\tilde{R}_1||_{D^\nu_{\rho_4}} \leq \tilde{m}_1, \tag{6.4}$$

$$||\epsilon^2 \hat{h}_{BF} \circ \tau_2 \circ \tau_3||_{D^\nu_{\rho_4}} \leq m_{BF}, \quad ||\epsilon^2 \hat{h}_I \circ \tau_2 \circ \tau_3 - \hat{h}_I||_{D^\nu_{\rho_4}} \leq m_I.$$

Remark 6.1 Condition $\rho_2 \leq 8\rho_5$ will be satisfied for an appropriate choice of ρ_j, δ_j, $j = 2, 3, 4$. This choice will be made in the proof of Theorem 3.1 below.

Proof of Lemma 6.1: By Lemma 5.1, the rate of change of h_ν and h_0 under the Hamiltonian flow of $H \circ \tau_1 \circ \tau_2 \circ \tau_3$ in $D^\nu_{\rho_3 - \delta_3}$ is given by

$$\frac{dh_\nu}{dt} = [h_\nu, R_2 + \tilde{R}_1 + \epsilon^2 \hat{h}_{BF} \circ \tilde{\tau} + \epsilon^2 (\hat{h}_I \circ \tilde{\tau} - \hat{h}_I)], \tag{6.5}$$

$$\frac{dh_0}{dt} = [h_0, \tilde{R}_1], \tag{6.6}$$

where $\hat{h}_{BF} \circ \tilde{\tau} = \hat{h}_{BF} \circ \tau_2 \circ \tau_3$ and $\hat{h}_I \circ \tilde{\tau} = \hat{h}_I \circ \tau_2 \circ \tau_3$.

Hence, using Lemma A2 in $D^\nu_{\rho_4 - \delta_4}$ with δ_4, ρ_4 satisfying $0 < \delta_4 < \rho_4 \leq \rho_3 - \delta_3$, we have that

$$\sup_{x \in D^\nu_{\rho_4 - \delta_4}} \left| \frac{dh_\nu}{dt} \right| \leq \frac{\tilde{c}_1}{\delta_4^2} ||h_\nu||_{D^\nu_{\rho_4}} \left(||R_2 + \tilde{R}_1||_{D^\nu_{\rho_4}} + \epsilon^2 ||\hat{h}_{BF} \circ \tilde{\tau}||_{D^\nu_{\rho_4}} + \epsilon^2 ||\hat{h}_I \circ \tilde{\tau} - \hat{h}_I||_{D^\nu_{\rho_4}} \right), \tag{6.7}$$

$$\sup_{x \in D^\nu_{\rho_4 - \delta_4}} \left| \frac{dh_0}{dt} \right| \leq \frac{\tilde{c}_1}{\delta_4^2} ||h_0||_{D^\nu_{\rho_4}} ||\tilde{R}_1||_{D^\nu_{\rho_4}}. \tag{6.8}$$

Consequently we have

$$|h_\nu(t)| \leq |h_\nu(0)| + \tilde{T}_D Q_\nu \tag{6.9}$$

with

$$Q_\nu = \frac{\tilde{c}_1}{\delta_4^2} ||h_\nu||_{D^\nu_{\rho_4}} \left(||R_2 + \tilde{R}_1||_{D^\nu_{\rho_4}} + \epsilon^2 ||\hat{h}_{BF} \circ \tilde{\tau}||_{D^\nu_{\rho_4}} + \epsilon^2 ||\hat{h}_I \circ \tilde{\tau} - \hat{h}_I||_{D^\nu_{\rho_4}} \right),$$

and

$$|h_0(t)| \leq |h_0(0)| + \tilde{T}_D Q_0, \tag{6.10}$$

with

$$Q_0 = \frac{\tilde{c}_1}{\delta_4^2} ||h_0||_{D^\nu_{\rho_4}} ||\tilde{R}_1||_{D^\nu_{\rho_4}},$$

where $|t| \leq \tilde{T}_D$, and $\tilde{T}_D$ is such that trajectories are in $D^\nu_{\rho_4 - \delta_4}$, $\forall |t| \leq \tilde{T}_D$. ¿From the definitions of h_ν and h_0, (6.9) and (6.10) imply that for $|t| \leq \tilde{T}_D$ we have

$$|J_\Gamma(t)| \leq |J_\Gamma(0)| + \frac{\tilde{T}_D}{2\epsilon^2 C_\Gamma \nu} Q_\nu, \tag{6.11}$$

and

$$|a_\gamma(t)|^2 \leq |a_\gamma(0)|^2 + \sum_{\substack{\gamma \in \mathcal{J} \\ \gamma \neq \Gamma}} |a_\gamma(0)|^2 + |J_\Gamma(t)| + |J_\Gamma(0)| + \frac{\tilde{T}_D}{\omega} Q_0, \quad \forall \gamma \neq \Gamma. \tag{6.12}$$

¿From assumption (6.1) on the initial conditions, (6.11) becomes

$$|J_\Gamma(t)| \leq \xi \rho_2 \rho_5 + \frac{\tilde{T}_D}{2\epsilon^2 C_\Gamma \nu} Q_\nu, \tag{6.13}$$

so that from (6.12) and (6.13) we have

$$|a_\gamma(t)| \leq 2n\xi^2 \rho_5^2 + 2\xi \rho_5 \rho_2 \frac{\tilde{T}_D}{2\epsilon^2 C_\Gamma \nu} Q_\nu + \frac{\tilde{T}_D}{\omega} Q_0, \quad \forall \gamma \neq \Gamma. \tag{6.14}$$

By our choice of ξ in (6.1), we have $2n\xi^2\rho_5^2 + 2\xi\rho_5\rho_2 \leq \frac{\rho_5^2}{2}$. Then, if $\tilde{T}_D$ is such that

$$\tilde{T}_D Q_\Gamma = \frac{\tilde{T}_D}{2\epsilon^2 C_\Gamma \nu} Q_\nu + \frac{\tilde{T}_D}{\omega} Q_0 \leq \frac{{\rho_5}^2}{2}, \tag{6.15}$$

(6.13) and (6.14) imply the bounds of (6.2) for $|t| \leq \tilde{T}_D$. On the other hand, estimating the size of h_ν and h_0 in $D^\nu_{\rho_4}$ we have

$$Q_\Gamma \leq \frac{\tilde{c}_1 \rho_4^2}{\delta_4^2} \left(||R_2||_{D^\nu_{\rho_4}} + (2n+2) ||\tilde{R}_1||_{D^\nu_{\rho_4}} + \epsilon^2 ||\hat{h}_{BF} \circ \tilde{\tau}||_{D^\nu_{\rho_4}} + \epsilon^2 ||\hat{h}_I \circ \tilde{\tau} - \hat{h}_I||_{D^\nu_{\rho_4}} \right), \tag{6.16}$$

so that (6.15) holds for any $\tilde{T}_D \leq T_D$, since T_D satisfies (6.3). Suppose now that some trajectory with initial conditions as in (6.1) crosses the boundary of $D^{\nu}_{\rho_4-\delta_4}$ for the first time at some time t_0, with $t_0 < T_D$. But then we have a contradiction since by (6.2) we must have that $x(t_0) \in D^{\nu}_{\rho_5} \subset D^{\nu}_{\rho_4-\delta_4}$. Hence choosing some T_D satisfying assumption (6.3), $\tilde{T}_D$ can be as large as T_D, and the lemma follows.

■

Proof of Theorem 3.1: We want to estimate the drift of J_Γ and a_γ, $\gamma \neq \Gamma$ using the transformed Hamiltonian $H \circ \tau_1 \circ \tau_2 \circ \tau_3$. Setting $\Delta = \frac{R}{2}$ in Lemma 4.1 and using assumption (3.1) of the theorem we see that the hypothesis of Lemma 4.1 is satisfied and therefore $H \circ \tau_1$ is well defined in $D_{\frac{R}{2}}$. Now let $\nu = \tau_1(\nu')$. ¿From the conditions (3.5) on ν', (3.2) of the statement, and the estimate for the difference between ν' and ν implied by (4.6) in Lemma 4.1, we can see that $D^{\nu}_{\rho_2} \in D_{\frac{R}{2}}$.

To specify the domains of τ_2 and τ_3 we set $\rho_2 = \epsilon^\sigma$, $\delta_2 = \frac{1}{2}\rho_2$ in Lemma 5.1, and $\rho_3 = \frac{1}{2}\rho_2$, $\delta_3 = \frac{1}{4}\rho_2$ in Lemma 5.2. Then, by the definition of C_{NBF} we find that condition (3.3) of the theorem implies (5.6) in the hypothesis of Lemma 5.1. To see this we note that

$$\hat{h}_{2,(\gamma,\Gamma)} = \hat{h}_2 - \hat{h}_{2,\Gamma} - \hat{h}_{2,\gamma} = \sum_{(\gamma_1,\gamma_2,\gamma_3,\gamma_4)\in\mathcal{Q}} I_{\gamma_1,\gamma_2,\gamma_3,\gamma_4} a_{\gamma_1} a_{\gamma_2} a^*_{\gamma_3} a^*_{\gamma_4}, \tag{6.17}$$

for some set $\mathcal{Q}$ of index quartets for which exactly one of them is Γ. Using the variables a_γ, a^*_γ for $\gamma \neq \Gamma$, and J_Γ, θ_Γ, the sum of (6.17) consists of monomials $\sqrt{\nu + J_\Gamma} e^{\pm i\theta_\Gamma} a_\phi a^*_\chi a^*_\psi$, for some ϕ, χ, $\psi \in \mathcal{J}$, so that

$$||\hat{h}_N||_{D^{\nu}_{\epsilon^\sigma}} \leq 2^4(\sqrt{\nu} + \epsilon^\sigma)\epsilon^{3\sigma} \left(\sum_{(\gamma_1,\gamma_2,\gamma_3,\gamma_4)\in\mathcal{Q}} |I_{\gamma_1,\gamma_2,\gamma_3,\gamma_4}| \right). \tag{6.18}$$

Then, from (3.5) and the estimate for $|\nu' - \nu|$, we have

$$||\hat{h}_N||_{D^{\nu}_{\epsilon^\sigma}} \leq \sqrt{\nu'}\frac{3}{2}2^5\epsilon^{3\sigma}||\hat{h}_N||_1. \tag{6.19}$$

¿From the definition of C_{NBF} in (3.11), the hypothesis of Lemma 5.1 follows, and therefore the Hamiltonian $H \circ \tau_1 \circ \tau_2$ is well defined in $D^{\nu}_{\frac{\rho_2}{2}}$.

Also, bounding $\tilde{h}_{(1)}$ and $\tilde{f}_{(1)}$, along the same lines in $D^{\nu}_{\rho_2}$ we find that condition (3.4) implies hypothesis (5.20) of Lemma 5.2. To see this we use the definition of $\tilde{h}_{(1)}$ in (5.16), to find that

$$||\tilde{h}_{(1)}||_{D^{\nu}_{\epsilon^\sigma}} \leq ||\tilde{h}_{(0)}||_{D^{\nu}_{\epsilon^\sigma}} + ||\hat{h}_N||_{D^{\nu}_{\epsilon^\sigma}} \leq$$

$$2^5\sqrt{\frac{3}{2}}\epsilon^{2+3\sigma}C_{NBF}+\epsilon^{2+4\sigma}C_\Gamma+2^4\epsilon^{2+4\sigma}||\hat{h}_{2,\gamma}||_1=\epsilon^{2+3\sigma}\tilde{H}_1, \tag{6.20}$$

while for $\tilde{f}_{(1)}$ we use (5.18) to see that

$$||\tilde{f}_{(1)}||_{D^{\nu}_{\epsilon^\sigma}}\le 2\epsilon^4\beta+3\dot{2}^5\sqrt{\frac{3}{2}}\frac{\pi\tilde{c}_1}{\nu' C_\Gamma}\epsilon^{2+4\sigma}C_{NBF}(2^6\sqrt{\frac{3}{2}}C_{NBF}+\tilde{H}_1)=\epsilon^{2+4\sigma}\tilde{F}_1. \tag{6.21}$$

Thus choosing $\tilde{H}_1$ and $\tilde{F}_1$ as in (3.12) and (3.13), condition (3.4) of the theorem implies assumption (5.3) in Lemma 5.1. Thus the Hamiltonian $H\circ\tau_1\circ\tau_2\circ\tau_3$ is well defined in $D^{\nu}_{\frac{\rho_2}{4}}$.

To estimate the drifts of J_Γ and a_γ, $\gamma\neq\Gamma$ we use Lemma 6.1 with $\rho_4=\frac{\rho_2}{4}$, $\delta_4=\frac{\rho_2}{8}$ and $\rho_5=\frac{\rho_2}{8}$ so that we consider the Hamiltonian flow of $H\circ\tau_1\circ\tau_2\circ\tau_3$ in a subset of $\mathbf{R}^{2n}\cap D^{\nu}_{\frac{\rho_2}{8}}$. We will be using the fact that the transformations and the Hamiltonian flow map real vectors to real vectors. Let now $y'=(\tau_2\circ\tau_3)^{-1}(y)$ and suppose that $y'(0)\in\mathbf{R}^{2n}\cap D^{\nu}_{\frac{\xi\rho_5}{2}}$. We can use (5.13) and (5.27) to estimate the difference between y' and y, obtaining

$$|q_\gamma\circ\tau_2\circ\tau_3-q_\gamma|\le||q_\gamma\circ\tau_2\circ\tau_3-q_\gamma\circ\tau_2||_{D^{\nu}_{\rho_2-\delta_2-\delta_3}}+||q_\gamma\circ\tau_2-q_\gamma||_{D^{\nu}_{\rho_2-\delta_2}}\le\rho_2(C_3+C_2),$$

$$|J_\gamma\circ\tau_2\circ\tau_3-J_\gamma|\le\rho_2^2(C_3+C_2),$$

where

$$C_3=\frac{2e^{\frac{2}{5e^2}}}{e^2}\frac{\tilde{F}_1}{\tilde{H}_1}\epsilon^{1-2\sigma},\quad C_2=\frac{32\tilde{c}_1\pi}{C_\Gamma\sqrt{\nu'}}C_{NBF}\epsilon^\sigma.$$

Using assumption (3.3) of the theorem, and $y'(0)\in\mathbf{R}^{2n}\cap D^{\nu}_{\frac{\xi\rho_5}{2}}$, we therefore have that

$$|q_\gamma\circ\tau_2\circ\tau_3|\le\xi\rho_5,\quad |J_\gamma\circ\tau_2\circ\tau_3|\le\xi\rho_2\rho_5,$$

i.e. $y(0)\in\mathbf{R}^{2n}\cap D^{\nu}_{\xi\rho_5}$. Then, from Lemma 6.1 we have that $y(t)\in\mathbf{R}^{2n}\cap D^{\nu}_{\rho_5}$, $\forall|t|\le T_D$, and using (5.13) and (5.27) and assumption (3.2) of the theorem we see that $y'(t)\in\mathbf{R}^{2n}\cap D^{\nu}_{\frac{\rho_2}{4}}$, i.e. $y'(t)$ is in the domain of $\tau_2\circ\tau_3$, $\forall|t|\le T_D$.

To pass to the original variables we let $y''=(\tau_1)^{-1}(y')$, and use (4.6) and assumption (3.1) to estimate the distance between y'' and y' (and $|\nu'-\nu|$). Arguing as previously, we have that $y''(0)\in\mathbf{R}^{2n}\cap D^{\nu'}_{\frac{\xi\rho_5}{4}}$ implies that $y'(0)\in\mathbf{R}^{2n}\cap D^{\nu'}_{\frac{\xi\rho_5}{2}}$. We can then apply the argument of the previous paragraph to see that $y'(t)\in\mathbf{R}^{2n}\cap D^{\nu}_{\frac{\rho_2}{4}}$, $\forall|t|\le T_D$. Using (4.6) and assumption (3.2) again we see that this implies that $y''(t)\in\mathbf{R}^{2n}\cap D^{\nu'}_{\frac{\rho_2}{2}}$, $\forall|t|\le T_D$.

We can also check that assumption (3.5) on ν' guarantees that ν satisfies the requirements of Lemma 5.1.

To obtain the theorem for the time T_m as in the statement we bound Q_Γ in (6.16) using the estimates (5.24) and (5.25) of R_2 and $\tilde{R}_1$ in Lemma 5.2 and the definitions of α, β, C_{NBF}, $\tilde{H}_1$ and $\tilde{F}_1$. Also, we easily estimate possible values for the constants of m_{BF} and m_I, obtaining $m_{BF} = \epsilon^{2+2\sigma}\tilde{C}_{BF}$ and $m_I = \epsilon^{2+2\sigma}\tilde{C}_I$, with $\tilde{C}_I$ and $\tilde{C}_{BF}$ given by (3.14) and (3.15) respectively. We then have that $T_M \leq T_D$ and the theorem follows.

■

7. Discussion

We conclude with a brief discussion of the physical motivation, applications, and possible extensions of the present work.

Our main motivation comes from a Hamiltonian model of small amplitude gravity water waves on the sphere. The model was developed in [LP], where it was shown that the equations of motion of gravity water waves can be written as a Hamiltonian system of an infinite set of coupled harmonic oscillators. The oscillators (modes) are in one to one correspondence with the spherical harmonics $Y_{[l,m]}$, with $l \in \mathbf{Z}^+$ and $m = -l, -l+1, \ldots, l$. It was also shown in the same work that the water wave system possesses families of approximate periodic orbits for which the amplitude of all but one of the modes is very small (i.e. quasi-monochromatic approximate solutions). In [P] we have seen numerically that trajectories starting near these sets stay nearby for long times. Moreover, we proved that the transfer of energy between modes with different wavenumber l is slow, in particular we showed that the sums I_l of amplitudes of modes with the same wavenumber l Poisson-commute with the quartic part of the second order Birkhoff normal form Hamiltonian. Thus the only possible way for a near-monochromatic motions to decay in a time scale smaller that $O(\epsilon^{-2})$ is through interaction with modes with the same wavenumber l, and we are led to study the projection of the full system to the space of the modes with a given wavenumber l. Such a projection preserves the Hamiltonian structure, moreover it leads to a Hamiltonian satisfying the resonance and symmetry properties (i) and (ii) of section 2, i.e. it belongs to the class of systems we considered here. Note that properties (i) and (ii) follow from the rotational symmetry of the water wave model, and will therefore hold for a variety of other dispersive wave equations on the sphere.

Application of the method we used to the water wave model studied in [P] is quite straightforward. Note that here we have aimed for a stability time scale that is much longer than the time scale needed to understand the numerical observations. The strategy however of performing the normal form once around the origin and then around a periodic orbit can still be applied. In the case where h_{BF} and h_I vanish for a particular mode Γ, we can start with the second order Birkhoff normal form of the water wave model around the origin, and then apply the transformation generated by $\tilde{\chi}_{(1)}$ of Lemma 5.1. It is then easy to see that, up to a remainder of $O(\epsilon^{2+4\sigma})$, the action J_Γ and the quantities I_l Poisson-commute with the Hamiltonian. Thus near-monochromatic trajectories stay $O(\epsilon^\sigma)$ close to periodic near monochromatic trajectories of the truncated quartic normal form for an $O(\epsilon^{-2-4\sigma})$ time. When h_{BF} and h_I are present we can reach a similar stability time provided that an additional bound on the size of h_{BF} is satisfied. A more detailed account of applications to the water wave problem will appear elsewhere.

The motions we studied here are also related to weakly non-linear traveling waves on the line. In the context of waves, a near-monochromatic solution corresponds to a traveling wave with an almost sinusoidal periodic profile and constant amplitude-dependent speed. Stokes derived an expansion for such solutions in surface water waves using a perturbation method that is similar to the Poincare-Linsted method for constructing periodic orbits in dynamical systems (see e.g. [KC]). In the approach we took here, secular terms are avoided by the normal form procedure, while the dependence of the frequency (or wave speed) ω_ϵ on the amplitude is $\omega_\epsilon = \omega + \epsilon^2 C_\Gamma |a_\Gamma|^2 + O(\epsilon^3)$. Also, the symmetry we imposed is analogous to translation invariance in non-linear wave equations, although the particular combination of resonance and symmetry properties (i) and (ii) of Section 2 is quite untypical, leading to the absence of odd order resonant terms. Clearly, the analogies hold at a formal level since the convergence properties of the expansions may be different, for instance the normal form argument we used does not suffice for convergence.

The linear stability of Stokes waves was studied by [BF], who assumed that the wave is given by the first terms in the Stokes expansion. This is analogous to considering the stability of the monochromatic periodic orbits of the truncated normal forms $\overline{h}^{(r)}$. [BF] found that mode interactions that have the form of the terms belonging to $\hat{h}_{BF}$ discussed here can lead to linear instability. Such terms force the modes with index $\gamma \neq \Gamma$ parametrically, with a forcing frequency that is subharmonic (with an $O(\epsilon^2)$ detuning). As we noted, in the case where the Benjamin-Feir quadratic part $\hat{h}_{BF}$ is present the stability time is not optimal, it can be prolonged by a factor $\epsilon^{4\sigma}$, provided that an additional bound of the ratio of $\tilde{C}_{BF}$ (i.e. the size of $\hat{h}_{BF}$) to $\nu' C_\Gamma$ (i.e. the detuning) is assumed. We hope to present these extensions in a future work. Note that a condition involving the ratio of coupling to detuning for avoiding the instability due to the Benjamin-Feir terms is easily obtained using the theory of the Mathieu equation, and is sometimes referred to as "Lighthill condition" (see e.g. [RT]). Our approach suggests that such a condition extends the time of (non-linear) stability, but is not enough for Nekhoroshev stability.

8. Acknowledgments

I would like to thank R. de la Llave, A. Minzoni and A. Olvera for helpful discussions and comments.

9. Appendix

The following technical lemmas are versions of well known consequences of the Cauchy estimates for analytic functions.

Lemma A1 (Poisson bracket estimate in Cartesian coordinates) Let F, G be analytic functions on D_ρ, $D_{\rho-\delta}$ respectively. and $0 < \delta, \delta' < \rho$. Then

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

and $\kappa_1 = 1$.

Lemma A2 (Poisson bracket estimate in mixed Cartesian, action-angle coordinates, see (5.1)) Let F, G be analytic functions on D_ρ^{ν,R_1}, $D_{\rho-\delta}^{\nu,R_1}$ respectively, and $0 < \delta, \delta' < \rho$, $R_1 > 0$. Then

$$||[F,G]||_{D^{\nu,R_1}_{\rho-\delta-\delta'}} \leq \frac{\tilde{c}_1}{\delta(\delta+\delta')} ||F||_{D^{\nu,R_1}_{\rho-\delta'}} ||G||_{D^{\nu,R_1}_\rho}, \quad \text{with} \quad \tilde{c}_1 = 4n(\kappa_1)^2,$$

and κ_1 as in Lemma $A1$.

The following lemma holds by majorizing the Lie-series, using either Lemma A1 for Cartesian coordinates or Lemma A2 for mixed Cartesian, action-angle coordinates. The domain $\tilde{D}_\rho$ below is either D_ρ, or D_ρ^{ν,R_1} (for fixed ν, R_1).

Lemma A3 (convergence of the Lie-series) Let S, H be analytic functions on $\tilde{D}_\rho$, and define $G_0 = H$, $G_r = \frac{1}{r}[S, G_{r-1}]$ for $r \geq 1$. Then

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

where $\kappa_2 = e^2\tilde{c}$ when $\tilde{D}_\rho$ denotes D_ρ, and $\kappa_2 = 2e^2\tilde{c}$ when $\tilde{D}_\rho$ denotes D_ρ^{ν,R_1}. The constants $\tilde{c}$, $\tilde{c}_1$ are as in Lemmas $A1$, $A2$. If $\frac{\kappa_2}{\delta^2}||S||_{\tilde{D}_\rho} < 1$ then $\exp(Ad_S H)$ is analytic in $\tilde{D}_{\rho-\delta}$).

10. REFERENCES

[B] D. Bambusi: Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations, preprint (1998)

[BF] T.B. Benjamin, J. E. Feir: The disintegration of wavetrains on deep water, J. Fluid Mech. 27, 417-430 (1967)

[D] A. Deprit: Canonical transformations depending on a small parameter, Cel. Mech. 20, 1-12 (1967)

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

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

[FGB] F. Fassò, M. Guzzo, G. Benettin: On the stability of elliptic equilibria, Math. Phys. Electron. J. 4, 16 pp., http://www.ma.utexas.edu/mpej/MPEJ.html (1998)

[L] P. Lochak: Canonical perturbation theory via simultaneous approximation, Russ. Math. Surv. 47, 57-133 (1992)

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

[KC] J. Kevorkian, J.D. Cole: Multiple scale and singular perturbation methods, Springer, New York (1996)

[MRS] J. Montaldi, M. Roberts, I. Stewart: Existence of nonlinear normal modes of symmetric Hamiltonian systems, Nonlinearity 3, 695-730 (1990)

[N] L. Niederman: Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system, Nonlinearity 11, 1465-1479 (1998)

[P] P. Panayotaros: Near-monochromatic water waves on the sphere, Physica D, to appear (1998)

[RT] M.I. Rabinowitz, D.I. Trubeskov: Oscillations and waves, Kluwer, Dordrecht (1989)