INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **18** (2005) 1141–1172 doi:10.1088/0951-7715/18/3/012

Quantitative estimates on the normal form around a non-semi-simple 1 : −1 resonant periodic orbit

Mercè Ollé, Juan R Pacha and Jordi Villanueva

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

E-mail: merce.olle@upc.edu, joanr@vilma.upc.edu and jordi@vilma.upc.edu

Received 9 July 2004, in final form 9 November 2004
Published 16 February 2005
Online at stacks.iop.org/Non/18/1141

Recommended by A Chenciner

Abstract

The purpose of this work is to give precise estimates for the size of the remainder of the normalized Hamiltonian around a non-semi-simple 1 : −1 resonant periodic orbit, as a function of the distance to the orbit.

We consider a periodic orbit of a real analytic three-degrees of freedom Hamiltonian system having a pairwise collision of its non-trivial characteristic multipliers on the unit circle. Under generic hypotheses of non-resonance and non-degeneracy of the collision, we present a constructive methodology to reduce the Hamiltonian around the orbit to its (integrable) normal form, up to any given order. This constructive process allows to obtain quantitative estimates for the size of the remainder of the normal form, as a function of the normalizing order. By selecting appropriately this order in terms of the distance R to the resonant orbit (measured using suitable coordinates), $r = r_{\text{opt}}(R) := 2 + \lfloor \exp(W(\log(1/R^{1/(\tau+1+\varepsilon)}))) \rfloor$, we have proved that the size of the remainder can be bounded (for small R) by $R^{r_{\text{opt}}(R)/2}$. Here, $W(\cdot)$ stands for Lambert's W function and verifies that $W(z)\exp(W(z)) = z$, $\tau \geqslant 1$ is the exponent of the required Diophantine condition and $\varepsilon > 0$ is any small quantity. The reasons leading to this bound instead of classical exponentially small estimates are also discussed.

Mathematics Subject Classification: 34C20, 37G05, 37G15

1. Introduction

In this paper we study real analytic three-degrees of freedom Hamiltonian systems having a 1 : −1 resonant periodic orbit. More precisely, we are considering a three-degrees-of-freedom Hamiltonian with a periodic orbit for which its four non-trivial characteristic multipliers

© 2005 IOP Publishing Ltd and London Mathematical Society Printed in the UK

(i.e. those different from one) collide pairwise in the unit circle or, in an equivalent way, its two corresponding normal frequencies are equal.

We note that the situation dealt with in this paper is very common in Hamiltonian mechanics. First of all, we recall that periodic orbits are not isolated in Hamiltonian systems, since if we have a periodic orbit of a Hamiltonian whose monodromy matrix does not have other eigenvalues equal to one except the two trivial ones, then this orbit is contained within a 1-parameter family of periodic orbits (parametrized by the energy, see [22]). What we are asking of this family is that when the parameter moves there is a transition from stability (the four non-trivial characteristic multipliers are different and of modulus one) to complex instability (one non-trivial characteristic multiplier is a complex number of modulus different from one and the other three are the complex conjugate number and the corresponding inverse numbers) through a pairwise collision of its non-trivial characteristic multipliers in the unit circle (at two conjugate points different from ± 1). This transition is usually referred to as the quasi-periodic Hamiltonian Andronov–Hopf bifurcation (see [18] for a broad study of this bifurcation). Hence, the transition orbit is the one we are considering.

This simple mechanism for the generation of such resonant orbits makes this phenomenon common in several mathematical models of science. Here we are not going to be more explicit about the context where this transition has been detected, but refer to the introductions of [10, 16, 17] for a wide description of previous works on the subject, even analytic or numeric.

In a primary classification of this resonance, we can distinguish between two cases depending on the semi-simple or non-semi-simple character of the four-dimensional box of the monodromy matrix associated to the non-trivial multipliers (the box corresponding to the so-called *normal directions* of the orbit). Of course the generic context is the non-semi-simple one, and it is the one we will focus on from now on. This is the non-degeneracy condition claimed in the abstract.

Now we consider a system of symplectic coordinates in $\mathbb{R}^6$ given by $(\theta, x, I, y) \in \mathbb{T}^1 \times \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2$, with the 2-form $\mathrm{d}\theta \wedge \mathrm{d}I + \mathrm{d}x \wedge \mathrm{d}y$, $\mathbb{T}^1 := \mathbb{R}/2\pi\mathbb{Z}$, $x = (x_1, x_2)$ and $y = (y_1, y_2)$. We want θ to be an angular variable describing the resonant periodic orbit, so that it corresponds to the circle $\{I = 0, x = y = 0\}$. Such a (local) system of coordinates always exists for a periodic orbit (see [2, 3]) and we refer to [12] for an example with an explicit construction of it. Using this system as a framework, we perform a Floquet transformation around the resonant periodic orbit in order to reduce the quadratic part of the Hamiltonian in the normal directions (the ones given by (x, y)) to constant coefficients. This is achieved by means of a symplectic change of coordinates, linear with respect to (x, y) and 2π-periodic in θ (see [1, 18] for the proof). The Hamiltonian expressed in the new variables takes the form:

$$\mathcal{H}(\theta, x, I, y) = \omega_1 I + \omega_2(y_1x_2 - y_2x_1) + \frac{\epsilon}{2}(y_1^2 + y_2^2) + \hat{\mathcal{H}}(\theta, x, I, y), \qquad (1)$$

where ω_1 is the angular frequency of the orbit and ω_2 is its (only) normal frequency, so that the non-trivial characteristic multipliers of the orbit are $\{\lambda, \bar{\lambda}, 1/\lambda, 1/\bar{\lambda}\}$, with $\lambda = \exp(2\pi \mathrm{i}\omega_2/\omega_1)$. The function $\hat{\mathcal{H}}$ contains higher order terms in (x, I, y). The sign $\epsilon = \pm 1$ is an invariant of the collision, but by means of a change of time, $t \to -t$, we exchange both contexts (see [16]). In the forthcoming, we shall assume $\epsilon = +1$.

A second criterion of classification for this resonance refers to the relation between ω_1 and ω_2. Hence, the resonance is called rational or irrational depending on the value of ω_1/ω_2. Again we will focus on the generic situation and we will suppose $\omega_1/\omega_2 \notin \mathbb{Q}$. But in this paper we will ask for something more. As usual when working with quantitative estimates on normal forms we will require ω_1 and ω_2 to satisfy a Diophantine condition (see (3)). This is the non-resonance condition we are asking for.

Our set up is Hamiltonian (1), with $\epsilon = +1$, that we assume can be analytically extended to a (complex) neighbourhood of the periodic orbit. Since Poincaré's dissertation [19], a natural question arising from a system like (1) refers to its normal form. Of course, the normal form associated to a $1 : -1$ resonance has been previously investigated and the answer to this question is known. We can quote [1] for the analogous case of a symplectic mapping and [24] for the Hamiltonian Hopf bifurcation at equilibrium points in two-degrees-of-freedom Hamiltonian systems (see also [14, 15, 21]). In [16] the specific case of a $1 : -1$ resonant periodic orbit of a three-degrees-of-freedom Hamiltonian is considered, the normal form is computed up to any order and a detailed analysis of the dynamics of the normal form is given.

The principal difficulties when computing this normal form come from the 'nilpotent' term $(y_1^2 + y_2^2)/2$, which gives rise to non-diagonalizable normal variational equations around the orbit, and hence, to non-diagonalizable homological equations (see (16)). This makes the analysis of the normal form more involved than, for instance, the case of an elliptic periodic orbit.

What we have not found in the literature are previous works where this normal form has been studied from the 'quantitative' point of view, that is, by controlling how the size of the remainder behaves as a function of the normalizing order and of the distance to the periodic orbit. We recall that in the presence of resonances normal forms computed up to infinite order are not convergent, in general. Thus, provided with these estimates on the remainder one can try to optimize the order of the normal form, as a function of the distance to the periodic orbit, in such a way that the size of the remainder becomes as small as possible. This is a classical issue that has been worked out by many researchers (see for instance [6, 7, 11, 23]).

Therefore, this quantitative approach requires not only the identification of the non-removable terms of the normal form (this can be done in a quite standard way) but also to compute explicitly the normalizing transformation, in such a way that we can control how its domain of definition 'shrinks' as a function of the normalizing order. The explicit (constructive) computation of the normal form transformation is not difficult if the homological equations are diagonal (which is the case in the previous works mentioned above), giving rise to simple bounds for the normal form. Thus, one of the main contributions of this paper is to fully develop a constructive algorithm to solve those homological equations, which allows one to compute the normalizing transformation for the Hamiltonian (1) (for instance, by means of the Lie series method). This algorithm can be implemented numerically using a computer (see [12] for a numerical computation of a normal form around an elliptic periodic orbit).

The quantitative result we have obtained is stated as follows.

Theorem 1.1. *We consider the real analytic Hamiltonian $\mathcal{H}$ of (1), with $\epsilon = +1$, that we suppose is defined in a complex domain of the form*

$$\mathcal{D}(\rho_0, R^{(0)}) := \{(\theta, x, I, y) \in \mathbb{C} \times \mathbb{C}^2 \times \mathbb{C} \times \mathbb{C}^2 : |\mathrm{Im}(\theta)| \leqslant \rho_0, \\ |I| \leqslant (R^{(0)})^2, |(x, y)| \leqslant R^{(0)}\} \tag{2}$$

for certain $\rho_0 > 0$ and $R^{(0)} > 0$, where $|\cdot|$ stands for the supremum norm. For this Hamiltonian we assume that the weighted norm introduced in (9) (defined from the Taylor–Fourier expansion of $\mathcal{H}$, see (6)) is finite in this domain, so that $\|\mathcal{H}\|_{\rho_0, R^{(0)}} < +\infty$. We also suppose that the vector $\omega = (\omega_1, \omega_2)$ verifies the Diophantine condition

$$|\langle k, \omega\rangle| \geqslant \gamma |k|_1^{-\tau}, \qquad \forall\, k \in \mathbb{Z}^2 \backslash \{(0, 0)\}, \tag{3}$$

for certain $\gamma > 0$ and $\tau \geqslant 1$, where $\langle\cdot, \cdot\rangle$ means the standard inner product of $\mathbb{R}^2$ and $|k|_1 := |k_1| + |k_2|$.

Then, given any $\varepsilon > 0$ and $\sigma > 1$, both fixed, there exists $0 < R^* < 1$, depending on ρ_0, $R^{(0)}$, $\|\mathcal{H}\|_{\rho_0,R^{(0)}}$, $|\omega_1|$, $|\omega_2|$, γ, τ, ε and σ, such that, for any $0 < R \leqslant R^*$, there is a real analytic canonical diffeomorphism $\Psi^{(R)}$ verifying:

(i) *$\Psi^{(R)}$ is defined in $\mathcal{D}(\sigma^{-2}\rho_0/2, R)$ with $\Psi^{(R)}(\mathcal{D}(\sigma^{-2}\rho_0/2, R)) \subset \mathcal{D}(\rho_0/2, \sigma R)$.*

(ii) *If we write $\Psi^{(R)} - \mathrm{Id} := (\Theta^{(R)}, \mathcal{X}^{(R)}, \mathcal{I}^{(R)}, \mathcal{Y}^{(R)})$, then all the components of this expression are 2π-periodic in θ and satisfy*

$$\|\Theta^{(R)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant \frac{(1-\sigma^{-2})\rho_0}{2}, \qquad \|\mathcal{I}^{(R)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant (\sigma^2-1)R^2,$$
$$\|\mathcal{X}_j^{(R)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant (\sigma-1)R, \quad \|\mathcal{Y}_j^{(R)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant (\sigma-1)R, \qquad j=1,2.$$

(iii) *The transformed Hamiltonian by the action of $\Psi^{(R)}$ takes the form:*

$$\mathcal{H}\circ\Psi^{(R)}(\theta,x,I,y) = Z^{(R)}(x,I,y) + \mathcal{R}^{(R)}(\theta,x,I,y),$$

where $Z^{(R)}$ (the normal form) is an integrable Hamiltonian system which looks like

$$Z^{(R)}(x,I,y) = \omega_1 I + \omega_2(y_1x_2 - y_2x_1) + \frac{1}{2}(y_1^2+y_2^2)$$
$$+\mathcal{Z}^{(R)}\left(I, -\frac{(x_1^2+x_2^2)}{2}, \frac{(y_1x_2-y_2x_1)}{2}\right)$$

with $\mathcal{Z}^{(R)}(u,v,w)$ a polynomial of degree less than or equal to $\lfloor r_{\mathrm{opt}}(R)/2\rfloor$ ($\lfloor\cdot\rfloor$ means the integer part) starting at degree two and $\mathcal{R}^{(R)}$ contains higher order terms on (x,I,y).

(iv) *The expression $r_{\mathrm{opt}}(R)$ is given by*

$$r_{\mathrm{opt}}(R) := 2 + \left\lfloor \exp\left(W\left(\log\left(\frac{1}{R^{1/(\tau+1+\varepsilon)}}\right)\right)\right)\right\rfloor$$

with $W : (0,+\infty) \to (0,+\infty)$ defined from the equation $W(z)\exp(W(z)) = z$.

(v) *$Z^{(R)}$ and $\mathcal{R}^{(R)}$ satisfy the bounds*

$$\|Z^{(R)}\|_R \leqslant \|\mathcal{H}\|_{\rho_0,R^{(0)}}, \qquad \|\mathcal{R}^{(R)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant R^{r_{\mathrm{opt}}(R)/2}. \tag{4}$$

(vi) *$\mathcal{R}^{(R)}$ goes to zero with R faster than any algebraic order, so that*

$$\lim_{R\to 0^+} \frac{\|\mathcal{R}^{(R)}\|_{\sigma^{-2}\rho_0/2,R}}{R^n} = 0, \qquad \forall n \geqslant 1.$$

Remark 1.2. The function W on the statement corresponds to the principal branch of a special function $W : \mathbb{C} \to \mathbb{C}$ known as the Lambert W function (see [4] for details on it). Moreover, for any $\tau > 1$ fixed, the Lebesgue measure of the complementary of the set of $\omega \in \mathbb{R}^2$ for which there is $\gamma > 0$ verifying (3) is zero (see [13], appendix 4).

Remark 1.3. We have formulated theorem 1.1 in terms of weighted norms instead of the supremum one because when working with them the proof becomes simpler. We note that they provide upper bounds for the supremum norm (see section 2 for more details).

At this point, it is natural to compare theorem 1.1 with previous quantitative results on normal forms. We select the case of an elliptic periodic orbit of a real analytic three-degrees-of-freedom Hamiltonian system. After a Floquet transformation the Hamiltonian around this orbit is

$$\mathcal{H}(\theta,x,I,y) = \omega_1 I + \frac{\omega_2}{2}(x_1^2+y_1^2) + \frac{\omega_3}{2}(x_2^2+y_2^2) + \hat{\mathcal{H}}(\theta,x,I,y). \tag{5}$$

If the vector of frequencies $\omega = (\omega_1, \omega_2, \omega_3)$ satisfies a Diophantine condition of the form $|\langle k, \omega\rangle| \geqslant \gamma |k|_1^{-\tau}$, $k \in \mathbb{Z}^3\backslash\{(0, 0, 0)\}$, with $\gamma > 0$ and $\tau \geqslant 2$, then we can compute, up to any arbitrary order r, the normal form around the orbit, that takes the form (we keep the same name for the transformed Hamiltonian and variables)

$$\mathcal{H}(\theta, x, I, y) = \mathcal{Z}^{(r)}\left(I, \frac{(x_1^2 + y_1^2)}{2}, \frac{(x_2^2 + y_2^2)}{2}\right) + \mathcal{R}^{(r)}(\theta, x, I, y),$$

where $\mathcal{Z}^{(r)}(u, v, w)$ is a polynomial of degree less than or equal to $\lfloor r/2\rfloor$ and $\mathcal{R}^{(r)}$ contains higher order terms in (x, I, y). What is proved in this case (under the same analyticity hypotheses as in theorem 1.1) is that if we pick the normalizing order $r = r_{\mathrm{opt}}(R) := \lfloor b/R^{2/(\tau+1)}\rfloor$, then $\|\mathcal{R}^{(r)}\|_{\rho_0/4,R} \leqslant a \exp(-b/R^{2/(\tau+1)})$, for certain $a, b > 0$ independent of R (see [11]). This implies that an exponentially small bound for the remainder is obtained in the elliptic case, which is something very much smaller than (4). Let us try to explain briefly the reason for this completely different behaviour (see also remark 8.2).

We recall that the non semi-simple character of the variational equations of the resonant periodic orbit of (1) is translated to the homological equations (16). Then, when solving those equations we find the same phenomenon as when solving an algebraic linear system with a matrix given by a n-dimensional Jordan box, with λ at the diagonal: the last component of the solution we compute has a factor of the form $1/\lambda^n$. In our context, we can identify n with the order of the normal form we are processing at the present time, so that the size of those boxes increases with the order, and λ is an integer combination of ω_1 and ω_2, depending also on the order (the so-called *small divisors*). But in the elliptic case homological equations are diagonal and only one small divisor appears in the denominator of any component of the solution. This is a colossal difference between the semi-simple and non-semi-simple context, that is responsible for the poor estimate (4). We do not claim that the estimates on theorem 1.1 are optimal, but we are convinced that they cannot be strongly improved with the standard quantitative approach to bound normal forms (see remarks 7.2 and A.9).

The contents of this paper are organized as follows. The first step is to introduce the basic notations and definitions we will use to formulate and prove the results of the paper. This is done in section 2. In section 3 we present the specific method we will use to construct the (canonical) normal form transformation, which is based on the Lie series method. Section 4 tackles the resolution of the homological equations related to the normal form process. More concretely, a constructive algorithm to solve the order s homological equation and to obtain the order s non-removable terms is presented, for any $s \geqslant 3$. The (formal) solutions thus obtained are studied in quantitative form in section 5. In section 6 the full normal form process of section 3 is considered from the quantitative point of view, giving bounds for the generating function of the transformation in terms of the order. In section 7 those bounds are converted into bounds for the remainder of the normal form, as a function of the order and of the distance to the resonant orbit. The optimal way to select the normal form order as a function of this distance is discussed in section 8. The proof of theorem 1.1 is given in section 9. Finally, appendix A is an appendix where we have compiled some technical results used in the course of the paper.

2. Basic notations and definitions

In this section we introduce some notations and definitions to be used throughout the paper.

We shall denote by $\mathbb{E}$ the complex vectorial space of formal Taylor–Fourier series of type

$$\mathbb{E} := \left\{ f = f(\theta, q, I, p) : f = \sum_{k,l,m,n} f_{k,l,m,n} I^l q^m p^n \exp(\mathrm{i}k\theta) \right\} \tag{6}$$

with $q = (q_1, q_2)$, $p = (p_1, p_2)$, $k \in \mathbb{Z}$, $l \in \mathbb{Z}_+$ and $m, n \in \mathbb{Z}_+^2$, where $Z_+ = \mathbb{N} \cup \{0\}$. Here and throughout the paper we will use extensively the standard multi-index notation: $q^m = q_1^{m_1} q_2^{m_2}$ and so on. Given a monomial $\alpha = I^l q^m p^n \exp(\mathrm{i}k\theta) \in \mathbb{E}$, we define its *adapted degree* or *weighted order* (so its degree from now on) as $\deg(\alpha) := 2l + |m|_1 + |n|_1$. Hence, the degree on I is counted twice with respect to the degree on (q, p). The expression $|u|_1$ means the *absolute norm* of a complex vector u, given by $|u|_1 := \sum_i |u_i|$. In the same way, we will write $|u|$ for the *supremum norm* of u. We will extend the same notations to denote the associated matrix norms.

This definition of degree is motivated by the action of the *Poisson bracket* on $\mathbb{E}$. Given $f, g \in \mathbb{E}$, we define its Poisson bracket by

$$\{f, g\} := \partial_\theta f \cdot \partial_I g - \partial_I f \cdot \partial_\theta g + \sum_{i=1}^{2} (\partial_{q_i} f \cdot \partial_{p_i} g - \partial_{p_i} f \cdot \partial_{q_i} g). \tag{7}$$

Thus, if $\alpha, \beta \in \mathbb{E}$ are two monomials, then $\deg(\{\alpha, \beta\}) = \deg(\alpha) + \deg(\beta) - 2$, giving the homogeneity of the Poisson bracket with respect to the adapted degree. Related to the Poisson bracket, we also introduce the (linear) Lie operator associated to each $u \in \mathbb{E}$:

$$\begin{aligned} L_u : \mathbb{E} &\longrightarrow \mathbb{E}, \\ f &\mapsto L_u f = \{f, u\}. \end{aligned}$$

During the normalizing process we will work with several subspaces of $\mathbb{E}$. Thus, $\mathbb{E}_s$ will stand for the subspace of $\mathbb{E}$ generated by the monomials of degree s. Moreover, given $k \in \mathbb{Z}$ and $l, M, N \in \mathbb{Z}_+$, we denote by $\mathbb{E}_{k,l,M,N}$ the subspace of $\mathbb{E}_{2l+M+N}$ spanned by the monomials of the form $I^l q^m p^n \exp(\mathrm{i}k\theta)$, with $|m|_1 = M$ and $|n|_1 = N$. Hence,

$$\mathbb{E}_s = \bigoplus_{\substack{k \in \mathbb{Z} \\ 2l + M + N = s}} \mathbb{E}_{k,l,M,N}. \tag{8}$$

We will also refer to general (complex) polynomials on the variables (q, I, p) or (q, p) using the standard notations $\mathbb{C}_s[q, I, p]$ (recall we are counting twice the degree on I) and $\mathbb{C}_s[q, p]$.

As we want to work with analytic functions, we are interested in convergent series expansions. Thus, given $\rho > 0$ and $R > 0$, we say that $f \in \mathbb{E}(\rho, R)$ if $f \in \mathbb{E}$ and the norm

$$\|f\|_{\rho,R} := \sum_{k,l,m,n} |f_{k,l,m,n}| R^{2l+|m|_1+|n|_1} \exp(|k|\rho) \tag{9}$$

is finite. As we have mentioned before in remark 1.3, $\|f\|_{\rho,R}$ is an upper bound for the supremum norm of f in $\mathcal{D}(\rho, R)$ (see (2) defined using (q, p) instead of (x, y)). If $\|f\|_{\rho,R} < +\infty$ it implies that f is analytic in the interior of $\mathcal{D}(\rho, R)$ and bounded on the boundary. Conversely, if f is analytic in a neighbourhood of $\mathcal{D}(\rho, R)$, then $\|f\|_{\rho,R} < +\infty$. We point out that the definitions of $\|\cdot\|_{\rho,R}$ and $\mathcal{D}(\rho, R)$ are coherent with the adapted degree.

In an analogous way, given $f \in \mathbb{E}_s$ we say that it belongs to $\mathbb{E}_s(\rho)$ if the norm

$$\|f\|_\rho := \sum_{\substack{k,l,m,n \\ 2l+m+n=s}} |f_{k,l,m,n}| \exp(|k|\rho) \tag{10}$$

is finite. For such s-homogeneous f we have $\|f\|_{\rho,R} = \|f\|_\rho R^s$. A very important case is when $f \in \mathbb{C}_s[q, I, p]$, where ρ plays no role. For such an f we define

$$\|f\| := \sum_{\substack{l,m,n\\ 2l+m+n=s}} |f_{0,l,m,n}|. \tag{11}$$

Some basic properties of these norms are surveyed in appendix A. Those properties are very similar to the ones for the classical supremum norm. Therefore, we have preferred to work with weighted norms instead of the supremum norm because several estimates become simpler with them, for instance those on small divisors. Thus, weighted norms have been widely used to quantify the effect of small divisors in normal forms (see, e.g. [6, 7, 11, 18, 23]). Alternatively, one can work with the supremum norm and use the estimates of Rüssmann on small divisors (see [20]).

The last definition in this section refers to the symmetries we will ask for the complex coefficients of $f \in \mathbb{E}$. In section 3 we will introduce a complexified version of Hamiltonian (1), so that the variables (q, p) will follow from (x, y) by means of a suitable linear but complex transformation (see (17)). This 'complexification' implies several symmetries on the coefficients of the functions involved. Let $\mathcal{S} : \mathbb{E} \to \mathbb{E}$ be the (linear) operator defined by

$$f \mapsto \mathcal{S}(f) := \sum_{k,l,m,n} (-1)^{m_1+n_2} \bar{f}_{-k,l,n_2,n_1,m_2,m_1} I^l q^m p^n \exp(\mathrm{i}k\theta), \tag{12}$$

where bar denotes complex conjugation. Then, we will use the notation $\mathbb{E}^{\mathcal{S}}$ to refer to the elements $f \in \mathbb{E}$ such that $\mathcal{S}(f) = f$. Equivalently, we will say that $f \in \mathbb{E}^{\mathcal{S}}$ satisfies the $\mathcal{S}$-symmetries. We extend the $\mathcal{S}$-notation to $\mathbb{E}_s^{\mathcal{S}}$ and $\mathbb{E}_{0,l,M,N}^{\mathcal{S}}$. It is important to note that if $f, g \in \mathbb{E}^{\mathcal{S}}$, then $\{f, g\} \in \mathbb{E}^{\mathcal{S}}$, so that the $\mathcal{S}$-symmetries are preserved by Poisson brackets.

As a final remark, we observe that the definition of $\mathbb{E}$ and the related ones have been made (for convenience) in terms of (q, p), but everything remains valid replacing (q, p) by (x, y).

3. Construction of the normalizing transformation

This section describes the specific method used to construct the normal form transformation.

A very natural way to compute the normalizing transformation is to look for it as a composition of a sequence of canonical transformations, in such a way that the normal form is computed 'degree-wise' by choosing the sth-transformation to remove the non-resonant terms of degree $s + 2$ from the Hamiltonian obtained after the previous step, for $s \geqslant 1$. However, if one is interested in further applications of the normal form (for instance a quantitative analysis or a numerical implementation), it is advisable to use some 'closed' algorithm which computes the normalizing transformation from a *single* canonical change. More precisely, throughout this work we shall use the Giorgilli–Galgani algorithm (see [7–9, 23]) applied to formal Taylor–Fourier series in $\mathbb{E}$.

Definition 3.1 (the Giorgilli–Galgani algorithm). *Let $G = \sum_{s\geqslant 3} G_s$, with $G_s \in \mathbb{E}_s$. We define the map $T_G : \mathbb{E} \to \mathbb{E}$ in the following way. Given $f = \sum_{l\geqslant 1} f_l$, with $f_l \in \mathbb{E}_l$, then $T_G f = \sum_{s\geqslant 1} F_s$, where $F_s = \sum_{l=1}^{s} f_{l,s-l}$ and the terms $f_{l,s-l} \in \mathbb{E}_s$ are computed recursively by*

$$f_{l,0} = f_l, \qquad f_{l,s} = \sum_{j=1}^{s} \frac{j}{s}\{f_{l,s-j}, G_{2+j}\}, \quad s \geqslant 1. \tag{13}$$

Usually, the sum G is known as the generating function of the transformation.

Remark 3.2. The very important property of the transformation T_G is that if f and G are convergent, then $T_G f = f \circ \Phi^{-\chi}_{\varepsilon=1}$, where $\Phi^{-\chi}_{\varepsilon}$ means the flow time ε of the non-autonomous Hamiltonian $-\chi$ (the time is ε), with $\chi = \sum_{s\geqslant 1} sG_{s+2}\varepsilon^{s-1}$. So, the coordinate transformation given by $\phi = T_G\theta$, $J = T_G I$, $Q_i = T_G q_i$ and $P_i = T_G p_i$ $(i = 1, 2)$ is canonical and 2π-periodic in θ. For an account of these properties, together with their corresponding proofs, see [8].

The idea is thus to take $f = \mathcal{H}$, the Hamilton function, to construct an *ad hoc* generating function G and to apply algorithm 3.1 to cast $\mathcal{H}$ into its normal form. This reduction process can be done recursively and the following algorithm can be given to determine both G and the normal form.

Proposition 3.3. *Consider* $H = \sum_{s\geqslant 2} H_s$, *with* $H_s \in \mathbb{E}_s$, *and the generating function* $G = \sum_{s\geqslant 3} G_s$. *If we write* $T_G H = \sum_{s\geqslant 2} Z_s$, *with* $Z_s \in \mathbb{E}_s$, *then the following relations are satisfied:*

$$Z_2 = H_2, \qquad \{G_s, H_2\} + Z_s = F_s, \quad s \geqslant 3, \tag{14}$$

where,

$$\begin{aligned} F_3 &= H_3, \\ F_s &= \sum_{j=1}^{s-3} \frac{j}{s-2}\{Z_{s-j}, G_{2+j}\} + \sum_{j=1}^{s-2} \frac{j}{s-2} H_{2+j,s-j-2}, \qquad s \geqslant 4. \end{aligned} \tag{15}$$

Here, the quantities $H_{l,k}$ *may be computed recursively from formulae (13).*

The proof of proposition 3.3 is formally identical to the proof of the corresponding classical one in [7] and [23], so the reader is referred to these.

From relation (14), it is clear that the important point is to investigate the solvability, in terms of G_s and Z_s, of the *homological equations*

$$L_{H_2} G_s + Z_s = F_s, \tag{16}$$

for a given $F_s \in \mathbb{E}_s$, in such a way that Z_s takes the simplest possible form.

In order to deal with such homological equations, it is convenient to introduce the following (complex) coordinates,

$$x_1 = \frac{q_1 - p_2}{\sqrt{2}}, \qquad x_2 = \frac{\mathrm{i}(q_1 + p_2)}{\sqrt{2}}, \qquad y_1 = \frac{q_2 + p_1}{\sqrt{2}}, \qquad y_2 = \frac{\mathrm{i}(q_2 - p_1)}{\sqrt{2}}. \tag{17}$$

These relations define a linear canonical change

$$(\theta, x, I, y) := \varphi(\theta, q, I, p), \tag{18}$$

which transforms the Hamiltonian $\mathcal{H}$ of (1) into $H := \mathcal{H} \circ \varphi$, which we expand as

$$H(\theta, q, I, p) = H_2(q, I, p) + \sum_{l\geqslant 3} H_l(\theta, q, I, p) \tag{19}$$

with $H_l \in \mathbb{E}_l$, where H_2 is given (for $\epsilon = +1$) by

$$H_2 = \omega_1 I + \mathrm{i}\omega_2(q_1 p_1 + q_2 p_2) + q_2 p_1. \tag{20}$$

These kind of transformations are usual in normal forms to put the homological equations in diagonal form. In our case this is not possible but, as we will explain in section 4, using complex coordinates we can re-write (16) as a family of algebraic systems of linear equations with lower-triangular matrices.

Remark 3.4. The price we pay for these simplifications on the homological equations is that Hamiltonian (19) is not given by a real analytic function but by a complex analytic one, so that real values of (x, y) correspond to complex values of (q, p) such that $\bar{q}_1 = -p_2$ and $\bar{q}_2 = p_1$. This is equivalent to saying that H verifies $\mathcal{S}(H) = H$ (see (12)), and hence $H \in \mathbb{E}^{\mathcal{S}}$. Conversely, given any convergent series $F \in \mathbb{E}^{\mathcal{S}}$, then $F \circ \varphi^{-1}$ is a real analytic function.

An alternative to the complexification of $\mathcal{H}$ is the chance to perform the normal form on the real Hamiltonian and complexify the system only when solving the homological equations. We have preferred to work with a complex Hamiltonian, but both approaches are equivalent.

4. Resolution of the homological equations

Given $F_s \in \mathbb{E}_s^{\mathcal{S}}$, $s \geqslant 3$, the target of this section is to look for the simplest expression for $Z_s \in \mathbb{E}_s^{\mathcal{S}}$ (*the normal form*) in such a way that there exists $G_s \in \mathbb{E}_s^{\mathcal{S}}$ verifying (16) (see section 2 for notations). We recall that we know *a priori* how Z_s looks because the normal form around a periodic orbit with a non-semi-simple $1 : -1$ resonance has been previously investigated (see [1, 16, 18]).

Proposition 4.1. *We consider H_2 defined in (20) with $\omega_1/\omega_2 \notin \mathbb{Q}$ and the decomposition $\mathbb{E}_s = \mathbb{E}_s^0 \oplus \mathbb{E}_s^+$, with*

$$\mathbb{E}_s^0 = \bigoplus_{M=N} \mathbb{E}_{0,l,M,N}, \qquad \mathbb{E}_s^+ = \bigoplus_{|k|+|M-N|\neq 0} \mathbb{E}_{k,l,M,N}, \qquad 2l + M + N = s,$$

so that $\mathbb{E}_s^0 = \{0\}$ if s is odd. Given $F_s \in \mathbb{E}_s^{\mathcal{S}}$, $s \geqslant 3$, we write it as

$$F_s = F_s^0 + F_s^+, \qquad F_s^0 \in \mathbb{E}_s^0, \qquad F_s^+ \in \mathbb{E}_s^+. \tag{21}$$

Then, we have:

(a) There exists a unique $G_s^+ \in \mathbb{E}_s^+$ such that $L_{H_2} G_s^+ = F_s^+$. Moreover, $G_s^+ \in \mathbb{E}_s^{\mathcal{S}}$.

(b) If s is even, there is a real homogeneous polynomial $\mathcal{Z}_s(u, v, w)$ of (standard) degree $s/2$, which is uniquely defined in terms of F_s^0, such that if we set

$$Z_s(q, I, p) := \mathcal{Z}_s\left(I, q_1 p_2, \frac{\mathrm{i}(q_1 p_1 + q_2 p_2)}{2}\right), \tag{22}$$

then $Z_s \in \mathbb{E}_s^0 \cap \mathbb{E}_s^{\mathcal{S}}$ and there is $G_s^0 \in \mathbb{E}_s^0$ (not unique) verifying $L_{H_2} G_s^0 + Z_s = F_s^0$. Moreover, G_s^0 can be chosen so that $G_s^0 \in \mathbb{E}_s^{\mathcal{S}}$.

Consequently, if s is odd $G_s := G_s^+$ and $Z_s := 0$ solve equation (16) in $\mathbb{E}_s^{\mathcal{S}}$, and if s is even so does $G_s := G_s^+ + G_s^0$ and Z_s given by (22).

Remark 4.2. To relate proposition 4.1 with the normal form given in theorem 1.1, we observe that if we return to real variables by means of the inverse change of (17) one gets

$$q_1 p_2 = -\frac{(x_1^2 + x_2^2)}{2}, \qquad \frac{\mathrm{i}(q_1 p_1 + q_2 p_2)}{2} = \frac{(y_1 x_2 - y_2 x_1)}{2}. \tag{23}$$

But the statement of proposition 4.1 does not completely solve our problem, as we want to compute *constructively* Z_s and G_s, so that we can give precise bounds on them. Thus, from here and till the end of section 4 we are going to perform a constructive proof of proposition 4.1.

First of all, to simplify notations, we assume degree $s \geqslant 3$ fixed and we skip the subscript s of F_s, G_s and Z_s. At this stage, we consider a generic $F \in \mathbb{E}_s$, whereas the role of the $\mathcal{S}$-symmetries will be considered later on.

Given a monomial $\alpha = I^l q^m p^n \exp(\mathrm{i}k\theta)$, direct computations show that

$$L_{H_2}\alpha = \left(\Omega + m_1\frac{q_2}{q_1} - n_2\frac{p_1}{p_2}\right)\alpha, \tag{24}$$

where H_2 is the one in (20) and Ω has been introduced as

$$\Omega \equiv \Omega_{k,|m|_1,|n|_1} := \mathrm{i}\omega_1 k + \mathrm{i}\omega_2(|m|_1 - |n|_1). \tag{25}$$

On the other hand, the quotient q_2/q_1 does not appear if the monomial α has $m_1 = 0$. Similarly for the quotient p_1/p_2. With this remark (24) is fully justified. Expression (24) implies that all the subspaces $\mathbb{E}_{k,l,M,N} \subset \mathbb{E}_s$ introduced in (8) are invariant under the action of L_{H_2}. Thus, given $F \in \mathbb{E}_s$ we decompose it according to its projection into the subspaces $E_{k,l,M,N}$, so that $F = \sum_{k,l,M,N} F_{k,l,M,N}$, with $2l + M + N = s$, and the same for G and Z, and we investigate separately the equations

$$L_{H_2}G_{k,l,M,N} + Z_{k,l,M,N} = F_{k,l,M,N}. \tag{26}$$

As will be discussed below, equations with $k \neq 0$ or $M \neq N$, so that $\mathbb{E}_{k,l,M,N} \subset \mathbb{E}_s^+$, can be solved by simply setting $Z_{k,l,M,N} \equiv 0$. The only possible non-removable (or *resonant*) monomials are those in $\mathbb{E}_{0,l,M,M}$, with $2l+2M = s$ (so s has to be even and then $\mathbb{E}_{0,l,M,M} \subset \mathbb{E}_s^0$). The explanation of this fact is that the condition $\omega_1/\omega_2 \notin \mathbb{Q}$ implies that $\Omega_{k,l,M,N} = 0$ if and only if $k = 0$ and $M = N$. In sections 4.1 and 4.2 we will consider both cases separately.

4.1. The case $\Omega \neq 0$

The first point we remark is that if $F \in \mathbb{E}_{k,l,M,N} \subset \mathbb{E}_s$, then its coefficients are readily determined by just a pair of subscripts. In view of (6), it is advisable to denote $f_{m,n} := f_{k,l,m,M-m,N-n,n}$, with $0 \leqslant m \leqslant M$ and $0 \leqslant n \leqslant N$ (we skip k and l since they are held fixed). With the notation above, equation (26) can be translated into a system of linear equations for the (complex) coefficients of $G_{k,l,M,N}$ ($\dim_{\mathbb{C}} \mathbb{E}_{k,l,M,N} = (M+1)(N+1)$). For this purpose we introduce $f \equiv f_{k,l,M,N}$ and $g \equiv g_{k,l,M,N}$ as the arrays holding the coefficients $f_{m,n}$ and $g_{m,n}$, respectively. We order those coefficients through the following claim: $g_{\mu,\beta} \prec g_{\alpha,\sigma}$ ($g_{\mu,\beta}$ precedes $g_{\alpha,\sigma}$) if $\mu > \alpha$ or, when $\mu = \alpha$, if $\beta > \sigma$ (the same for $f_{m,n}$). Therefore, g takes the form $g^* = (g_M^*, g_{M-1}^*, \ldots, g_0^*)$, still with $g_j^* = (g_{j,N}, g_{j,N-1}, g_{j,N-2}, \ldots, g_{j,1}, g_{j,0})$, for $j = 0, \ldots, M$, where the star means transposition (and identically for f). Hence, if we set $Z_{k,l,M,N} \equiv 0$, then equation (26) is equivalent to solve $\Lambda g = f$, with $\Lambda \equiv \Lambda_{k,M,N}$ a square matrix of dimension $(M+1)(N+1)$. It is straightforward to check that this system is written block-wise (with $M+1$ blocks of dimension $N+1$) as

$$\begin{pmatrix} D_N & & & & \\ E_M & D_N & & & \\ & E_{M-1} & D_N & & \\ & & \ddots & \ddots & \\ & & & E_1 & D_N \end{pmatrix}\begin{pmatrix} g_M \\ g_{M-1} \\ g_{M-2} \\ \vdots \\ g_0 \end{pmatrix} = \begin{pmatrix} f_M \\ f_{M-1} \\ f_{M-2} \\ \vdots \\ f_0 \end{pmatrix}. \tag{27}$$

The different blocks stand for $E_j = j \cdot Id_{N+1}$, with Id_{N+1} the $(N+1) \times (N+1)$ identity matrix, whereas $D_N = \Omega \cdot Id_{N+1} - P_N$, with $\Omega \equiv \Omega_{k,M,N}$ (see (25)) and

$$P_N = \begin{pmatrix} 0 & & & & \\ N & 0 & & & \\ & N-1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}. \tag{28}$$

It follows from this description that the matrix Λ is a band lower-triangular matrix, where all the entries different from zero are placed at the main diagonal and on two bands (sub-diagonals) below the main diagonal. Moreover, the elements of the diagonal are all equal to Ω. Hence, if $\Omega \neq 0$, this specific part of the homological equations (we mean their components on the subspace $\mathbb{E}_{k,l,M,N}$) has a unique solution which can be easily obtained.

At this point, we have proved part (a) of proposition 4.1, thus solving the projection of equation (16) into $\mathbb{E}^{+}$ by setting $Z = 0$. Clearly,

$$G^{+} = \sum_{\substack{2l+M+N=s \\ |k|+|M-N|\neq 0}} G_{k,l,M,N}. \tag{29}$$

Furthermore, G^{+} satisfies the $\mathcal{S}$-symmetries if F^{+} does. To ensure this one can check that if $G_{k,l,M,N}$ is (the only) solution of $L_{H_2}G_{k,l,M,N} = F_{k,l,M,N}$, then $L_{H_2}\mathcal{S}(G_{k,l,M,N}) = \mathcal{S}(F_{k,l,M,N})$.

Remark 4.3. We point out that the Diophantine condition (3) is not required to ensure convergence of G^{+} if F^{+} does. This is because, for a fixed s, in the divisors $\mathrm{i}\omega_1 k + \mathrm{i}\omega_2(M-N)$ appearing in the homological equation we have that $M - N$ ranges between $-s$ and s, and therefore, for all $k \in \mathbb{Z}$, these divisors are bounded from below whenever ω_1/ω_2 is irrational. But we need (3) to control the amplification factor in the norm of G^{+} with respect to the norm of F^{+}.

4.2. The case $\Omega = 0$

To complete the normal form reduction process we are forced to study the projection of (16) into $\mathbb{E}_s^0$ (i.e. we have to consider (26) with $k = 0$ and $M = N$). Then, we have to inquire which are the terms in F^0 that can be eliminated with a suitable G (even though $\Omega = 0$), and which ones are non-removable and have to remain as resonant terms.

To simplify notations we fix the values of l and M, with $2l + 2M = s$ even, and we denote by L the linear operator L_{H_2} restricted to $\mathbb{E}_{0,0,M,M}$, so that $L : \mathbb{E}_{0,0,M,M} \to \mathbb{E}_{0,0,M,M}$. Thus, we observe that given $F \in \mathbb{E}_{0,l,M,M}$ we can express it as $F = I^l \hat{F}$, with $\hat{F} \in \mathbb{E}_{0,0,M,M}$. Then, we have $L_{H_2}F = I^l L\hat{F}$, and hence the expression I^l is uncoupled from equation (26). Consequently, we only have to investigate the solvability in $\mathbb{E}_{0,0,M,M}$, for $M \geqslant 1$ (the case $M = 0$ is trivial), of

$$L\hat{G} + \hat{Z} = \hat{F}. \tag{30}$$

The important point now is that $\operatorname{Ker} L \neq \{0\}$, which implies that Range L is strictly contained in $\mathbb{E}_{0,0,M,M}$. What we have to do is to find a complementary subspace of Range L in $\mathbb{E}_{0,0,M,M}$. A standard way to do this is to introduce a Hermitian product $\langle\cdot|\cdot\rangle$ in $\mathbb{E}_{0,0,M,M}$ and to consider the decomposition $\mathbb{E}_{0,0,M,M} = \operatorname{Ker} L^{\dagger} \oplus \operatorname{Range} L$, where $L^{\dagger}$ is the adjoint operator of L with respect to $\langle\cdot|\cdot\rangle$. In this paper we do not consider this method because we will construct $\hat{G}$ and $\hat{Z}$ explicitly from $\hat{F}$, but full details of this approach in our context can be found in [18].

Given $\hat{F} \in \mathbb{E}_{0,0,M,M}$ we express it in the form:

$$\hat{F} = \sum_{m=0}^{M}\sum_{n=0}^{M} f_{m,n} q_1^m q_2^{M-m} p_1^{M-n} p_2^n = \sum_{i=0}^{M} (q_1 p_2)^i \sum_{j=0}^{M-i} f_{M-j,i+j} (q_1 p_1)^{M-i-j} (q_2 p_2)^j$$
$$+ \sum_{i=1}^{M} (q_2 p_1)^i \sum_{j=0}^{M-i} f_{M-i-j,j} (q_1 p_1)^{M-i-j} (q_2 p_2)^j. \tag{31}$$

Now, we introduce $\xi = (\xi_1, \xi_2, \xi_3, \xi_4)$ defined by

$$\xi_1 = q_1 p_2, \qquad \xi_2 = q_2 p_1, \qquad \xi_3 = \frac{\mathrm{i}(q_1 p_1 + q_2 p_2)}{2}, \qquad \xi_4 = \frac{(q_1 p_1 - q_2 p_2)}{2}, \tag{32}$$

so that $q_1 p_1 = -\mathrm{i}\xi_3 + \xi_4$ and $q_2 p_2 = -\mathrm{i}\xi_3 - \xi_4$. By replacing (q, p) in terms of ξ in (31), we obtain an expression of the form

$$\hat{F} = \sum_{i=0}^{M} \xi_1^i \sum_{j=0}^{M-i} \hat{f}_{i,j} \xi_3^{M-i-j} \xi_4^j + \sum_{i=1}^{M} \xi_2^i \sum_{j=0}^{M-i} \tilde{f}_{i,j} \xi_3^{M-i-j} \xi_4^j, \tag{33}$$

where the coefficients $\{\hat{f}_{i,j}\}_{0\leqslant j\leqslant M-i}$ are easily obtained from $\{f_{M-j,i+j}\}_{0\leqslant j\leqslant M-i}$—the coefficients in (31)—for any $0 \leqslant i \leqslant M$, and the coefficients $\{\tilde{f}_{i,j}\}_{0\leqslant j\leqslant M-i}$ follow from $\{f_{M-i-j,j}\}_{0\leqslant j\leqslant M-i}$, for any $1 \leqslant i \leqslant M$ (for explicit expressions, see [18]). Moreover, expression (33) is a real polynomial in ξ if $\hat{F}$ verifies the $\mathcal{S}$-symmetries (check that the components of ξ are *real* under those symmetries: $\mathcal{S}(\xi) = \xi$). Now, we look for

$$\hat{G} = \sum_{i=1}^{M} \xi_1^i \sum_{j=0}^{M-i} \hat{g}_{i,j} \xi_3^{M-i-j} \xi_4^j + \sum_{i=0}^{M} \xi_2^i \sum_{j=0}^{M-i} \tilde{g}_{i,j} \xi_3^{M-i-j} \xi_4^j. \tag{34}$$

We stress that the sums defining $\hat{F}$ and $\hat{G}$ are arranged in a different way. This will ease the computation of the coefficients of $\hat{G}$. At this point it is convenient to know how the operator L acts on the ξ-monomials involved in $\hat{G}$. First we have,

$$\{\xi_1, H_2\} = -2\xi_4, \qquad \{\xi_2, H_2\} = 0, \qquad \{\xi_3, H_2\} = 0, \qquad \{\xi_4, H_2\} = \xi_2$$

and hence, if $i \geqslant 1$,

$$\begin{aligned} L(\xi_1^i \xi_3^\nu \xi_4^j) &= \partial_{\xi_1}(\xi_1^i \xi_3^\nu \xi_4^j)\{\xi_1, H_2\} + \partial_{\xi_3}(\xi_1^i \xi_3^\nu \xi_4^j)\{\xi_3, H_2\} + \partial_{\xi_4}(\xi_1^i \xi_3^\nu \xi_4^j)\{\xi_4, H_2\} \\ &= -2i\xi_1^{i-1}\xi_3^\nu \xi_4^{j+1} + j\xi_1^i \xi_2 \xi_3^\nu \xi_4^{j-1} = -(2i+j)\xi_1^{i-1}\xi_3^\nu \xi_4^{j+1} - j\xi_1^{i-1}\xi_3^{\nu+2}\xi_4^{j-1}, \end{aligned}$$

where we have used that $\xi_1\xi_2 = -(\xi_3^2 + \xi_4^2)$. Analogously, $L(\xi_2^i \xi_3^\nu \xi_4^j) = j\xi_2^{i+1}\xi_3^\nu \xi_4^{j-1}$. These computations together on $\hat{G}$ give the following expression for $L\hat{G}$:

$$\begin{aligned} &-\sum_{i=1}^{M} \xi_1^{i-1} \sum_{j=0}^{M-i} \hat{g}_{i,j} \left\{(2i+j)\xi_3^{M-i-j}\xi_4^{j+1} + j\xi_3^{M-i-j+2}\xi_4^{j-1}\right\} + \sum_{i=0}^{M} \xi_2^{i+1} \sum_{j=0}^{M-i} j\tilde{g}_{i,j}\xi_3^{M-i-j}\xi_4^{j-1} \\ &\quad = -\sum_{\nu=0}^{M-1} \xi_1^\nu \Bigg[\hat{g}_{\nu+1,1}\xi_3^{M-\nu} + \sum_{\mu=1}^{M-\nu-2} \left\{(2\nu+\mu+1)\hat{g}_{\nu+1,\mu-1} + (\mu+1)\hat{g}_{\nu+1,\mu+1}\right\} \\ &\qquad \times \xi_3^{M-\nu-\mu}\xi_4^\mu + (\nu+M)\hat{g}_{\nu+1,M-\nu-2}\xi_3\xi_4^{M-\nu-1} + (\nu+M+1)\hat{g}_{\nu+1,M-\nu-1}\xi_4^{M-\nu} \Bigg] \\ &\qquad + \sum_{\nu=1}^{M} \xi_2^\nu \sum_{\mu=0}^{M-\nu} (\mu+1)\tilde{g}_{\nu-1,\mu+1}\xi_3^{M-\nu-\mu}\xi_4^\mu. \end{aligned} \tag{35}$$

Remark 4.4. We observe that the coefficients $\{\tilde{g}_{i,0}\}_{i=0,\ldots,M}$ are missing in $L\hat{G}$, so that they play no role in equation (30). They are responsible for the non-uniqueness of $\hat{G}$. From now on we set them $\tilde{g}_{i,0} = 0$ (to keep the symmetries and to have minimal norm). Moreover, we also point out that there is a small 'abuse' of notation for the expression inside the square brackets when $\nu = M - 1$, because the term $(\nu + M)\hat{g}_{\nu+1,M-\nu-2}\xi_3\xi_4^{M-\nu-1}$ has no sense and has to be removed. Similarly, the sum from $\mu = 1$ to $\mu = M - \nu - 2$ is empty for $\nu = M - 2$ and $\nu = M - 1$.

Now, we have to compare this expression for $L\hat{G}$ with the one for $\hat{F}$ in (33), trying to determine the coefficients $\{\hat{g}_{i,j}\}$ and $\{\tilde{g}_{i,j}\}$ and the *resonant terms* $\hat{Z}$ from (30). Our first try is to set $\hat{Z} \equiv 0$. Then by comparison of the coefficients of $\xi_2^\nu \xi_3^{M-\nu-\mu} \xi_4^\mu$, one gets the relations,

$$(\mu + 1)\tilde{g}_{\nu-1,\mu+1} = \tilde{f}_{\nu,\mu}, \qquad \nu = 1, \ldots, M, \quad \mu = 0, \ldots, M - \nu. \tag{36}$$

Hence, the coefficients $\{\tilde{g}_{i,j}\}$ of $\hat{G}$ in (34) are now fully determined (see also remark 4.4). Similarly, a comparison of the coefficients of $\xi_1^\nu \xi_3^{M-\nu-\mu} \xi_4^\mu$ leads to the following family of linear systems ($0 \leqslant \nu \leqslant M - 1$ is the parameter):

$$\begin{aligned} &-(2\nu + \mu + 1)\hat{g}_{\nu+1,\mu-1} - (\mu + 1)\hat{g}_{\nu+1,\mu+1} = \hat{f}_{\nu,\mu}, \\ &-(\nu + M)\hat{g}_{\nu+1,M-\nu-2} = \hat{f}_{\nu,M-\nu-1}, \\ &-(\nu + M + 1)\hat{g}_{\nu+1,M-\nu-1} = \hat{f}_{\nu,M-\nu} \end{aligned} \tag{37}$$

with $\mu = 1, \ldots, M - \nu - 2$ in the first equation (see remark 4.4 to understand this system for $\nu = M - 2$ and $\nu = M - 1$). Linear systems (37) are easily solved by backwards substitution with respect to μ and they determine all the coefficients $\{\hat{g}_{i,j}\}$ of $\hat{G}$ in (34). In particular, for a fixed ν, the equation corresponding to $\mu = 2$ gives the solution for $\hat{g}_{\nu+1,1}$, so that the term $\hat{g}_{\nu+1,1}\xi_1^\nu \xi_3^{M-\nu}$ inside the square brackets of (35) cannot be used to remove $\hat{f}_{\nu,0}\xi_1^\nu \xi_3^{M-\nu}$ from $\hat{F}$. Thus, these terms have to be adjusted by taking a non-zero $\hat{Z}$. From here it is very simple to identify the non-removable terms, which are those given by

$$\hat{Z} := \sum_{i=0}^{M} \hat{Z}_i \xi_1^i \xi_3^{M-i} = \sum_{i=0}^{M} \hat{f}_{i,0} \xi_1^i \xi_3^{M-i} + \sum_{\nu=0}^{M-1} \hat{g}_{\nu+1,1} \xi_1^\nu \xi_3^{M-\nu}. \tag{38}$$

Finally, if we consider a generic $F^0 \in \mathbb{E}_s^0$ (with s even), then we express it as

$$\begin{aligned} F^0 &= \sum_{l=0}^{s/2} I^l \hat{F}_{s/2-l} \\ &= \sum_{l=0}^{s/2} \sum_{m=0}^{s/2-l} \sum_{n=0}^{s/2-l} f_{0,l,m,s/2-l-m,s/2-l-n,n} I^l q_1^m q_2^{s/2-l-m} p_1^{s/2-l-n} p_2^n \end{aligned} \tag{39}$$

with $\hat{F}_{s/2-l} \in \mathbb{E}_{0,0,s/2-l,s/2-l}$. Then, we solve $L\hat{G}_{s/2-l} + \hat{Z}_{s/2-l} = \hat{F}_{s/2-l}$ in $\mathbb{E}_{0,0,s/2-l,s/2-l}$ as described above, and we get

$$G^0 = \sum_{l=0}^{s/2} I^l \hat{G}_{s/2-l}, \qquad Z = \sum_{l=0}^{s/2} I^l \hat{Z}_{s/2-l}, \tag{40}$$

solving the projection of (16) into $\mathbb{E}_s^0$. Furthermore, it is clear that if F^0 satisfies the $\mathcal{S}$-symmetries, then it is also true for the above-defined G^0 and Z.

5. Bounds for the solution of the homological equations

At this point we know in a very precise way not only the structure of the normal form but also an effective method to compute the generating function of the normalizing transformation, according to definition 3.1 and proposition 3.3. Now it is time to start the quantitative part of the process. The first step is, of course, to bound the solutions for G_s and Z_s of equation (16) in terms of F_s. This will be done by using the weighted norms introduced in section 2. The result we have obtained is stated as follows.

Proposition 5.1. *With the same hypotheses of proposition 4.1, let* $F_s \in \mathbb{E}_s^{\mathcal{S}}(\rho)$ *for certain* $\rho > 0$ *(see (10)) and consider equation (16). Then, the solutions for* G_s *and* Z_s *given by proposition 4.1 can be chosen (as constructed in section 4) such that* $G_s \in \mathbb{E}_s^{\mathcal{S}}(\rho)$ *with the following bounds (recall* $Z_s = 0$ *if s is odd),*

$$\|G_s\|_\rho \leqslant \frac{2^s s!}{\Omega_s^{s+1}} \|F_s\|_\rho, \qquad \|Z_s\| \leqslant 2^{s/2} \|F_s\|_\rho, \tag{41}$$

where

$$\Omega_s := \min_{\substack{k \in \mathbb{Z}^2 \setminus \{(0,0)\} \\ |k_2| \leqslant s}} \{|\langle k, \omega\rangle|, 1\}. \tag{42}$$

Remark 5.2. Let us assume for a moment that ω also verifies the Diophantine condition (3), so that $\Omega_s \geqslant \tilde{\gamma} s^{-\tau}$, $s \geqslant 1$, for certain $0 < \tilde{\gamma} \leqslant 1$ (see lemma A.7). This estimate on Ω_s is very pessimistic for a lot of values of s, but we can think that 'morally' it is 'quite sharp' for an infinite number of them (maybe by decreasing the exponent τ). Given a generic $F \in \mathbb{E}^{\mathcal{S}}(\rho_0, R_0)$ (see (9)) for certain $\rho_0, R_0 > 0$, we can formally solve equation $L_{H_2} G + Z = F$, order by order, obtaining the bound (41) for the terms of degree s. However, from this bound and the corresponding one of Ω_s, one cannot ensure the convergence of $G = \sum_s G_s$ and hence cannot prove that $G \in \mathbb{E}^{\mathcal{S}}(\rho_1, R_1)$ for any $0 < \rho_1 \leqslant \rho_0$ and $0 < R_1 \leqslant R_0$. This makes a strong difference between the semi-simple and non-semi-simple case, where the corresponding bounds of G_s in terms of F_s, combined with the Diophantine condition, lead to the convergence of G.

The proof of proposition 5.1 has been divided into two subsections according to decomposition (21) of F_s and G_s. In section 5.1 we will bound G_s^+ in terms of F_s^+ and in section 5.2 we will bound G_s^0 and Z_s in terms of F_s^0 (we observe that $\|F_s\|_\rho = \|F_s^+\|_\rho + \|F_s^0\|$ and the same for G_s).

Again, as done in section 4, we will skip the subscript s everywhere to simplify the notation.

5.1. Bounds for G^+

We start considering the projection of equation (16) into $\mathbb{E}_{k,l,M,N} \subset \mathbb{E}_s^+$, with $2l + M + N = s$ and $\Omega \equiv \Omega_{k,M,N} \neq 0$ (see (25)). Hence, what we have to do is to solve (27). For this purpose, we compute explicitly the inverse of the $(M+1)(N+1)$-square matrix $\Lambda \equiv \Lambda_{k,M,N}$ associated to this system. This inverse can be written block-wise as

$$\Lambda^{-1} = \begin{pmatrix} \binom{M}{0}\mathcal{D}_1 & & & & \\ \binom{M}{1}\mathcal{D}_2 & \binom{M-1}{0}\mathcal{D}_1 & & & \\ \binom{M}{2}\mathcal{D}_3 & \binom{M-1}{1}\mathcal{D}_2 & \binom{M-2}{0}\mathcal{D}_1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \binom{M}{M}\mathcal{D}_{M+1} & \binom{M-1}{M-1}\mathcal{D}_M & \binom{M-2}{M-2}\mathcal{D}_{M-1} & \cdots & \binom{0}{0}\mathcal{D}_1 \end{pmatrix}, \tag{43}$$

where we have introduced $\mathcal{D}_\nu = (-1)^{\nu+1}(\nu-1)!D_N^{-\nu}$. To compute the powers $D_N^{-\nu}$ (for $\nu = 1, \ldots, M+1$) we use the definition of D_N and the binomial formula:

$$\begin{aligned} D_N^{-\nu} &= (\Omega \cdot Id_{N+1} - P_N)^{-\nu} \\ &= \frac{1}{\Omega^\nu}\left(Id_{N+1} - \frac{1}{\Omega}P_N\right)^{-\nu} = \sum_{j=0}^{N} \frac{(-1)^j}{\Omega^{\nu+j}} \binom{-\nu}{j} P_N^j, \end{aligned} \tag{44}$$

where we have used that the matrix P_N (see (28)) is $(N+1)$-nilpotent. More concretely, direct computation of the powers P_N^j, for $j = 1, \ldots, N$, yields

$$P_N^j = \begin{pmatrix} 0 & & & & & & & \\ 0 & 0 & & & & & & \\ \vdots & \vdots & \ddots & & & & & \\ j!\binom{N}{j} & 0 & \cdots & 0 & & & & \\ 0 & j!\binom{N-1}{j} & \cdots & 0 & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & j!\binom{j}{j} & 0 & \cdots & 0 \end{pmatrix}.$$

Hence, the only coefficients different from zero of P_N^j are those in a sub-diagonal starting at the $(j+1)$th row. Moreover,

$$\binom{-\nu}{j} = (-1)^j \frac{\nu(\nu+1)\cdots(\nu+j-1)}{j!} = (-1)^j \binom{\nu+j-1}{j}.$$

Thus, defining

$$a_j \equiv a_j(\nu, \Omega) := \frac{j!}{\Omega^{\nu+j}} \binom{\nu+j-1}{j}, \qquad j = 0, \ldots, N$$

and by substitution in (44), one obtains an explicit expression for these matrices

$$D_N^{-\nu} = \begin{pmatrix} \binom{N}{0}a_0 & & & & \\ \binom{N}{1}a_1 & \binom{N-1}{0}a_0 & & & \\ \binom{N}{2}a_2 & \binom{N-1}{1}a_1 & \binom{N-2}{0}a_0 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \binom{N}{N}a_N & \binom{N-1}{N-1}a_{N-1} & \binom{N-2}{N-2}a_{N-2} & \cdots & \binom{0}{0}a_0 \end{pmatrix}.$$

The next lemma furnishes an estimate on the absolute norm of Λ^{-1} (see (43)).

Lemma 5.3. *With the notations and definitions above, the following inequality holds*

$$|\Lambda^{-1}|_1 \leqslant \left(1 + \frac{1}{|\Omega|}\right)^{M+N} \frac{(M+N)!}{|\Omega|}. \tag{45}$$

Proof. From the definition of the norm $|\cdot|_1$ for matrices, one has to look at the column of Λ^{-1} with the biggest absolute norm, and it is clear that it is the first one. Then,

$$
\begin{aligned}
|\Lambda^{-1}|_1 &= \sum_{\nu=0}^{M} \binom{M}{\nu} |\mathcal{D}_{\nu+1}|_1 = \sum_{\nu=0}^{M} \nu! \binom{M}{\nu} |D_N^{-(\nu+1)}|_1 \\
&= \sum_{\nu=0}^{M} \nu! \binom{M}{\nu} \sum_{j=0}^{N} \binom{N}{j} |a_j(\nu+1,\Omega)| \\
&= \sum_{\nu=0}^{M}\sum_{j=0}^{N} \nu! j! \binom{M}{\nu}\binom{N}{j}\binom{\nu+j}{j} \frac{1}{|\Omega|^{\nu+j+1}} \\
&\leqslant \frac{(M+N)!}{|\Omega|} \left(\sum_{\nu=0}^{M} \binom{M}{\nu} \frac{1}{|\Omega|^{\nu}}\right) \left(\sum_{j=0}^{N} \binom{N}{j} \frac{1}{|\Omega|^{j}}\right) \\
&\leqslant \frac{(M+N)!}{|\Omega|} \left(1+\frac{1}{|\Omega|}\right)^{M+N},
\end{aligned}
$$

where it has been used that, if $0 \leqslant \nu \leqslant M$ and $0 \leqslant j \leqslant N$,

$$\nu! j! \binom{\nu+j}{j} \leqslant (M+N)!.$$

This ends the proof of lemma 5.3. □

Remark 5.4. We observe that bound (45) is 'quite sharp' in the following sense: the most 'dangerous term' coming from (45) is the power $M+N+1$ of the small divisor $1/|\Omega|$, and the coefficient of this power cannot be improved because the sum given by the expression of $|\Lambda^{-1}|_1$ contains a summand of the form $(M+N)!/|\Omega|^{M+N+1}$.

We have opted for bounding $|\Lambda^{-1}|_1$ because from the definition of the absolute norm we obtain

$$\|G_{k,l,M,N}\|_\rho \leqslant |\Lambda^{-1}_{k,l,M,N}|_1 \cdot \|F_{k,l,M,N}\|_\rho \leqslant \mu_{k,M,N}\|F_{k,l,M,N}\|_\rho$$

with

$$\mu_{k,M,N} := \left(1+\frac{1}{|\Omega_{k,M,N}|}\right)^{M+N} \frac{(M+N)!}{|\Omega_{k,M,N}|}.$$

Therefore, summation of all those terms of G which are in $\mathbb{E}_s^+$ (see (29)) yields to:

$$
\begin{aligned}
\|G^+\|_\rho &= \sum_{\substack{2l+M+N=s\\ |k|+|M-N|\neq 0}} \|G_{k,l,M,N}\|_\rho \leqslant \sum_{\substack{2l+M+N=s\\ |k|+|M-N|\neq 0}} \mu_{k,M,N}\|F_{k,l,M,N}\|_\rho \\
&\leqslant \left(\max_{\substack{2l+M+N=s\\ |k|+|M-N|\neq 0}} \{\mu_{k,M,N}\}\right)\|F^+\|_\rho \leqslant \frac{2^s s!}{\Omega_s^{s+1}}\|F^+\|_\rho, \qquad (46)
\end{aligned}
$$

where Ω_s is defined in (42).

5.2. *Bounds for G^0 and Z*

Now it is time to bound, in terms of $\|F^0\|$ (see (11) for the definition of the norm), the solutions for the projection of equation (16) into $\mathbb{E}_s^0$, namely G^0 and Z.

We start bounding the solutions in $\mathbb{E}_{0,0,M,M}$ of (30). To do this, we find two main difficulties. On the one hand, we have to bound the solutions for the coefficients $\{\tilde{g}_{i,j}\}$ in (36) and $\{\hat{g}_{i,j}\}$ in (37) in terms of $\{\tilde{f}_{i,j}\}$ and $\{\hat{f}_{i,j}\}$, respectively (these coefficients give $\hat{F}$ and $\hat{G}$ expressed in powers of ξ, see (33) and (34)). On the other hand, we have to translate these bounds into a bound for $\|\hat{G}\|$ in terms of $\|\hat{F}\|$ (recall that $\|\hat{G}\|$ and $\|\hat{F}\|$ are computed from the coefficients of $\hat{G}$ and $\hat{F}$ expressed in powers of (q, p)).

In what concerns $\{\tilde{g}_{i,j}\}$, we have from (36) (recall also that $\tilde{g}_{i,0} = 0$, see remark 4.4),

$$\sum_{i=0}^{M}\sum_{j=0}^{M-i} |\tilde{g}_{i,j}| \leqslant \sum_{i=1}^{M}\sum_{j=0}^{M-i} |\tilde{f}_{i,j}|. \tag{47}$$

The control of the $\{\hat{g}_{i,j}\}$ is more involved. It requires doing some work with the recursive formula defining the solutions of (37).

Lemma 5.5. *For the solutions of (37), we have*

$$\sum_{i=1}^{M}\sum_{j=0}^{M-i} |\hat{g}_{i,j}| \leqslant \frac{1}{2}\sum_{i=0}^{M-1}\sum_{j=1}^{M-i} |\hat{f}_{i,j}|.$$

Proof. We consider system (37) for $0 \leqslant \nu \leqslant M-3$ (the particular cases $\nu = M-2$ and $\nu = M-1$ can be easily discussed, see remark 4.4). Thus, we obtain

$$\hat{g}_{\nu+1,M-\nu-1} = -\frac{1}{\nu+M+1}\hat{f}_{\nu,M-\nu}, \qquad \hat{g}_{\nu+1,M-\nu-2} = -\frac{1}{\nu+M}\hat{f}_{\nu,M-\nu-1},$$

plus the (backwards) recurrence

$$\hat{g}_{\nu+1,\mu-1} = -\frac{1}{2\nu+\mu+1}\hat{f}_{\nu,\mu} - \frac{\mu+1}{2\nu+\mu+1}\hat{g}_{\nu+1,\mu+1}, \qquad \mu = 1, \ldots, M-\nu-2.$$

By taking the complex modulus, we derive the recursive bounds

$$|\hat{g}_{\nu+1,\mu-1}| \leqslant \frac{1}{2\nu+\mu+1}|\hat{f}_{\nu,\mu}| + |\hat{g}_{\nu+1,\mu+1}|, \qquad \mu = 1, \ldots, M-\nu-2.$$

Then, it can be easily checked by induction that the general term of the solution verifies:

$$|\hat{g}_{\nu+1,\mu-1}| \leqslant \sum_{j=0}^{\lfloor (M-\nu-\mu)/2 \rfloor} \frac{1}{2\nu+\mu+2j+1}|\hat{f}_{\nu,\mu+2j}|, \qquad \mu = 1, \ldots, M-\nu.$$

Moreover, we note that these inequalities also work for $\nu = M-2$ and $\nu = M-1$. Now, we sum all those components for $1 \leqslant \mu \leqslant M-\nu$, obtaining

$$\begin{aligned}\sum_{\mu=1}^{M-\nu} |\hat{g}_{\nu+1,\mu-1}| &\leqslant \sum_{\mu=1}^{M-\nu}\sum_{j=0}^{\lfloor (M-\nu-\mu)/2 \rfloor} \frac{1}{2\nu+\mu+2j+1}|\hat{f}_{\nu,\mu+2j}| = \sum_{\sigma=1}^{M-\nu} \frac{\lfloor (\sigma+1)/2 \rfloor}{2\nu+\sigma+1}|\hat{f}_{\nu,\sigma}| \\ &\leqslant \sum_{\sigma=1}^{M-\nu} \frac{\sigma+1}{2(2\nu+\sigma+1)}|\hat{f}_{\nu,\sigma}| \leqslant \frac{1}{2}\sum_{\sigma=1}^{M-\nu} |\hat{f}_{\nu,\sigma}|.\end{aligned}$$

The desired result follows at once by summing these bounds for $0 \leqslant \nu \leqslant M-1$. □

By combining (47) and lemma 5.5, we derive the following bound for the coefficients of $\hat{G}$ in (34) in terms of those of $\hat{F}$ in (33):

$$\sum_{i=1}^{M}\sum_{j=0}^{M-i}|\hat{g}_{i,j}| + \sum_{i=0}^{M}\sum_{j=0}^{M-i}|\tilde{g}_{i,j}| \leqslant \sum_{i=0}^{M}\sum_{j=0}^{M-i}|\hat{f}_{i,j}| + \sum_{i=1}^{M}\sum_{j=0}^{M-i}|\tilde{f}_{i,j}|. \tag{48}$$

Moreover, using lemma 5.5 we also can control the coefficients of the non-removable terms $\hat{Z}$ given in (38), obtaining that the same bound holds,

$$\sum_{i=0}^{M}|\hat{Z}_i| \leqslant \sum_{i=0}^{M}|\hat{f}_{i,0}| + \sum_{\nu=0}^{M-1}|\hat{g}_{\nu+1,1}| \leqslant \sum_{i=0}^{M}\sum_{j=0}^{M-i}|\hat{f}_{i,j}| + \sum_{i=1}^{M}\sum_{j=0}^{M-i}|\tilde{f}_{i,j}|. \tag{49}$$

The next step is to translate inequalities (48) and (49) into bounds for $\|\hat{G}\|$ and $\|\hat{Z}\|$ as a function of $\|\hat{F}\|$. It forces us to control the changes of variables that relate the coefficients of these functions in powers of (q, p) with the ones in powers of ξ. It is done in lemma A.4, and we obtain

$$\|\hat{G}\| \leqslant 2^M\|\hat{F}\|, \qquad \|\hat{Z}\| \leqslant 2^M\|\hat{F}\|.$$

To end this section we only have to consider the expressions of F^0, G^0 and Z given in (39) and (40), thus obtaining

$$\|G^0\| = \sum_{l=0}^{s/2}\|\hat{G}_{s/2-l}\| \leqslant \sum_{l=0}^{s/2}2^{s/2-l}\|\hat{F}_{s/2-l}\| \leqslant 2^{s/2}\sum_{l=0}^{s/2}\|\hat{F}_{s/2-l}\| = 2^{s/2}\|F^0\|.$$

Analogously, we derive the bound $\|Z\| \leqslant 2^{s/2}\|F^0\|$. Now, by joining these bounds with (46) the proof of proposition 5.1 follows straightforwardly.

6. Bounds for the normalizing process

In this section we resume, from the quantitative point of view, the normalizing process for the Hamiltonian H given in (19). The recursive formulae of proposition 3.3 provide us with a closed algorithm to compute G_s and Z_s as a function of the coefficients of the Hamiltonian H. Moreover, proposition 5.1 gives bounds for the solutions G_s and Z_s of the homological equations (16) in terms of the bounds on F_s which, in view of the recursive formula (15), can be computed from those on the previous solutions.

The target of the present section is to convert those recursive formulae into bounds for G_s and Z_s, for $s \geqslant 3$, depending only on initial data H and on s.

Proposition 6.1. *We consider the Hamiltonian H given in (19), with H_2 given in (20). We suppose that $H \in \mathbb{E}^S(\rho_0, R_0)$ for certain $\rho_0 > 0$ and $R_0 > 0$, with $\|H\|_{\rho_0,R_0} \leqslant cR_0^2$, being $c \geqslant 1$. Moreover, we assume that $\omega_1/\omega_2 \notin \mathbb{Q}$. We compute the functions G_s and Z_s of the normalizing process introduced in proposition 3.3, for $s \geqslant 3$, where equation (16) has been solved as stated in proposition 4.1, according to the constructive method described in section 4 and thus with the bounds given by proposition 5.1. Then, there exists $\lambda > 1$ (universal constant) such that,*

$$\|G_s\|_{3\rho_0/4} \leqslant \frac{1}{2}(s-1)!\beta_3\cdots\beta_s\frac{\lambda^{s-3}c^{s-2}\Delta_s^{s-3}}{R_0^{s-2}},$$

$$\|Z_s\| \leqslant \frac{1}{2}(s-1)!\beta_3\cdots\beta_{s-1}\frac{2^{s/2}\lambda^{s-3}c^{s-2}\Delta_s^{s-3}}{R_0^{s-2}}$$

for any $s \geqslant 3$, with

$$\beta_j := \frac{2^j j!}{\Omega_j^{j+1}}, \quad \Delta_j := 4 + \frac{4}{\mathrm{e}\rho_0}\sum_{l=3}^{j-1}\frac{1}{l}, \qquad j \geqslant 3, \tag{50}$$

where Ω_j is defined in (42).

The proof of proposition 6.1 is a straightforward consequence of lemmas 6.2 and A.5. First, lemma 6.2 gives more precise quantitative information on the objects involved in the normal form algorithm of proposition 3.3.

Lemma 6.2. *With the same hypotheses of proposition 6.1, let us consider a fixed integer $s \geqslant 3$. From this value of s we introduce*

$$\mu_s := \frac{\rho_0}{4\sum_{l=3}^{s-1} 1/l}$$

and define the s-depending quantities $\delta_j^{(s)} := \mu_s/j$, $j = 3, \ldots, s-1$. Then, we have the following bounds for the algorithm given in proposition 3.3,

$$\|H_{l,m}\|_{\rho_0-\delta_3^{(s)}-\cdots-\delta_{m+2}^{(s)}} \leqslant b_{l,m}\frac{(l+m-1)!}{(l-1)!}\beta_3\cdots\beta_{m+2}\frac{3^m c^{m+1}\Delta_s^m}{R_0^{l+m-2}} \tag{51}$$

for any $l \geqslant 3$ and $m \geqslant 0$, with $3 \leqslant l+m \leqslant s$, and for any $k \geqslant 3$ we have

$$\|F_k\|_{\rho_0-\delta_3^{(s)}-\cdots-\delta_{k-1}^{(s)}} \leqslant \frac{a_k}{2}(k-1)!\beta_3\cdots\beta_{k-1}\frac{3^{k-3}c^{k-2}\Delta_s^{k-3}}{R_0^{k-2}}, \tag{52}$$

$$\|G_k\|_{\rho_0-\delta_3^{(s)}-\cdots-\delta_{k-1}^{(s)}} \leqslant \frac{a_k}{2}(k-1)!\beta_3\cdots\beta_{k}\frac{3^{k-3}c^{k-2}\Delta_s^{k-3}}{R_0^{k-2}}, \tag{53}$$

$$\|Z_k\| \leqslant \frac{a_k}{2}(k-1)!\beta_3\cdots\beta_{k-1}\frac{2^{k/2}3^{k-3}c^{k-2}\Delta_s^{k-3}}{R_0^{k-2}}, \tag{54}$$

where β_j, $j = 3, \ldots, s$ and Δ_s are those in (50). The quantities $\{a_k\}$ and $\{b_{l,m}\}$ are defined through the following recurrences:

$$a_k = \frac{2^{3/2}}{6(k-2)(k-1)!}\sum_{j=1}^{k-3} j(j+2)!(k-j)!a_{2+j}a_{k-j} + \frac{2}{k-2}\sum_{j=1}^{k-2}\frac{j}{(j+1)!}\frac{b_{2+j,k-j-2}}{3^{j-1}}, \tag{55}$$

$$b_{l,m} = \frac{1}{6m(l+m-1)!}\sum_{j=1}^{m} j(j+2)!(l+m-j)!a_{2+j}b_{l,m-j} \tag{56}$$

with $a_3 = 1$ and $b_{l,0} = 1$ for any $l \geqslant 3$.

Remark 6.3. Again, several 'abuses of notation' have been done in the statement of lemma 6.2. First, the definition of μ_s has no sense for $s = 3$, but in this case it plays no role. Moreover, we have to understand that $\rho_0 - \delta_3^{(s)} - \cdots - \delta_2^{(s)}$ equals to ρ_0 and $\beta_3\cdots\beta_2 = 1$. Finally, we recall that $Z_s = 0$ if s is odd, but we have not taken advantage of this fact.

On its turn, lemma A.5 gives a geometrical bound for the quantities $\{a_k\}$ and $\{b_{l,m}\}$, so that $a_k \leqslant \tilde{\lambda}^{k-3}$ and $b_{l,m} \leqslant \tilde{\lambda}^m$, for certain $\tilde{\lambda} > 1$ (see remark A.6 for a numerical value of $\tilde{\lambda}$). From here, we deduce proposition 6.1 in an easy way, by taking $\lambda := 3\tilde{\lambda}$ (we observe that $\delta_3^{(s)} + \cdots + \delta_{s-1}^{(s)} = \rho_0/4$ in (53)).

Next to that, we prove lemma 6.2. The precise statement and proof of lemma A.5 is given in appendix A.

Proof of lemma 6.2. We keep fixed the value of $s \geqslant 3$ along the proof, so that we can remove the s-dependence from the notation ($\mu \equiv \mu_s$, $\delta_j \equiv \delta_j^{(s)}$, $\Delta_s \equiv \Delta$ and so on). Moreover, we consider the following simplification for the notation of the domains. We define:

$$\sigma_3 := \rho_0, \qquad \sigma_j := \rho_0 - \delta_3 - \cdots - \delta_j, \quad j = 4, \ldots, s-1.$$

Now let us clarify the role of μ, Δ and of $\{\delta_j\}$ in the proof. Let us take $f_l \in \mathbb{E}_l(\rho + \delta_l)$ and $f_m \in \mathbb{E}_m(\rho + \delta_m)$, for certain $\rho > 0$ and $l, m \in \{3, \ldots, s-1\}$. Then, we have (see (7)),

$$\|\{f_l, f_m\}\|_\rho \leqslant \left(\frac{m}{2\mathrm{e}\delta_l} + \frac{l}{2\mathrm{e}\delta_m} + 4lm\right)\|f_l\|_{\rho+\delta_l}\|f_m\|_{\rho+\delta_m} = lm\Delta\|f_l\|_{\rho+\delta_l}\|f_m\|_{\rho+\delta_m}, \tag{57}$$

where we have used lemma A.1 to bound the θ-derivatives and the norm of the product. Hence, by choosing the values for $\{\delta_j\}$ of the statement, we have that Δ acts as a homogenization factor for the norm of the Poisson bracket with respect to the degree of f_l and f_m. Finally, as mentioned in remark 6.3, μ is selected so that $\sigma_{s-1} = 3\rho_0/4$.

The last observation before starting the proof refers to the bound on H. The hypothesis $\|H\|_{\rho_0,R_0} \leqslant cR_0^2$ and the definition $H_{l,0} = H_l$ imply (recall we are using a weighted norm),

$$\|H_{l,0}\|_{\rho_0} = \|H_l\|_{\rho_0} \leqslant \frac{c}{R_0^{l-2}}, \qquad l \geqslant 3. \tag{58}$$

We perform the proof by induction on the following property:

$\mathcal{P}_\mu := \{$(51) holds for any $3 \leqslant l + m \leqslant \mu, m \geqslant 0$;
(52), (53) and (54) hold for any $3 \leqslant k \leqslant \mu\}$.

For $\mu = 3$ it is clear that $\mathcal{P}_3$ is true. Indeed, $H_{3,0} = H_3$, $F_3 = H_{3,0}$ and $L_{H_2}G_3 + Z_3 = F_3$. Thus,

$$\|H_{3,0}\|_{\rho_0} \leqslant \frac{c}{R_0}, \qquad \|F_3\|_{\rho_0} \leqslant \frac{c}{R_0}$$

and by applying proposition 5.1 we obtain

$$\|Z_3\| = 0, \qquad \|G_3\|_{\rho_0} \leqslant \beta_3\|F_3\|_{\rho_0} \leqslant \beta_3\frac{c}{R_0}.$$

From here, $\mathcal{P}_3$ follows.

Now we suppose $\mathcal{P}_{k-1}$ true, for certain $k \geqslant 4$, and we want to check that $\mathcal{P}_k$ also holds. We start bounding $H_{l,m}$, with $l + m = k$ ($l \geqslant 3$, $m \geqslant 0$). The case $l = k$ and $m = 0$ is clear from (58). Hence, we suppose $m \geqslant 1$ and we use formula (13) on $H_{l,m}$, thus obtaining (see (57)),

$$\|H_{l,m}\|_{\sigma_{m+2}} \leqslant \sum_{j=1}^{m}\frac{j}{m}(j+2)(l+m-j)\Delta\|G_{2+j}\|_{\sigma_{m+2}+\delta_{2+j}}\|H_{l,m-j}\|_{\sigma_{m+2}+\delta_{l+m-j}}. \tag{59}$$

In order to use the inductive hypothesis $\mathcal{P}_{k-1}$ on (59), we have to check that the domains where we want to control the norms of G_{2+j} and $H_{l,m-j}$, for $j = 1, \ldots, m$, are contained inside the domains for which we have the bounds given by $\mathcal{P}_{k-1}$. For G_{2+j} we need,

$$\sigma_{m+2} + \delta_{j+2} \leqslant \sigma_{j+1} \iff \delta_3 + \cdots + \delta_{j+2} \leqslant \delta_3 + \cdots + \delta_{m+2}.$$

It holds if $1 \leqslant j \leqslant m$. For $H_{l,m-j}$ the condition is

$$\sigma_{m+2} + \delta_{l+m-j} \leqslant \sigma_{m-j+2} \iff \delta_3 + \cdots + \delta_{m-j+2} + \delta_{l+m-j} \leqslant \delta_3 + \cdots + \delta_{m+2}.$$

The worst case for l is clearly when $l = 3$ (because δ_l decreases with l), so that we have to check

$$\delta_3 + \cdots + \delta_{m-j+2} + \delta_{m-j+3} \leqslant \delta_3 + \cdots + \delta_{m+2}.$$

It is obviously true if $1 \leqslant j \leqslant m$. Thus,

$$\|H_{l,m}\|_{\sigma_{m+2}} \leqslant \sum_{j=1}^{m} \frac{j}{m}(j+2)(l+m-j)\Delta \frac{a_{2+j}}{2}(j+1)!\beta_3\cdots\beta_{2+j}\frac{3^{j-1}c^j\Delta^{j-1}}{R_0^j} \times b_{l,m-j}\frac{(l+m-j-1)!}{(l-1)!}\beta_3\cdots\beta_{m-j+2}\frac{3^{m-j}c^{m-j+1}\Delta^{m-j}}{R_0^{l+m-j-2}} \leqslant b_{l,m}\frac{(l+m-1)!}{(l-1)!}\beta_3\cdots\beta_{m+2}\frac{3^m c^{m+1}\Delta^m}{R_0^{l+m-2}},$$

where we have used that if $1 \leqslant j \leqslant m$ then $\beta_3\cdots\beta_{2+j}\beta_3\cdots\beta_{m-j+2} \leqslant \beta_3\cdots\beta_{m+2}$ (we observe that β_j is increasing as a function of j, see (50)) and the inductive definition of $b_{l,m}$ in (56). Now it is the turn of F_k. Formula (15) gives

$$\|F_k\|_{\sigma_{k-1}} \leqslant \sum_{j=1}^{k-3}\frac{j(2+j)(k-j)}{k-2}\Delta\|G_{2+j}\|_{\sigma_{k-1}+\delta_{2+j}}\|Z_{k-j}\| + \sum_{j=1}^{k-2}\frac{j}{k-2}\|H_{2+j,k-j-2}\|_{\sigma_{k-1}}.$$

Again, to use the inductive bounds $\mathcal{P}_{k-1}$ we require certain compatibility conditions on the domains where the norms are evaluated:

$$\sigma_{k-1}+\delta_{2+j} \leqslant \sigma_{1+j}, \qquad j=1,\ldots,k-3,$$
$$\sigma_{k-1} \leqslant \sigma_{k-j}, \qquad j=1,\ldots,k-2,$$

which are not difficult to verify. Then, we have

$$\|F_k\|_{\sigma_{k-1}} \leqslant \sum_{j=1}^{k-3}\frac{j}{k-2}(2+j)(k-j)\Delta\frac{a_{2+j}}{2}(j+1)!\beta_3\cdots\beta_{2+j}\frac{3^{j-1}c^j\Delta^{j-1}}{R_0^j} \times \frac{a_{k-j}}{2}(k-j-1)!\beta_3\cdots\beta_{k-j-1}\frac{2^{(k-j)/2}3^{k-j-3}c^{k-j-2}\Delta^{k-j-3}}{R_0^{k-j-2}} + \sum_{j=1}^{k-2}\frac{j}{k-2}b_{2+j,k-j-2}\frac{(k-1)!}{(j+1)!}\beta_3\cdots\beta_{k-j}\frac{3^{k-j-2}c^{k-j-1}\Delta^{k-j-2}}{R_0^{k-2}} \leqslant \frac{a_k}{2}(k-1)!\beta_3\cdots\beta_{k-1}\frac{3^{k-3}c^{k-2}\Delta^{k-3}}{R_0^{k-2}}.$$

Here we have used, in the first sum, that $\beta_3\cdots\beta_{2+j}\beta_3\cdots\beta_{k-j-1}2^{(k-j)/2} \leqslant \beta_3\cdots\beta_{k-1}2^{3/2}$, for $1 \leqslant j \leqslant k-3$. Indeed, the case $j=k-3$ is clear and the other ones are easily proved by observing that $\beta_j\sqrt{2} \leqslant \beta_{j+1}$ for any $j \geqslant 3$ (see (50)), and hence $\beta_3 2^{(k-j-3)/2} \leqslant \beta_{k-j}$. The increasing character of the sequence $\{\beta_j\}$ ends the proof. Moreover, in the second sum, we have used that $\Delta > 1$, $c \geqslant 1$ and $\beta_3\cdots\beta_{k-j} \leqslant \beta_3\cdots\beta_{k-1}$ for $1 \leqslant j \leqslant k-2$. Then, we use the definition of a_k in (55). Finally, the bounds on $\|G_k\|_{\sigma_{k-1}}$ and $\|Z_k\|$ are a straightforward application of proposition 5.1. □

7. Bounds for the remainder of the normal form

We compute a finite order normalizing process for the (complexified) Hamiltonian H of (19). Thus, we have a generating function $G^{(r)} = G_3 + \cdots + G_r$, for a certain fixed order (degree) $r \geqslant 3$, and the corresponding finite order normal form $Z^{(r)} = Z_2 + \cdots + Z_r$, with $Z_2 = H_2$, according to the algorithm of proposition 3.3.

Now, we consider the canonical transformation $T_{G^{(r)}}H$ (see definition 3.1) or equivalently $H \circ \Phi_{\varepsilon=1}^{-\chi^{(r)}}$, with $\chi^{(r)} = \sum_{s=1}^{r-2} sG_{s+2}\varepsilon^{s-1}$ (see remark 3.2), then

$$(T_{G^{(r)}}H)(\theta, q, I, p) = H \circ \Phi_{\varepsilon=1}^{-\chi^{(r)}}(\theta, q, I, p) = Z^{(r)}(q, I, p) + \mathcal{R}^{(r)}(\theta, q, I, p), \tag{60}$$

where from proposition 4.1 we have that $Z^{(r)}$ takes the form

$$Z^{(r)}(q, I, p) = \omega_1 I + \mathrm{i}\omega_2(q_1p_1 + q_2p_2) + q_2p_1 + \mathcal{Z}^{(r)}\left(I, q_1p_2, \frac{\mathrm{i}(q_1p_1 + q_2p_2)}{2}\right), \tag{61}$$

where $\mathcal{Z}^{(r)}(u, v, w)$ is a polynomial of (standard) degree less than or equal to $\lfloor r/2\rfloor$ in (u, v, w), starting at degree two.

By applying proposition 6.1 we can derive bounds for G_s, $3 \leqslant s \leqslant r$, as a function of s and of some initial data on H. The purpose of this section is to use those bounds on G_s to estimate the size of the remainder $\mathcal{R}^{(r)}$ as a function of r and of the distance to the resonant orbit.

Therefore, before doing that let us return to *real* variables. We recall that the complexified Hamiltonian H of (19) has been obtained from $\mathcal{H}$ in (1), with $\epsilon = +1$, by replacing the variables (x, y) by (q, p) through the canonical change φ of (18). As the $\mathcal{S}$-symmetries induced by the complexification (see section 2) have been preserved by the normal from algorithm of section 4 (see proposition 4.1), we can go back to real variables by means of φ^{-1} (see remark 3.4). Hence, equation (60) is converted into the real expression

$$H \circ \Phi_{\varepsilon=1}^{-(\chi^{(r)}\circ\varphi^{-1})}(\theta, x, I, y) = Z^{(r)} \circ \varphi^{-1} + \mathcal{R}^{(r)} \circ \varphi^{-1}. \tag{62}$$

Then, a bound for $\mathcal{R}^{(r)} \circ \varphi^{-1}$ is given by the next result.

Proposition 7.1. *With the same hypotheses of proposition 6.1 and the notation above, we also assume that the vector ω satisfies the Diophantine condition (3). Given any $\sigma > 1$ and $0 < \varepsilon'' < \varepsilon'$, all fixed, then there exists $\tilde{d} > 0$, depending on ρ_0, c, $|\omega_1|$, $|\omega_2|$, γ, τ, ε' and ε'', such that if we define,*

$$\sigma'' = \min\{1 - \sigma^{-1}, (1 - \sigma^{-2})\rho_0\}, \qquad \sigma' = \frac{\sigma''}{\tilde{d} + \sigma''}, \qquad \hat{R}^{(r)} = \frac{\sigma^{-1}\sigma' R_0}{(r-2)^{(\tau+1+\varepsilon')(r-2)/2}},$$

then, for any $r \geqslant 3$ and $0 < R \leqslant \hat{R}^{(r)}$,

(a) The canonical transformation $\Phi_{\varepsilon=1}^{-(\chi^{(r)}\circ\varphi^{-1})}$ is a real analytic diffeomorphism defined in the domain $\mathcal{D}(\sigma^{-2}\rho_0/2, \hat{R}^{(r)})$ (see (2)), with $\Phi_{\varepsilon=1}^{-(\chi^{(r)}\circ\varphi^{-1})}(\mathcal{D}(\sigma^{-2}\rho_0/2, R)) \subset \mathcal{D}(\rho_0/2, \sigma R)$.

(b) If we set $\Phi_{\varepsilon=1}^{-(\chi^{(r)}\circ\varphi^{-1})} - Id := (\Theta^{(r)}, \mathcal{X}^{(r)}, \mathcal{I}^{(r)}, \mathcal{Y}^{(r)})$, then we have $\Theta^{(r)}, \mathcal{X}_j^{(r)}, \mathcal{I}^{(r)}, \mathcal{Y}_j^{(r)} \in \mathbb{E}(\sigma^{-2}\rho_0/2, \hat{R}^{(r)})$, $j = 1, 2$, verifying

$$\|\Theta^{(r)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant \frac{(1-\sigma^{-2})\rho_0}{2}, \qquad \|\mathcal{I}^{(r)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant (\sigma^2 - 1)R^2,$$

$$\|\mathcal{X}_j^{(r)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant (\sigma - 1)R, \qquad \|\mathcal{Y}_j^{(r)}\|_{\sigma^{-2}\rho_0/2,R} \leqslant (\sigma - 1)R.$$

(c) The following bounds hold for the normal form and for the remainder:

$$\|Z^{(r)} \circ \varphi^{-1}\|_R \leqslant \|H\|_{\rho_0,R_0} \qquad \|\mathcal{R}^{(r)} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,R} \leqslant cR_0^2\left(\frac{R}{\hat{R}^{(r)}}\right)^{r+1}.$$

Proof. As the value of $r \geqslant 3$ is held fixed throughout this proof, the superscript '(r)' will be omitted in all the r-dependent expressions.

Now, let us start describing the main trick in the proof (it has also been used in [12]). After we have fixed the order r of the normal form, we can use the following process to compute the remainder: we take $G_s = 0$ and $Z_s = F_s$, for $s > r$, as solutions of (16), and thus $\mathcal{R} = \sum_{s>r} Z_s$, according to the recurrences of proposition 3.3. Unfortunately, the bounds on those recurrences provided by lemma 6.2 only apply if we consider the solution of (16) given by proposition 4.1, but not if we set $G_s = 0$. Of course, we have the chance of still using the recurrences of definition 3.1 and proposition 3.3 to bound the remainder, by adapting lemma 6.2 (see, for instance [7] for an example of this and [18] for the application of this methodology to the present context). However, we have preferred to follow a different method, trying to overcome the tedious estimates on the above mentioned algorithm.

Let us suppose it is proved that the canonical transformation $\Phi_{\varepsilon=1}^{-(\chi\circ\varphi^{-1})}$ acts as described in items (a) and (b) on the statement. Then, by using lemma A.2 and by taking (weighted) norms on (60), we obtain:

$$\begin{aligned}\|H \circ \Phi_{\varepsilon=1}^{-(\chi\circ\varphi^{-1})}\|_{\sigma^{-2}\rho_0/2,\hat{R}} &= \|Z \circ \varphi^{-1}\|_{\hat{R}} + \|\mathcal{R} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,\hat{R}} \\ &\leqslant \|H\|_{\rho_0/2,\sigma\hat{R}} \leqslant \|H\|_{\rho_0,R_0} \leqslant cR_0^2.\end{aligned}$$

This inequality gives bounds on the transformed Hamiltonian without doing estimates on the normal form algorithm. *A priori*, this bound on $\mathcal{R} \circ \varphi^{-1}$ does not look quite small, but we recall that we know by construction that the Taylor expansion on (q, I, p) of the remainder starts at degree $r + 1$. It gives item (c) on the statement. This fact is another of the advantages of working with a weighted norm. Hence, the important point is to prove (a) and (b).

From lemma A.3 we know that the action of $\Phi_{\varepsilon=1}^{-(\chi\circ\varphi^{-1})}$ can be discussed in terms of bounds on the derivatives of $\tilde{\chi}(\theta, x, I, y; \varepsilon) := \chi \circ \varphi^{-1}$. More precisely, we have to control $\tilde{\chi}_1 := \tilde{\chi}(\cdot; 1)$,

$$\begin{aligned}\|\partial_I \tilde{\chi}_1\|_{\rho_0/2,\sigma R} &\leqslant \sum_{s=3}^{r} \frac{s^2}{2} \|G_s \circ \varphi^{-1}\|_{3\rho_0/4} (\sigma R)^{s-2}, \\ \|\partial_\theta \tilde{\chi}_1\|_{\rho_0/2,\sigma R} &\leqslant \sum_{s=3}^{r} \frac{4s}{\mathrm{e}\rho_0} \|G_s \circ \varphi^{-1}\|_{3\rho_0/4} (\sigma R)^{s}, \\ \|\partial_{x_j} \tilde{\chi}_1\|_{\rho_0/2,\sigma R} &\leqslant \sum_{s=3}^{r} s^2 \|G_s \circ \varphi^{-1}\|_{3\rho_0/4} (\sigma R)^{s-1}, \\ \|\partial_{y_j} \tilde{\chi}_1\|_{\rho_0/2,\sigma R} &\leqslant \sum_{s=3}^{r} s^2 \|G_s \circ \varphi^{-1}\|_{3\rho_0/4} (\sigma R)^{s-1}\end{aligned}$$

for $j = 1, 2$. Thus, the key expression to be studied is, for any $0 < R \leqslant \hat{R}$,

$$g(\sigma R) := \sum_{s=3}^{r} s^2 \|G_s \circ \varphi^{-1}\|_{3\rho_0/4} (\sigma R)^{s-2} \leqslant \sum_{s=3}^{r} 2^{s/2} \frac{s}{2} s! \beta_3 \cdots \beta_s \lambda^{s-3} c^{s-2} \Delta_s^{s-3} \left(\frac{\sigma R}{R_0}\right)^{s-2}.$$

To derive this estimate we have used the bounds on $\|G_s\|_{3\rho_0/4}$ coming from proposition 6.1. Furthermore, we observe that the effect of φ^{-1} in the bounds can be controlled in terms of the degree of G_s as $\|G_s \circ \varphi^{-1}\|_{3\rho_0/4} \leqslant 2^{s/2} \|G_s\|_{3\rho_0/4}$. The control of the product $\beta_3 \cdots \beta_s$ is carried out in lemma A.8 (see also lemma A.7). Moreover, we also have (see (50)),

$$\Delta_s = 4 + \frac{4}{\mathrm{e}\rho_0} \sum_{j=3}^{s-1} \frac{1}{j} \leqslant 4 + \frac{4}{\mathrm{e}\rho_0} \int_2^{s-1} \frac{\mathrm{d}x}{x} = 4 + \frac{4}{\mathrm{e}\rho_0} \log\left(\frac{(s-1)}{2}\right).$$

Thus, if $d_{\varepsilon''}$ is the constant provided by lemma A.8 we obtain

$$g(\sigma R) \leqslant \sum_{s=3}^{r} d_{\varepsilon''} 2^{s/2} \frac{s}{2} s! \lambda^{s-3} c^{s-2} \left(4 + \frac{4}{\mathrm{e}\rho_0} \log\left(\frac{s-1}{2}\right)\right)^{s-3} (s+1)^{(\tau+1+\varepsilon'')(s+1)^2/2} \left(\frac{\sigma R}{R_0}\right)^{s-2}$$
$$\leqslant \sum_{s=3}^{r} \tilde{d}(s-2)^{(\tau+1+\varepsilon')(s-2)^2/2} \left(\frac{\sigma R}{R_0}\right)^{s-2},$$

where $\tilde{d}$ is the value asked on the statement, whose existence is clear because $\varepsilon' > \varepsilon''$ (compare with the determination of d_ε in the proof of lemma A.8). Now, we observe that from the definition of $\hat{R}$ we have

$$g(\sigma \hat{R}) \leqslant \sum_{s=3}^{r} \tilde{d}(s-2)^{(\tau+1+\varepsilon')(s-2)^2/2} \left(\frac{\sigma'}{(r-2)^{(\tau+1+\varepsilon')(r-2)/2}}\right)^{s-2}$$
$$\leqslant \sum_{s=3}^{r} \tilde{d}(\sigma')^{s-2} \leqslant \tilde{d}\frac{\sigma'}{1-\sigma'} = \sigma''.$$

Hence, from the definition of σ'' we obtain, for any $0 < R \leqslant \hat{R}$,

$$\max\left\{\frac{2}{\rho_0}\|\partial_I \tilde{\chi}_1\|_{\rho_0/2,\sigma R}, \frac{1}{(\sigma R)^2}\|\partial_\theta \tilde{\chi}_1\|_{\rho_0/2,\sigma R}\right\} \leqslant 1 - \sigma^{-2},$$
$$\max_{j\in\{1,2\}}\left\{\frac{1}{\sigma R}\|\partial_{x_j} \tilde{\chi}_1\|_{\rho_0/2,\sigma R}, \frac{1}{\sigma R}\|\partial_{y_j} \tilde{\chi}_1\|_{\rho_0/2,\sigma R}\right\} \leqslant 1 - \sigma^{-1}.$$

Then, a direct application of lemma A.3 gives (a) and (b), ending the proof. □

Remark 7.2. The bound on the product $\beta_3 \cdots \beta_s$ provided by lemma A.8 is the key one in the paper: it is responsible for the estimate of theorem 1.1 for the size of the remainder of the normal form in terms of the Lambert function (instead of, for instance, an exponentially small one). Thus, it is quite natural to wonder if (A3) can be improved. As discussed in remark A.9, no remarkable improvement can be expected for this estimate.

8. Optimal normalization order

After section 7 we have identified the normal form up to any order and we have also controlled the canonical normal form transformation. Hence, the only thing that remains to do before proving theorem 1.1 is to ask the main question of this paper: we have to discuss, according to the bound for $\mathcal{R}^{(r)}$ given by proposition 7.1, what is the '*optimal*' choice for r as a function of the distance, R, to the resonant periodic orbit, in such a way that the remainder of the normal form becomes as small as possible. The answer to this question is given by the next proposition.

Proposition 8.1. *With the same statement of proposition 7.1, let us consider a fixed* ε, *with* $0 < \varepsilon'' < \varepsilon' < \varepsilon$. *If we select the order of the normal form,* r, *as a function of* R *so that* $r = r_{\mathrm{opt}}(R)$, *with*

$$r_{\mathrm{opt}}(R) := 2 + \left\lfloor \exp\left(W\left(\log\left(\frac{1}{R^{1/(\tau+1+\varepsilon)}}\right)\right)\right)\right\rfloor,$$

where $W(\cdot)$ *verifies* $W(z)\exp(W(z)) = z$, *for any* $z > 0$ *(see remark 1.2), then we obtain the following estimate for the remainder of the (real) normal form (62),*

$$\|\mathcal{R}^{r_{\mathrm{opt}}(R)} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,R} \leqslant R^{r_{\mathrm{opt}}(R)/2}, \qquad 0 < R \leqslant R^*,$$

where R^* *depends on* ρ_0, R_0, c, $|\omega_1|$, $|\omega_2|$, γ, τ, ε, ε', ε'' *and* σ.

Proof. By using the definition of $\hat{R}^{(r)}$ and part (c) of proposition 7.1, we obtain

$$\|\mathcal{R}^{(r)} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,R} \leqslant cR_0^2 \left(\frac{(r-2)^{(\tau+1+\varepsilon')(r-2)/2} R}{\sigma^{-1}\sigma' R_0} \right)^{r+1}, \qquad 0 < R \leqslant \hat{R}^{(r)}.$$

Now, we define r_0 as the first integer such that $(r-2)^{(\varepsilon'-\varepsilon)(r-2)/2} \leqslant \sigma^{-1}\sigma' R_0$, and thus, if $r \geqslant r_0$,

$$\|\mathcal{R}^{(r)} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,R} \leqslant cR_0^2((r-2)^{(\tau+1+\varepsilon)(r-2)/2} R)^{r+1}, \qquad 0 < R \leqslant \hat{R}^{(r)}. \tag{63}$$

Given any $R > 0$ fixed, we define the following function:

$$h(r) := \log(cR_0^2) + \frac{(\tau+1+\varepsilon)(r-2)(r+1)}{2} \log(r-2) + (r+1)\log R,$$

which equals the logarithm of the present bound for $\|\mathcal{R}^{(r)} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,R}$. What we want to do is to minimize (as much as possible) $h(r)$. Thus, we compute its derivative,

$$h'(r) = \frac{(\tau+1+\varepsilon)(2r-1)}{2} \log(r-2) + \frac{(\tau+1+\varepsilon)(r+1)}{2} + \log R$$

and try to look for the value of r such that $h'(r) = 0$. It leads to a complicated expression, and hence, we only consider the dominant part and then

$$(\tau+1+\varepsilon)(r-2)\log(r-2) + \log R = 0$$
$$\iff \log(r-2)\exp(\log(r-2)) = \log\left(\frac{1}{R^{1/(\tau+1+\varepsilon)}}\right),$$

which suggests the choice $r = r_{\mathrm{opt}}(R)$ on the statement. For this (integer) value of r we have $(r-2)^{(\tau+1+\varepsilon)(r-2)} \leqslant 1/R$, giving from (63) the bound for $\|\mathcal{R}^{(r)} \circ \varphi^{-1}\|_{\sigma^{-2}\rho_0/2,R}$ on the statement if R is small enough (we observe that the condition $r_{\mathrm{opt}}(R) \geqslant r_0$ means R small enough). □

Remark 8.2. Now, we can compare proposition 8.1 with the context of an elliptic periodic orbit of a real analytic three-degrees of freedom Hamiltonian (see (5) and text below). Of course, we do not plan to give full details for the elliptic case, but a reader familiar with bounds on normal forms will find no difficulties in filling the gaps. Thus, in the elliptic context we obtain the same bounds for $\|G_s\|_{3\rho_0/4}$ as in proposition 6.1, but by setting in (50)

$$\beta_j^{-1} := \min_{\substack{k \in \mathbb{Z}^3\backslash\{(0,0,0)\} \\ |k_2|+|k_3| \leqslant j}} \{|\langle k, \omega\rangle|, 1\}.$$

If ω is Diophantine, we have an estimate like $\beta_j \leqslant \tilde{\gamma}^{-1} j^\tau$, $j \geqslant 1$ (compare with lemma A.7). A re-formulation of proposition 7.1 using this estimate leads to an expression of the form $\hat{R}^{(r)} = b/r^{\tau+1}$, for certain $b > 0$, from which follows a bound like $\|\mathcal{R}^{(r)}\|_{\rho_0/4,R} \leqslant a(r^{\tau+1}R/b)^{r+1}$, with $a > 0$. Thus, an exponentially small estimate of the form $\|\mathcal{R}^{(r)}\|_{\rho_0/4,R} \leqslant a\exp(-(c/R)^{1/(\tau+1)})$ is obtained for the remainder. This result coincides with the one of [6] for an elliptic fixed point of a Hamiltonian system (obtained also using the Giorgilli–Galgani algorithm), but did not match the exponent $2/(\tau+1)$ of [11], which has been obtained using a different strategy to quantify the normal form, which cannot be applied to the present context (read the reason in remark 5.2).

9. Proof of theorem 1.1

The last issue that remains is to prove theorem 1.1. The proof follows from propositions 7.1 and 8.1 (see also proposition 6.1). We take the *real* Hamiltonian $\mathcal{H}(\theta, x, I, y)$ and consider the *complexified* one $H(\theta, q, I, p) = \mathcal{H} \circ \varphi$ (see (18)). In order to apply the results mentioned above, we set $R_0 := R^{(0)}/\sqrt{2}$ and $c := \max\{\|\mathcal{H}\|_{\rho_0, R^{(0)}}/R_0^2, 1\}$. Hence (see lemma A.2),

$$\|H\|_{\rho_0,R_0} = \|\mathcal{H} \circ \varphi\|_{\rho_0,R_0} \leqslant \|\mathcal{H}\|_{\rho_0,R^{(0)}} \leqslant cR_0^2.$$

Moreover, we choose (for example) $\varepsilon' = \varepsilon/2$ and $\varepsilon'' = \varepsilon'/2$, thus defining R^* from proposition 8.1. Now, if we set $r := r_{\mathrm{opt}}(R)$ as a normalizing order, $\Psi^{(R)} := \Phi_{\varepsilon=1}^{-(\chi^{(r)}\circ\varphi^{-1})}$, $Z^{(R)} := Z^{(r)} \circ \varphi^{-1}$ and $\mathcal{R}^{(R)} := \mathcal{R}^{(r)} \circ \varphi^{-1}$ (see (60)–(62)), we have that items (i)–(v) on the statement of theorem 1.1 follow directly from propositions 7.1 and 8.1. In particular, we observe that (61) gives the desired expression for the normal form if we use (23). Part (vi) is a consequence of (v) and of the definition of $r_{\mathrm{opt}}(R)$.

Acknowledgments

The authors thank Pere Gutiérrez for interesting discussions and comments. The authors are partially supported by the Spanish MCyT/FEDER grant BFM2003-07521-C02-1 and the Catalan CIRIT grant 2001SGR-00070.

Appendix A. Lemmata

In this appendix we include some of the technical results used throughout the paper that we have preferred not to state in the body of the paper. Our idea has been to stress in the main part of the paper the key points of the constructions we present, and to postpone the most technical details till this section. We hope that this presentation helps to add to the readability of the paper.

In the first three results, that are given without proof, we want to survey some basic properties of the weighted norms introduced in section 2. The proof of lemma A.1 is quite direct and the proof of lemma A.2 is not difficult using lemma A.1 (just expand and bound). The proof of lemma A.3 is more delicate, but it is very similar to a result that is well known in terms of the supremum norm (see, for instance [5]). The differences refer to the use of the weighted norm instead of the supremum one when bounding the composition of functions (we have to use lemma A.2) and thus we also omit the proof (just bound the integral expressions given by the solutions of the flow). For full details, see [18].

Lemma A.1. *Let* $f, g \in \mathbb{E}(\rho, R)$*, and* $0 < \delta \leqslant \rho$*. Then we have:*

$$\text{(i) } \|f \cdot g\|_{\rho,R} \leqslant \|f\|_{\rho,R} \cdot \|g\|_{\rho,R}, \qquad \text{(ii) } \|\partial_\theta f\|_{\rho-\delta,R} \leqslant \frac{\|f\|_{\rho,R}}{\delta \mathrm{e}}.$$

Moreover, analogous properties can be generalized to the norms $\|\cdot\|_\rho$ *and* $\|\cdot\|$ *in* $\mathbb{E}_s(\rho)$ *and* $\mathbb{C}_s[q, I, p]$*, respectively.*

Lemma A.2. *Let us take* $0 < \rho_0 < \rho$ *and* $0 < R_0 < R$ *and consider the analytic functions* Θ, $\mathcal{I}$, $\mathcal{Q}_j$ *and* $\mathcal{P}_j$, $j = 1, 2$, *all belonging to* $\mathbb{E}(\rho_0, R_0)$ *and verifying* $\|\Theta\|_{\rho_0,R_0} \leqslant \rho - \rho_0$, $\|\mathcal{I}\|_{\rho_0,R_0} \leqslant R^2$ *and* $\max_{j=1,2}\{\|\mathcal{Q}_j\|_{\rho_0,R_0}, \|\mathcal{P}_j\|_{\rho_0,R_0}\} \leqslant R$*. Then, given any* $f \in \mathbb{E}(\rho, R)$*, the function* $F(\theta, q, I, p)$ *defined by* $F := f(\theta + \Theta, \mathcal{Q}, \mathcal{I}, \mathcal{P})$*, with* $\mathcal{Q} = (\mathcal{Q}_1, \mathcal{Q}_2)$ *and* $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$*, is such that* $F \in \mathbb{E}(\rho_0, R_0)$ *and* $\|F\|_{\rho_0,R_0} \leqslant \|f\|_{\rho,R}$.

Lemma A.3. *Let* $\tilde{\chi}(\theta, x, I, y; \varepsilon)$ *be a real analytic function of the form*

$$\tilde{\chi} = \sum_{s\geqslant 1} \tilde{\chi}_{s+2}\varepsilon^{s-1}$$

with $\tilde{\chi}_j \in \mathbb{E}_j(\rho)$, $j \geqslant 3$, *for certain* $\rho > 0$. *Given* $R > 0$, *we define*

$$\delta := \frac{1}{\rho}\|\partial_I \tilde{\chi}_{|\varepsilon=1}\|_{\rho,R},$$
$$\alpha := \max_{j\in\{1,2\}}\left\{\frac{1}{R}\|\partial_{x_j}\tilde{\chi}_{|\varepsilon=1}\|_{\rho,R}, \frac{1}{R}\|\partial_{y_j}\tilde{\chi}_{|\varepsilon=1}\|_{\rho,R}\right\},$$
$$\beta := \frac{1}{R^2}\|\partial_\theta \tilde{\chi}_{|\varepsilon=1}\|_{\rho,R}$$

and suppose that $\delta, \alpha, \beta < 1$. *Then, the canonical transformation* $\Phi_\varepsilon^{-\tilde{\chi}}$, *defined as the flow time* ε *of the non-autonomous Hamiltonian* $-\tilde{\chi}$ *(the time is* ε*) is a real analytic diffeomorphism, defined in* $\mathcal{D}(\rho_\varepsilon, R_\varepsilon)$, *with*

$$\rho_\varepsilon = (1-|\varepsilon|\delta)\rho \qquad \text{and} \qquad R_\varepsilon = \min\{1-|\varepsilon|\alpha, \sqrt{1-|\varepsilon|\beta}\}R$$

for any $-1 \leqslant \varepsilon \leqslant 1$, *and such that* $\Phi_\varepsilon^{-\tilde{\chi}}(\mathcal{D}(\rho_\varepsilon, R_\varepsilon)) \subset \mathcal{D}(\rho, R)$. *Moreover, if we set* $\Phi_\varepsilon^{-\tilde{\chi}} - Id := (\Theta^{(\varepsilon)}, \mathcal{X}^{(\varepsilon)}, \mathcal{I}^{(\varepsilon)}, \mathcal{Y}^{(\varepsilon)})$, *then* $\Theta^{(\varepsilon)}, \mathcal{X}_j^{(\varepsilon)}, \mathcal{I}^{(\varepsilon)}, \mathcal{Y}_j^{(\varepsilon)} \in \mathbb{E}(\rho_\varepsilon, R_\varepsilon)$, $j = 1, 2$, *with*

$$\|\Theta^{(\varepsilon)}\|_{\rho_\varepsilon,R_\varepsilon} \leqslant |\varepsilon|\delta\rho, \qquad \|\mathcal{I}^{(\varepsilon)}\|_{\rho_\varepsilon,R_\varepsilon} \leqslant \beta R^2, \qquad \|\mathcal{X}_j^{(\varepsilon)}\|_{\rho_\varepsilon,R_\varepsilon} \leqslant \alpha R, \qquad \|\mathcal{Y}_j^{(\varepsilon)}\|_{\rho_\varepsilon,R_\varepsilon} \leqslant \alpha R.$$

The remaining results of this section, lemmas A.4, A.5, A.7 and A.8, are technical lemmas used throughout the paper.

Lemma A.4. *Let* $f \in \mathbb{E}_{0,0,M,M}$ *with* $M \in \mathbb{N}$ *(see section 2). We consider the following equivalent expressions for* f *(see (31)–(33)),*

$$\begin{aligned} f &= \sum_{m=0}^{M}\sum_{n=0}^{M} f_{m,n} q_1^m q_2^{M-m} p_1^{M-n} p_2^n \\ &= \sum_{i=0}^{M}\eta_1^i \sum_{j=0}^{M-i} f_{M-j,i+j}\eta_3^{M-i-j}\eta_4^j + \sum_{i=1}^{M}\eta_2^i\sum_{j=0}^{M-i} f_{M-i-j,j}\eta_3^{M-i-j}\eta_4^j \\ &= \sum_{i=0}^{M}\xi_1^i\sum_{j=0}^{M-i}\hat{f}_{i,j}\xi_3^{M-i-j}\xi_4^j + \sum_{i=1}^{M}\xi_2^i\sum_{j=0}^{M-i}\tilde{f}_{i,j}\xi_3^{M-i-j}\xi_4^j, \end{aligned}$$

where $\eta_1 = \xi_1 = q_1p_2$, $\eta_2 = \xi_2 = q_2p_1$, $\eta_3 = q_1p_1$, $\eta_4 = q_2p_2$, $\xi_3 = \mathrm{i}(q_1p_1 + q_2p_2)/2$ *and* $\xi_4 = (q_1p_1 - q_2p_2)/2$. *Then, we have:*

$$\sum_{m=0}^{M}\sum_{n=0}^{M}|f_{m,n}| \leqslant \sum_{i=0}^{M}\sum_{j=0}^{M-i}|\hat{f}_{i,j}| + \sum_{i=1}^{M}\sum_{j=0}^{M-i}|\tilde{f}_{i,j}| \leqslant 2^M\sum_{m=0}^{M}\sum_{n=0}^{M}|f_{m,n}|.$$

Proof. We denote by $\tilde{f} \equiv \tilde{f}(\eta)$ and $\hat{f} \equiv \hat{f}(\xi)$ the polynomial expressions of $f \equiv f(q, p)$ in powers of $\eta = (\eta_1, \eta_2, \eta_3, \eta_4)$ and $\xi = (\xi_1, \xi_2, \xi_3, \xi_4)$ as given in the statement, with $\tilde{f} \in \mathbb{C}_M[\eta]$ and $\hat{f} \in \mathbb{C}_M[\xi]$. Thus, we have that f, $\tilde{f}$ and $\hat{f}$ represent the same function but expressed in different variables. We observe that if we compute the absolute norms $\|f(q, p)\|$, $\|\tilde{f}(\eta)\|$ and $\|\hat{f}(\xi)\|$ (see (11)), using their expansions in powers of the corresponding variables, we obtain:

$$\|f(q,p)\| = \|\tilde{f}(\eta)\| = \sum_{m=0}^{M}\sum_{n=0}^{M}|f_{m,n}|, \qquad \|\hat{f}(\xi)\| = \sum_{i=0}^{M}\sum_{j=0}^{M-i}|\hat{f}_{i,j}| + \sum_{i=1}^{M}\sum_{j=0}^{M-i}|\tilde{f}_{i,j}|.$$

Now, we consider ξ as a function of (q, p), $\xi \equiv \xi(q, p)$, so that $\xi_j(q, p) \in \mathbb{C}_2[q, p]$, $j = 1, \ldots, 4$. Thus, using the multiplicative character of the $\|\cdot\|$-norm (see lemma A.1) and $\|\xi_j(q, p)\| = 1$,

$$\|f(q, p)\| = \|\hat{f}(\xi(q, p))\| \leqslant \sum_{i=0}^{M} \|\xi_1(q, p)\|^i \sum_{j=0}^{M-i} |\hat{f}_{i,j}| \|\xi_3(q, p)\|^{M-i-j} \|\xi_4(q, p)\|^j$$
$$+ \sum_{i=1}^{M} \|\xi_2(q, p)\|^i \sum_{j=0}^{M-i} |\tilde{f}_{i,j}| \|\xi_3(q, p)\|^{M-i-j} \|\xi_4(q, p)\|^j = \|\hat{f}(\xi)\|.$$

This gives the first inequality on the statement. Now, we consider $\eta \equiv \eta(\xi)$ and compute

$$\|\hat{f}(\xi)\| = \|\tilde{f}(\eta(\xi))\| \leqslant \sum_{i=0}^{M} \|\eta_1(\xi)\|^i \sum_{j=0}^{M-i} |f_{M-j,i+j}| \|\eta_3(\xi)\|^{M-i-j} \|\eta_4(\xi)\|^j$$
$$+ \sum_{i=1}^{M} \|\eta_2(\xi)\|^i \sum_{j=0}^{M-i} |f_{M-i-j,j}| \|\eta_3(\xi)\|^{M-i-j} \|\eta_4(\xi)\|^j \leqslant 2^M \|f(q, p)\|,$$

because $\eta_j(\xi) \in \mathbb{C}_1[\xi]$ with $\|\eta_j(\xi)\| \leqslant 2$, $j = 1, \ldots, 4$. This ends the proof. □

Lemma A.5. *We consider the recurrences for* $\{a_k\}_{k\geqslant 3}$ *and* $\{b_{l,m}\}_{l\geqslant 3, m\geqslant 0}$ *introduced in (55) and (56) (see lemma 6.2). Then, there exists* $\tilde{\lambda} > 1$ *such that* $a_k \leqslant \tilde{\lambda}^{k-3}$ *and* $b_{l,m} \leqslant \tilde{\lambda}^m$.

Proof. We proceed by induction with respect to $\mu \equiv k - 3 \equiv m$, for $\mu \geqslant 0$. We observe that this approach is coherent with the two recurrences because $b_{l,m}$ (it means $\mu = m$) depends on $a_3, \ldots, a_{m+2}, b_{l,0}, \ldots, b_{l,m-1}$ (it means $\mu \leqslant m - 1$), and a_k (it means $\mu = k - 3$) depends on $a_3, \ldots, a_{k-1}, b_{k,0}, \ldots, b_{4,k-4}$ (it means $\mu \leqslant k - 4$) and also on $b_{3,k-3}$, which corresponds to the same inductive case than a_k, $\mu = k - 3$, but whose computation only requires data from the previous cases, $\mu \leqslant k - 4$.

The case $\mu = 0$ is trivial. Let us suppose now that such a value of $\tilde{\lambda}$ exists up to order $\mu - 1$ and we check the case $\mu \equiv k - 3 \equiv m$, with $\mu \geqslant 1$. We have

$$a_k \leqslant \tilde{\lambda}^{k-3} \left(\frac{2^{3/2}\tilde{\lambda}^{-1}}{6(k-2)(k-1)!} \sum_{j=1}^{k-3} j(j+2)!(k-j)! + \frac{2}{k-2} \sum_{j=1}^{k-2} \frac{j}{(j+1)!} \left(\frac{1}{3\tilde{\lambda}}\right)^{j-1} \right), \tag{A1}$$

$$b_{l,m} \leqslant \tilde{\lambda}^m \left(\frac{\tilde{\lambda}^{-1}}{6m(l+m-1)!} \sum_{j=1}^{m} j(j+2)!(l+m-j)! \right). \tag{A2}$$

Hence, the conditions that we have to impose on $\tilde{\lambda}$ become clear: we want $\tilde{\lambda} > 1$ such that the expressions between brackets in (A1) and (A2) are both less than or equal to one.

We consider first (A2). As we have to deal with values of $l \geqslant 3$, we have:

$$\frac{1}{m(l+m-1)!} \sum_{j=1}^{m} j(j+2)!(l+m-j)! \leqslant \frac{1}{m(m+2)!} \sum_{j=1}^{m} j(j+2)!(m+3-j)! := \Sigma_m,$$

because the case $l = 3$ is the one that gives the worst possibility for this expression. Now, we observe that the term in the sum defining Σ_m corresponding to the index j contains a product of two factorial, given by $(j+2)!(m+3-j)!$, that does not change if we replace j by $m - j + 1$. This motivates us to consider a different alignment for the sum, by joining the contribution of both terms. To do this, we have to distinguish between the cases m even and

m odd. If m is even,

$$\Sigma_m = \frac{m+1}{m} \sum_{j=1}^{m/2} \frac{(j+2)!(m+3-j)!}{(m+2)!} \leqslant \frac{3}{2} \sum_{j=1}^{+\infty} \frac{(j+2)!(j+3)!}{(2j+2)!} := A < +\infty,$$

where we have used that the quotient $(m+3-j)!/(m+2)!$ is decreasing as a function of m, and thus, we can bound it by replacing m in the sum by the first value of m for which the script j appears in the sum, that is, $m = 2j$. Then, we only have to bound it for $m \geqslant 2$.

If m is odd then the indexes $j = (m+1)/2$ and $m - j + 1$ are coincident, but we obtain an upper bound for Σ_m counting twice their contribution. Hence,

$$\Sigma_m \leqslant \frac{m+1}{m} \sum_{j=1}^{(m+1)/2} \frac{(j+2)!(m+3-j)!}{(m+2)!} \leqslant 2 \sum_{j=1}^{+\infty} \frac{(j+2)!(j+2)!}{(2j+1)!} := B < +\infty,$$

using analogous arguments as in the previous bound and that $m \geqslant 1$.

The control of the first sum inside the brackets of (A1) is very similar to the control of Σ_m, and we omit the details. We only mention that if we set $C := \max\{(2/3)A, (3/4)B\}$, then

$$\frac{1}{(k-2)(k-1)!} \sum_{j=1}^{k-3} j(j+2)!(k-j)! \leqslant C, \qquad k \geqslant 4.$$

Finally, the second sum inside the brackets of (A1) verifies:

$$\frac{2}{k-2} \sum_{j=1}^{k-2} \frac{j}{(j+1)!} \left(\frac{1}{3\tilde{\lambda}}\right)^{j-1} \leqslant f\left(\frac{1}{3\tilde{\lambda}}\right) \leqslant f\left(\frac{1}{3}\right) = 9 - 6\mathrm{e}^{1/3} := D < 1, \qquad k \geqslant 4,$$

where

$$f(x) := \sum_{j=1}^{+\infty} \frac{j}{(j+1)!} x^{j-1} = \frac{(x-1)\mathrm{e}^x + 1}{x^2}$$

is positive and strictly increasing for $x \geqslant 0$.

Thus, it is clear that we can define $\tilde{\lambda} > 1$ as $\tilde{\lambda} := \max\{2^{3/2}C/(1-D), A, B\}/6$. □

Remark A.6. The sums defining A and B can be expressed in terms of hypergeometric functions so that a numerical value for them can be easily computed. Thus, we obtain the following numerical estimate for $\tilde{\lambda} = 20.362\,07\ldots$.

Lemma A.7. *Let $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$ verify the Diophantine condition (3), that is, $|\langle k, \omega\rangle| \geqslant \gamma |k|_1^{-\tau}$, $\forall k \in \mathbb{Z}^2 \backslash \{(0,0)\}$. Then, we have the following estimate for the quantities Ω_s defined in (42),*

$$\Omega_s := \min_{\substack{k \in \mathbb{Z}^2 \backslash \{(0,0)\} \\ |k_2| \leqslant s}} \{|\langle k, \omega\rangle|, 1\} \geqslant \tilde{\gamma} s^{-\tau}, \qquad s \geqslant 1,$$

where $0 < \tilde{\gamma} \leqslant 1$ is given by $\tilde{\gamma} := \min\{(3/2 + |\omega_2/\omega_1|)^{-\tau}\gamma, |\omega_1|, 1\}$.

Proof. We only have to take the minimum of the following two cases:

- If $0 < |k_2| \leqslant s$, then there is a sole $k_1 \in \mathbb{Z}$ such that $|k_1 + k_2\omega_2/\omega_1| \leqslant 1/2$. This k_1 verifies

$$|k_1| \leqslant \left|k_1 + \frac{k_2\omega_2}{\omega_1}\right| + \left|\frac{k_2\omega_2}{\omega_1}\right| \leqslant \frac{1}{2} + |k_2| \left|\frac{\omega_2}{\omega_1}\right|.$$

Thus, using the Diophantine condition (3) we obtain,

$$\min_{k_1\in\mathbb{Z}}\{|\langle k,\omega\rangle|\} \geqslant \gamma\left(\frac{1}{2}+|k_2|\left|\frac{\omega_2}{\omega_1}\right|+|k_2|\right)^{-\tau} \geqslant \tilde{\gamma}|k_2|^{-\tau} \geqslant \tilde{\gamma}s^{-\tau}.$$

- If $k_2=0$, $|\langle k,\omega\rangle| = |k_1\omega_1| \geqslant |\omega_1| \geqslant \tilde{\gamma}s^{-\tau}$

and then the proof is complete. □

Lemma A.8. *With the same hypotheses of lemma A.7, we consider the quantities* $\beta_j := 2^j j!/\Omega_j^{j+1}$*, defined in (50). Then, given any* $\varepsilon > 0$*, there is a constant* $d_\varepsilon \equiv d_\varepsilon(\tilde{\gamma},\tau)$*, such that*

$$\beta_3\cdots\beta_s \leqslant d_\varepsilon(s+1)^{(\tau+1+\varepsilon)(s+1)^2/2}, \qquad s\geqslant 3. \tag{A3}$$

Proof. Applying lemma A.7 on the definition of β_j, $3\leqslant j\leqslant s$, we have,

$$\beta_3\cdots\beta_s \leqslant 2^{3+\cdots+s}\tilde{\gamma}^{-4-\cdots-(s+1)}[3^4\cdots s^{s+1}]^\tau 3!\cdots s!.$$

Now, we use Stirling's formula, $j! = \sqrt{2\pi}\, j^{j+1/2}\mathrm{e}^{-j+\xi_j/(12j)}$, with $0<\xi_j<1$, thus obtaining:

$$\beta_3\cdots\beta_s \leqslant 2^{-3+s(s+1)/2}\mathrm{e}^{3-s(s+1)/2}\tilde{\gamma}^{6-(s+1)(s+2)/2}(\sqrt{2\pi}\mathrm{e}^{1/36})^{s-2}(s!/2)^{\tau+1/2}[3^3\cdots s^s]^{\tau+1}.$$

Moreover,

$$3^3\cdots s^s = \exp\left(\sum_{j=3}^{s} j\log j\right) \leqslant \exp\left(\int_3^{s+1} x\log x\,\mathrm{d}x\right) = \frac{\mathrm{e}^{9/4}}{3^{9/2}}\frac{(s+1)^{(s+1)^2/2}}{\mathrm{e}^{(s+1)^2/4}},$$

where we have used that $f(x) := x\log x$ is increasing if $x\geqslant 3$. Now, if we combine all:

$$\beta_3\cdots\beta_s \leqslant \frac{(2/\mathrm{e})^{-3+s(s+1)/2}}{\tilde{\gamma}^{-6+(s+1)(s+2)/2}}(\sqrt{2\pi}\mathrm{e}^{1/36})^{s-2}\left(\frac{s!}{2}\right)^{\tau+1/2}\left(\frac{\mathrm{e}^{9/4}}{3^{9/2}}\frac{(s+1)^{(s+1)^2/2}}{\mathrm{e}^{(s+1)^2/4}}\right)^{\tau+1}.$$

Then, it is clear that $\lim_{s\to+\infty}(\beta_3\cdots\beta_s)(s+1)^{-(\tau+1+\varepsilon)(s+1)^2/2} = 0$, showing the existence of d_ε. □

Remark A.9. As has been pointed in remark 7.2, the estimate (A3) is a very important one. Let us define

$$\alpha_s := \frac{\log(\beta_3\cdots\beta_s)}{s^2\log s}.$$

Then, numerics suggests that for a Diophantine vector $\omega\in\mathbb{R}^2$ the behaviour

$$\limsup_{s\to+\infty}\alpha_s = \alpha$$

for certain $\alpha > 0$, is the correct one. Thus, no remarkable improvement can be expected for (A3). The most simple case is when $\omega_1 = 1$ and $\omega_2 = (\sqrt{5}-1)/2$ (the golden mean), for which $\lim_{s\to+\infty}\alpha_s$ can be easily computed. Indeed, by using Stolz criteria and Stirling's formula we have that if

$$\lim_{s\to+\infty}\frac{1}{2}\left(1-\frac{\log\Omega_s}{\log s}\right)$$

exists, then it is equal to α. To compute this limit, one proceeds as follows. Let us denote by $\{F_k\}_{k\geqslant 1}$ the Fibonacci numbers. From the well-known Diophantine properties of the golden mean we have that there exists $\gamma > 0$ making the approximation

$$\Omega_{F_k} = |F_{k-1}-\omega_2F_k| \approx \gamma F_{k+1}^{-1},$$

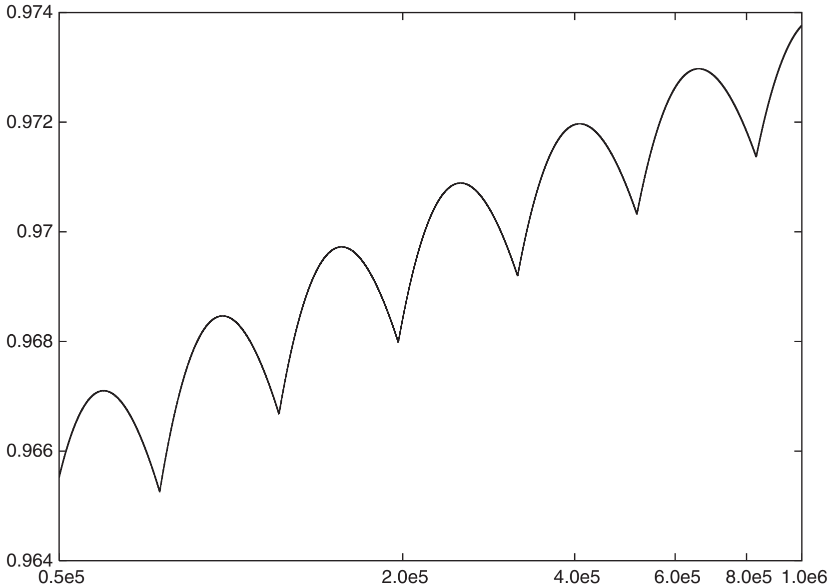

Figure A1. α_s (vertical axis) plotted with respect to s (horizontal axis in logarithmic scale).

asymptotically valid for large values of k (recall the definition of Ω_s in (42)). With the above expression in mind

$$\lim_{k\to+\infty} \frac{\log \Omega_k}{\log F_k} = -\lim_{k\to+\infty} \frac{\log F_{k+1}}{\log F_k} = -1.$$

To extend this limit to non-Fibonacci numbers let $F_{k(s)}$ be, for a given s, the largest Fibonacci number not bigger than s, so $F_{k(s)} \leqslant s < F_{k(s)+1}$. Hence,

$$\frac{\log \Omega_{F_{k(s)+1}}}{\log F_{k(s)+1}} \leqslant \frac{\log \Omega_s}{\log s} \leqslant \frac{\log \Omega_{F_{k(s)}}}{\log F_{k(s)}}.$$

But $k(s) \to \infty$ as s does, giving $\alpha = 1$ (recall we can take $\tau = 1$ for $\omega = (1, (\sqrt{5}-1)/2)$. In figure A1 we plot α_s as a function of $\log s$ for values of s from 5×10^4 to 1×10^6.

References

[1] Bridges T J, Cushman R H and Mackay R S 1995 Dynamics near an irrational collision of eigenvalues for symplectic mappings *Fields Inst. Commun.* **4** 61–79
[2] Bruno A D 1989 Normalization of a Hamiltonian system near a cycle or a torus *Russ. Math. Surveys* **44** 53–89
[3] Bruno A D 1994 *The Restricted 3-Body Problem: Plane Periodic Orbits* (Berlin: de Gruyter)
[4] Corless R M, Gonnet G H, Hare D E G, Jeffrey D J and Knuth D E 1996 On the Lambert *W* function *Adv. Comput. Math.* **5** 329–59
[5] Delshams A and Gutiérrez P 1996 Effective stability and KAM theory *J. Diff. Eqns* **128** 415–90
[6] Delshams A and Gutiérrez P 1996 Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system *J. Diff. Eqns* **131** 277–303
[7] Giorgilli A, Delshams A, Fontich E, Galgani L and Simó C 1988 Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem *J. Diff. Eqns* **77** 167–98
[8] Giorgilli A and Galgani L 1978 Formal integrals for an autonomous Hamiltonian system near an equilibrium point *Cel. Mech.* **17** 267–80

[9] Giorgilli A and Galgani L 1985 Rigorous estimates for the series expansions of Hamiltonian perturbation theory *Cel. Mech.* **37** 95–112

[10] Jorba A and Ollé M 2004 Invariant curves near Hamiltonian–Hopf bifurcations of four-dimensional symplectic mappings *Nonlinearity* **17** 691–710

[11] Jorba À and Villanueva J 1997 On the normal behavior of partially elliptic lower-dimensional tori of Hamiltonian systems *Nonlinearity* **10** 783–822

[12] Jorba À and Villanueva J 1998 Numerical computation of normal forms around some periodic orbits of the restricted three-body problem *Physica* D **114** 197–229

[13] Lochak P and Meunier C 1988 *Multiphase Averaging for Classical Systems: With Applications to Adiabatic Theorems* (*Applied Mathematical Sciences* vol 72) (New York: Springer)

[14] Meyer K R 1998 The evolution of the stable and unstable manifold of an equilibrium point *Cel. Mech. Dyn. Astron.* **70** 159–65

[15] Meyer K R and Schmidt D S 1971 Periodic orbits near $\mathcal{L}_4$ for mass ratios near the critical mass ratio of Routh *Cel. Mech.* **4** 99–109

[16] Ollé M, Pacha J R and Villanueva J 2005 Dynamics close to a non semi-simple $1 : -1$ resonant periodic orbit *Discrete Contin. Dyn. Syst.* B in press (*Preprint* http://www.ma1.upc.edu/recerca/2003.html)

[17] Ollé M, Pacha J R and Villanueva J 2004 Motion close to the Hopf bifurcation of the vertical family of periodic orbits of L_4 *Cel. Mech. Dyn. Astron.* **90** 89–109

[18] Pacha J R 2002 On the quasi-periodic Hamiltonian Andronov–Hopf bifurcation *PhD Thesis* Universitat Politècnica de Catalunya, http://tdx.cesca.es

[19] Poincaré H 1928 *Thèse* (*Oeuvres* vol 1) (Paris: Gauthier-Villars) pp IL–CXXXII

[20] Rüssmann H 1975 On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus *Dynamical Systems, Theory and Applications* (*Lecture Notes in Physics* vol 38) ed J Moser (Berlin: Springer) pp 598–624

[21] Schmidt D 1994 Versal normal form of the Hamiltonian function of the restricted problem of three bodies near $\mathcal{L}_4$. Oscillations in nonlinear systems: applications and numerical aspects *J. Comput. Appl. Math.* **52** 155–76

[22] Siegel C L and Moser J K 1971 *Lectures on Celestial Mechanics* (New York: Springer)

[23] Simó C 1989 Estabilitat dels sistemes hamiltonians *Mem. Real Acad. Cienc. Artes Barcelona* vol 48, No 7 (Barcelona: Impremta Juvenil S.A.) pp 303–44

[24] Van der Meer J C 1985 *The Hamiltonian Hopf Bifurcation* (*Lecture Notes in Mathematics* vol 1160) (Berlin: Springer)