UNIVERSITAT POLITÈCNICA DE CATALUNYA

DEPARTAMENT DE MATEMÀTICA APLICADA I

Memòria per a aspirar al grau de doctor en Matemàtiques

ON THE QUASI-PERIODIC HAMILTONIAN ANDRONOV-HOPF BIFURCATION

Directors:
Dra. Mercè Ollé Torner
Dr. Jordi Villanueva Castelltort

Doctorand:
Juan Ramón Pacha Andújar

Barcelona, Juliol, 2002

Acknowledgements

I would like to express my gratitude to all the persons who have helped me these years along.

First of all, I wish to thank my supervisors, Mercè Ollé and Jordi Villanueva, for their teaching and support beyond their merely academic duties, for the uncountable many hours they lost with me, for their patience! I'll be in debt to both.

I also owe much to Pere Gutiérrez, he gave me good ideas and advices ...and always let me take part in his (usually wild) mountain-adventures.

And a lot of people: all the professors of the *Departamento de Matemàtica Aplicada I* here at the UPC, and those at the UB-UPC Dynamical System group: they never denied me neither their knowledge nor their time.

And last, but not least, I thank my buddies at the day-by-day work –specially my officemates, Tomás Lázaro and Marta Peña–, for their support and friendship.

Contents

List of Figures

Notation and list of symbols

As a rule, each notation is explained where it first appears. Nevertheless, we collect here (see below) some of the notations used frequently in the text. The symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ stand for the basic sets (positive integers, integers, rationals, real and complex numbers respectively). Vectors are boldfaced, so typically we note $\boldsymbol{z}^* = (x, y) \in \mathbb{R}^2$ and an asterisk superscripting a vector or a matrix denotes its transpose. Holomorphic functions that are real valued for real arguments will be called "real analytic". Equations and formulas in chapters are numbered in the following way: the leftmost digit corresponds to the chapter, the second correspond to the section and the rightmost one is the number the formula makes up in this section. So, for example, label (3.5.4) denotes the fourth formula of the fifth section in the third chapter. Statements of theorems, propositions, lemmas and definitions are slanted. This is not the case neither for examples nor for remarks, so the symbols ($\diamond$) and ($\spadesuit$) are used as endpoints of the former and of the latter respectively. Similarly, the end of proofs (of theorems, propositions, lemmas...) are marked by a square ($\square$). In the following list of symbols, we do not include those with highly specific and long definitions as, for instance (see appendix 3.2.41), those of the norms $|\cdot|_\rho, |\cdot|_{\rho,R}$,... and so on.

List of symbols

$(\alpha)_k$	the generalized factorial (Pochhamer symbol): $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1)$, $(\alpha)_0 = 1$
$[m, n]$	(m, n integers; $m \geq 0$, $n > 0$) the remainder of the integer division of m by n
$\overline{\boldsymbol{z}}$	for $\boldsymbol{z} \in \mathbb{C}$, its complex conjugate
$\cdot$	between real numbers, their ordinary product $(3 \cdot 2 = 6)$
$\dot{x}$	differentiation with respect to time: $\dot{x} = \frac{dx}{dt}$
$\langle \boldsymbol{u}, \boldsymbol{v} \rangle$	the standard inner product of two vectors, so $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j=1}^{n} u_j v_j$ for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$
e	appearing in formulas: the base of the natural logarithm, i. e., $\mathrm{e} = \exp(1)$. In the text both notations e and $\exp(1)$ are used indistinctly
i	the imaginary unit: $\mathrm{i} = \sqrt{-1}$
meas	Lebesgue measure
$\lfloor x \rfloor$	the integral part of $x \in \mathbb{R}$

$\mathbb{T}^n$ the standard n torus: $\mathbb{T}^n = S^n = (\mathbb{R}/2\pi\mathbb{Z})^n$

$\times$ between sets, the set (Cartesian) product; between real numbers, their ordinary product ($3 \times 2 = 6$)

$\mathbb{Z}_+$ the non-negative integers, i. e.: $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$

$\{f, g\}$ Poisson bracket of the functions f and g

A^* if A is a $n \times m$ matrix, A^* denotes its transposed

I_n the identity $n \times n$ matrix

J_n the matrix of the standard symplectic form, i. e.: $J_n = \left(\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right)$

$L_g f$ $L_g f = \{f, g\}$

W_0 The *principal branch* of the Lambert W function. See page 78

Introduction

In this memoir, the main topic is the study of the dynamics close to resonant (in the sense we are going to precise) periodic orbits. We focus on real analytic three-degree of freedom Hamiltonian systems and our objective is to investigate the quasi-periodic bifurcation phenomena linked to one parameter families of periodic orbits undergoing $1 : -1$ resonance. To be more concrete, we assume that the orbits of the family are first linearly stable; for a *critical* value of the parameter, the nontrivial (i. e., those different from one) characteristic multipliers of the corresponding periodic orbit collide (the so called Krein collision) on the unit circle: this corresponds to the *critical* $1 : -1$ *resonant* or simply *resonant* periodic orbit. Then, if certain generic conditions are met, the characteristic multipliers leave out the unit circle to the complex plane, hence the family looses its (linear) stability and the periodic orbits become *complex unstable*.

This transition stable-complex unstable is not a strange or uncommon phenomenon, so examples can be found in several fields of science, from astronomy –galactic dynamics (see Martinet, 1984; Pfenniger, 1985b, 1990; Ollé and Pfenniger, 1998), planetary theory (e. g. in Hadjidemetriou, 1985)–, to particle accelerators (Howard et al., 1986). Moreover, not only in three degrees of freedom Hamiltonian systems, but also in higher dimensional problems. For example, in Ollé et al. (1999) were found families of periodic orbits with transitions stable-complex unstable for the spatial elliptic three-body problem (three and a half degrees of freedom): two pair of characteristic multipliers collide, while the third stays on the unit circle.

On the other hand, three-degree of freedom Hamiltonian systems can be investigated through Poincaré (or first return) four dimensional maps (see appendix B, section B.2 for a short description of Poincaré maps paraphernalia). This reduces the study of a Hamiltonian system in $\mathbb{R}^6$ to that of symplectic maps defined on a certain four dimensional *surface of section*. It turns out that each element in the dynamical backbone of the flow has its map counterpart. So for instance (and particularly interesting for us): periodic orbits and two dimensional invariant tori on the system are pictured as fixed points and invariant curves on the map respectively. Even more, the eigenvalues of a fixed point (we mean, those of the linearization of the map around that fixed point) are given by the nontrivial characteristic multipliers of the periodic orbit of the flow it comes from.

Hence (despite the interest they could have on their own), the study of symplectic maps are often used to envisage some qualitative properties of Hamiltonian systems. With this aim, in order to explore the motion close to complex instability, several researchers have investigated one-parameter families of symplectic diffeomorphisms with a fixed point undergoing Krein collisions between its characteristic multipliers. Particularly, in Pfenniger (1985a), and Ollé and Pfenniger (1999) two symplectic generalizations of the Froeschlé's map, $T_s, T_t : \mathbb{R}^4 \to \mathbb{R}^4$, were explored (see section 1.3.2 of chapter 1 for a more detailed description). Both T_s and T_t have a fixed point at 0 and depend on a parameter L in such a way that, for values L less than some critic value, L_{crit}, the fixed point is linearly stable (its characteristic multipliers lie on the unit circle), when $L = L_{crit}$ a pairwise eigenvalue collision there takes place, and for $L > L_{crit}$ they leave the unit circle to the complex plane. In addition, we shall assume that the argument of the two complex conjugate multipliers of the fixed point for L_{crit} (i. e., when eigenvalue collision takes place) are not commensurable with 2π –one speaks then about *irrational* collision–. Therefore, for T_s, stable invariant curves unfold for each $L > L_{crit}$ (so, on the "unstable side", since for these values of the parameter, the fixed point is complex-unstable). This is known as the *direct* bifurcation. For the map T_t, though, unstable invariant curves rise, for $L < L_{crit}$ "on the stable side": they appear thus in a similar way as limit cycles do in the classical Andronov-Hopf bifurcation (see Andronov et al., 1959, chap. VI, §4). We shall also mention that rational collisions (where the argument of the characteristic multipliers is 2π-commensurable at collision) can occur. Then, generically, multiple periodic points unfold. This situation is likewise investigated at the first of the papers quoted at the beginning of the paragraph (and analytically, in a more general context, in Bridges, 1990, 1991). However, as we are interested in quasi-periodic motions, irrationality is assumed throughout.

With T_s and T_t as paradigmatic examples, we wonder if such behavior (well understood from *numerical* research) could be dealt analytically in a general symplectic context, as Bridges, Cushman and Mackay (1995) did for symplectic maps and Van der Meer (1985) for the Hamiltonian Hopf bifurcation at equilibrium points in two-degrees of freedom Hamiltonian systems (also, see Meyer and Schmidt, 1971; Schmidt, 1994; Meyer, 1998). The improvements of this work can be summarized as: (i) We rely on normal forms as *one* of the key tools of our approach, deriving in a *constructive* way and up to *any* (arbitrary) order, a versal normal form of the Hamiltonian around the resonant periodic orbit. Analyzing the (truncated) normal form, we describe the mechanism of the 2D-invariant tori unfolding, according to the Hamiltonian Hopf pattern, identifying those parameters which govern not only the bifurcation, but also its character. We remark that this is not a merely qualitative process for, in addition, accurate parametrizations of the families of invariant tori are derived in this way. (ii) We compute "optimal" bounds for the remainder of the normal form, so one expects to prove the preservation of a higher number (in measure sense) of invariant tori –than, indeed, with a less sharp estimates–. (iii) And, finally, we apply KAM methods to establish the persistence of most (in the measure sense) of the bifurcated invariant tori. A more detailed explanation of these points is given below.

In chapter 1 we state the problem. So let us consider a real analytic three-degree of freedom Hamiltonian, $H(\boldsymbol{\zeta})$, $\boldsymbol{\zeta}^* = (\xi_1, \xi_2, \xi_3, \eta_1, \eta_2, \eta_3) \in \mathbb{R}^6$ with the corresponding system

$$\dot{\boldsymbol{\zeta}} = J_3 \operatorname{grad} H(\boldsymbol{\zeta}), \tag{1}$$

being J_3 the matrix of the standard canonical 2-form in $\mathbb{R}^6$ (see appendix B, section B.1). Suppose that this system has a non-degenerate family of periodic orbits, $\{\mathcal{M}_\sigma\}_{\sigma\in\mathbb{R}}$, such that for some value of the parameter, say $\sigma = 0$, the corresponding (critical, resonant) periodic orbit, $\mathcal{M}_0$, has a collision of its nontrivial characteristic multipliers. To be more precise (see figure 1.2 in chapter 1), suppose that, for $\sigma < 0$, these characteristic multipliers of $\mathcal{M}_\sigma$ lie on the unit circle, they approach pairwise as σ goes to 0, for this value they collide and separate towards the complex plane when $\sigma > 0$. Moreover:

(i) As it has been already mentioned above, we assume the collision is irrational: more precisely, if $2\pi\nu_0$ is the characteristic exponent corresponding to the characteristic multiplier λ_0 of the resonant periodic orbit $\mathcal{M}_0$ (so $\lambda_0 = \mathrm{e}^{2\pi\, i\nu_0}$), then $\nu_0 \notin \mathbb{Q}$.

(ii) We suppose both, non-degeneracy of the family of periodic orbits, –that is, we ask the twist condition[(1)]–, and non-degeneracy of the collision, since one requires the eigenvalues to leave the unit circle for $\sigma > 0$. These assumptions imply that the monodromy matrix of $\mathcal{M}_0$ should not have trivial (diagonal) Jordan blocks, neither those corresponding to the (double) eigenvalues λ, $\overline{\lambda}$ nor the (non-diagonal) block associated to the (double) eigenvalue 1.

Prior to the normal form computations, we need to perform some previous reductions. The objective is to transform the initial raw Hamiltonian into a new one that eases the application of the normal form algorithm. This process involves: (i) the introduction of local coordinates around the orbit $\mathcal{M}_0$, (ii) Floquet reduction of the quadratic part of the Hamiltonian and (iii) complexification to make this quadratic part as simple as possible.

Next, we use Giorgilli-Galgani algorithm (see definition 1.7 in chapter 1 and the references given there) to carry out the nonlinear reduction. The reasons for the choice of the Giorgilli-Galgani machinery to compute the nonlinear normalization are basically two: one of them is motivated by the practical implementation of this methodology. So, if one plans to apply normal form computations to our problem (we mean, a three-degree of freedom Hamiltonian system with a family of periodic orbits undergoing a transition stable to complex-unstable), then the Giorgilli-Galgani algorithm is a very efficient way to carry this process out (see Giorgilli et al., 1989; Simó, 1989; Delshams and Gutiérrez, 1996). In this sense, we want to stress that the solvability of the homological equations (with the intricacies due to its Jordan block structure) has been discussed in chapter 1 in a constructive way, we mean: not only the resonant terms are identified, but also we issue an algorithm to compute them explicitly, as well as the coefficients of the generating function (compare with Schmidt, 1994; Bridges, Cushman and Mackay, 1995).

The second reason is of a deeper technical nature, so it is a more involved task to discuss it here. However, we can try to give a systematic outline: first of all, it is well known that in presence of resonances, normal forms do not converge in general. Hence, a natural question (that in the present context is discussed in chapter 2) is to ask what is the optimal order up to which the normal form should be computed. As this order is not known *a priori*, a good idea to obtain it *a fortiori* is to perform nonlinear normalization as a composition of canonical transformations such that, at any step, the first order correction of the corresponding transformation kills all the terms not left in normal form after the previous step (see, for instance Nekhoroshev, 1977, 1979; Perry and Wiggins, 1994). In the process, one sets up and solves homological equations holding monomials of arbitrary high order, and finally is precisely the number of steps what determines the order of the normal form. However, though this process works well in the semi-simple[(2)] case, in our context, if we try to solve the homological equations at any order, the solution is no longer convergent (even when we have good arithmetic properties for the frequencies). Thus it will be clear that, to derive the optimal normalization order, one has to proceed order by

[(1)]At least locally at $\sigma = 0$, so if $\omega(0)$ is the frequency of the critical periodic orbit, $\mathcal{M}_0$, then it must be $\omega'(0) \neq 0$

[(2)]We mean, when the monodromy matrix of the periodic orbit has pairwise different normal eigenvalues.

order and then, closed algorithms become much more efficient and, among them, that of Giorgilli-Galgani, fits specially well to our purposes.

Consequently, it is worth remarking here that the *homological equations* we have to solve in order to determine the generating function cannot be transformed into diagonal form –as happens when one normalizes around semi-simple periodic orbits–. This makes the derivation of the structure of the normal form a more involved task. At the end, we are able to give, in theorem 1.24 on page 37, a versal normal form for three-degree of freedom Hamiltonian around the $1 : -1$ resonant periodic orbit $\mathcal{M}_0$. Below, a (short) version of this result follows.

Theorem 1. *Consider the above Hamiltonian system* (1). *Under the forementioned conditions, H can be reduced, by means of a symplectic change defined in a neighborhood of the periodic orbit $\mathcal{M}_0$, to a real Hamiltonian which (keeping the same name for the old and the transformed one), is given by*

$$H(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}) = \omega_1 I_1 + \omega_2 \boldsymbol{y} \times \boldsymbol{x} + \tfrac{1}{2}|\boldsymbol{y}|_2^2 + \mathcal{Z}_r\left(\tfrac{1}{2}|\boldsymbol{x}|_2^2, I_1, \boldsymbol{y} \times \boldsymbol{x}\right) + \mathfrak{R}^{(r)}(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}),$$

with the notation,

$$|\boldsymbol{x}|_2 = (x_1^2 + x_2^2)^{1/2}, \quad |\boldsymbol{y}|_2 = (y_1^2 + y_2^2)^{1/2}, \quad \boldsymbol{x} \times \boldsymbol{y} = x_1 y_2 - x_2 y_1,$$

ω_1 being the frequency of the resonant periodic orbit and ω_2/ω_1 (assumed not commensurable with 2π), the characteristic exponent of the critical periodic orbit, $\mathcal{M}_0$, respectively; $\mathcal{Z}_r(u, v, w)$ a polynomial of degree[(3)] *$\lfloor r/2 \rfloor$ beginning with quadratic terms and $\mathfrak{R}^{(r)}(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y})$ 2π-periodic in θ_1 holding terms of "degree" higher than r (see chapter 1, section 1.7, for a precise definition of what means degree in our context).*

The quadratic part of the Hamiltonian plus $\mathcal{Z}_r(\cdots)$ in the statement of this last theorem will be denoted by $Z^{(r)}$, i. e.,

$$Z^{(r)}(\boldsymbol{x}, I_1, \boldsymbol{y}) = \omega_1 I_1 + \omega_2 \boldsymbol{y} \times \boldsymbol{x} + \tfrac{1}{2}|\boldsymbol{y}|_2^2 + \mathcal{Z}_r\left(\tfrac{1}{2}|\boldsymbol{x}|_2^2, I_1, \boldsymbol{y} \times \boldsymbol{x}\right).$$

Moreover, to formulate some results, it will be convenient to express the polynomial $\mathcal{Z}_2$ as,

$$\mathcal{Z}_2(u, v, w) = \frac{1}{2}(au^2 + bv^2 + cw^2) + duv + euw + fvw, \tag{2}$$

with a, b, c, d, e f real coefficients. $Z^{(r)}$ is the (real) normal form up to order (degree) r and is proven to be integrable, whereas $\mathfrak{R}^{(r)}$ stands for the (non-integrable) remainder. Then, from section 1.8, and up to the end of chapter 1, $\mathfrak{R}^{(r)}$ is dropped so only the dynamics of the normal form is accounted for.

Parametrization of the family of periodic orbits: it is immediately seen that

$$\left.\begin{aligned} &\theta_1 = (\omega_1 + \partial_2 \mathcal{Z}_r(0, I_1, 0))t + \theta_1^0, \\ &I_1 = \text{const.}, \\ &x_1 = x_2 = y_1 = y_2 = 0, \end{aligned}\right\} \tag{3}$$

[(3)] Throughout the text, the symbol $\lfloor \cdot \rfloor$ will mean the integral part, i. e., for $x \in \mathbb{R}$,

$$\lfloor x \rfloor = \max\{z \in \mathbb{Z} : z \leq x\}.$$

is a family (with I_1 as parameter) of periodic orbits of the Hamiltonian system $Z^{(r)}$, which will match the family $\mathcal{M}_\sigma$ of the complete Hamiltonian. Whence, the characteristic exponents associated to the normal directions can be computed in terms of the action I_1 to give

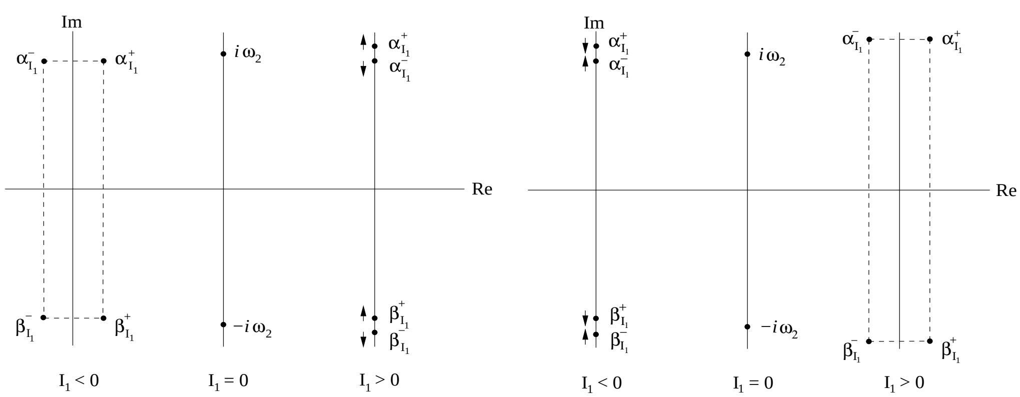
$$\alpha^{\pm}_{I_1} = \mathrm{i}(\omega_2 + fI_1) \pm \sqrt{-dI_1 + \mathrm{O}(I_1^2)} + \mathrm{O}(I_1^2), \tag{4a}$$

$$\beta^{\pm}_{I_1} = -\mathrm{i}(\omega_2 + fI_1) \pm \sqrt{-dI_1 + \mathrm{O}(I_1^2)} + \mathrm{O}(I_1^2), \tag{4b}$$

from these expressions we see that if $|I_1|$ is small enough, the sign of the content inside the square roots, depends mainly on the sign of $-dI_1$. This opens two possibilities:

Case 1. $d > 0$, The family of periodic orbits (3) is complex unstable for $I_1 < 0$, and (linearly) stable for $I_1 > 0$. See figure below.

Case 2. $d < 0$. The family turns out to be (linearly) stable for $I_1 < 0$ and complex unstable for $I_1 > 0$, as it can be appreciated in the figure.



(a) $d > 0$. (b) $d < 0$.

Figure 1: We note that when $I_1 = 0$, then $\alpha_0^- = \alpha_0^+ = i\omega_2$ and $\beta_0^- = \beta_0^+ = -i\omega_2$ (collision of characteristics exponents). Therefore, the family changes its linear character from complex-unstable to stable (when $d > 0$, fig. 1(a)), or vice-versa (when $d < 0$, fig.1(b)).

On the quasi-periodic solutions: further, the quasi-periodic solutions (note, solutions of $Z^{(r)}$) we seek are more easily derived if first the symplectic change,

$$\begin{aligned} x_1 &= \sqrt{2q}\cos\theta_2, & y_1 &= -\frac{I_2}{\sqrt{2q}}\sin\theta_2 + p\sqrt{2q}\cos\theta_2, \\ x_2 &= -\sqrt{2q}\sin\theta_2, & y_2 &= -\frac{I_2}{\sqrt{2q}}\cos\theta_2 - p\sqrt{2q}\sin\theta_2, \end{aligned} \tag{5}$$

is applied to $Z^{(r)}$ introducing thus an extra angle, θ_2, and its conjugate action, I_2. Taking these new coordinates: $(\theta_1, \theta_2, q, I_1, I_2, p)$, the corresponding Hamiltonian equations are bound to be,

$$\begin{aligned}
\dot{\theta}_1 &= \omega_1 + \partial_2 \mathcal{Z}_r(q, I_1, I_2), \\
\dot{\theta}_2 &= \omega_2 + \frac{I_2}{2q} + \partial_3 \mathcal{Z}_r(q, I_1, I_2), \\
\dot{q} &= 2qp, \\
\dot{I}_1 &= 0, \\
\dot{I}_2 &= 0, \\
\dot{p} &= -p^2 + \frac{I_2^2}{4q^2} - \partial_1 \mathcal{Z}_r(q, I_1, I_2).
\end{aligned} \tag{6}$$

So one might compute 2D-invariant tori as equilibrium points of $\dot{q} = 0$, $\dot{p} = 0$. With this idea, the next theorem follows naturally (see section 1.9, page 41).

Theorem 2. *Let the coefficients a, b and d be those in $\mathcal{Z}_2$. Then:*

(i) If $a \neq 0$, there exists a real analytic function $\Upsilon : \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$, defined in some neighborhood of the origin $\mathcal{U} \subset \mathbb{R}^2$, such that the functions $\{\mathcal{T}(t; \boldsymbol{J}, \boldsymbol{\theta}^0), t \in \mathbb{R}\}_{\boldsymbol{J} \in \mathcal{U}, \boldsymbol{\theta}^0 \in \mathbb{T}^2}$, with $\boldsymbol{J}^* = (J_1, J_2)$ and,

$$\mathcal{T}(t; \boldsymbol{J}, \boldsymbol{\theta}^0) = \begin{pmatrix} \boldsymbol{\Omega}(\boldsymbol{J})t + \boldsymbol{\theta}^0 \\ \Upsilon(\boldsymbol{J}) \\ J_1 \\ 2J_2\Upsilon(\boldsymbol{J}) \\ 0 \end{pmatrix} \in \mathbb{T}^2 \times \mathbb{R}^4 \qquad (\mathbb{T}^2 = R^2/2\pi\mathbb{Z}^2), \tag{7}$$

$$\Omega_1(\boldsymbol{J}) = \omega_1 + \partial_2 \mathcal{Z}_r \left(\Upsilon(\boldsymbol{J}), J_1, 2J_2\Upsilon(\boldsymbol{J})\right), \tag{8}$$

$$\Omega_2(\boldsymbol{J}) = \omega_2 + J_2 + \partial_3 \mathcal{Z}_r \left(\Upsilon(\boldsymbol{J}), J_1, 2J_2\Upsilon(\boldsymbol{J})\right), \tag{9}$$

are solutions winding a two-parameter family of 2D-tori of the system (6) *with intrinsic (or basic) frequencies given by* (8) *and* (9)*.*

(ii) If, in addition, $b - \frac{d^2}{a} \neq 0$, the invariant tori of the first item are non-degenerate, in the sense that the matrix of the derivatives of $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$ with respect the parameters $\boldsymbol{J}^* = (J_1, J_2)$ is not singular at the origin, i. e., $\det\left(\partial_{\boldsymbol{J}} \boldsymbol{\Omega}(\boldsymbol{0})\right) \neq 0$.

In particular, the second item means that the frequencies Ω_1, Ω_2 in (8), (9) map diffeomorphically (locally at the origin) the space of the parameters into the space of frequencies, so the tori in the family could be described equally well using the frequencies $\boldsymbol{\Omega}$ instead of the parameters $\boldsymbol{J}$.

Remark 3. From the expression for the polar coordinates in (5), and the parametrization of the family of invariant tori follows that only those values of the parameters J_1, J_2 making $\Upsilon(J_1, J_2) > 0$ will determine real quasi-periodic solutions. As $\Upsilon(J_1, J_2)$ is obtained applying the implicit function theorem (section 1.9, proof of theorem 1.27) at $(0, 0)$, in particular, one can compute its expansion around the origin. Up to second order in J_1, J_2, results:

$$\Upsilon(J_1, J_2) = -\frac{d}{a}J_1 + \frac{1}{a}J_2^2 + \dots,$$

so real invariant tori are assured for $\boldsymbol{J}$ in a neighborhood of the origin if, for instance, the two additional conditions

$$-\frac{d}{a}J_1 > 0 \quad \text{and} \quad |J_2| \leq |J_1|^{\alpha},$$

with $\alpha > 1/2$ are simultaneously fulfilled. ♠

Before studying the linear character of the invariant tori, we note that, if through the change (5) we transform back to the normal form coordinates, the family of invariant tori given above is expressed as $\theta_1(t;\boldsymbol{J},\boldsymbol{\theta}^0) = \Omega_1(\boldsymbol{J})t + \theta_1^{(0)}$, $I_1 = J_1$ and,

$$\begin{aligned}
\boldsymbol{x}(t;\boldsymbol{J},\boldsymbol{\theta}^0) &= \begin{pmatrix} -\sqrt{\Upsilon(\boldsymbol{J})}\sin(\Omega_2(\boldsymbol{J})t+\theta_2^0) \\ \sqrt{\Upsilon(\boldsymbol{J})}\cos(\Omega_2(\boldsymbol{J})t+\theta_2^0) \end{pmatrix}, \\
\boldsymbol{y}(t;\boldsymbol{J},\boldsymbol{\theta}^0) &= \begin{pmatrix} -\sqrt{2J_2\Upsilon(\boldsymbol{J})}\sin(\Omega_2(\boldsymbol{J})t+\theta_2^0) \\ -\sqrt{2J_2\Upsilon(\boldsymbol{J})}\cos(\Omega_2(\boldsymbol{J})t+\theta_2^0) \end{pmatrix}.
\end{aligned}$$

Therefore, in the phase plane (x_1, x_2), the family of tori plots as a family of invariant circles of radii $(\Upsilon(\boldsymbol{J}))^{1/2}$ and centered at the origin where the periodic orbits are placed (analogously in (y_1, y_2)). In this sense, the (bifurcated) 2D-invariant tori appear "around" the periodic orbits.

Normal behavior of the invariant tori. Next, to study the linear stability of the family of invariant tori (7) one first sets up the variational equations of system (6) around one of the real torus of the family in the normal directions, say $\mathcal{T}(t;\boldsymbol{J},\boldsymbol{\theta}^0)$, for some $J \in \mathcal{U}$:

$$\begin{aligned}
\dot{X} &= 2\Upsilon(\boldsymbol{J})\,Y, \\
\dot{Y} &= -\left(\frac{2J_2^2}{\Upsilon(\boldsymbol{J})} + \partial^2_{1,1}\mathcal{Z}_r(\Upsilon(\boldsymbol{J}), J_1, 2J_2\Upsilon(\boldsymbol{J}))\right) X,
\end{aligned} \tag{10}$$

and sees that,

$$\varrho^{\pm}_{J_1,J_2} = \pm\sqrt{-4J_2^2 - 2\Upsilon(\boldsymbol{J})\partial^2_{1,1}\mathcal{Z}_r(\Upsilon(\boldsymbol{J}), J_1, 2J_2\Upsilon(\boldsymbol{J}))}$$

are its characteristic exponents, but expansion of the stuff inside the square root with respect to $\boldsymbol{J}$ yields:

$$\varrho^{\pm}_{J_1,J_2} = \pm\sqrt{2dJ_1 - 6J_2^2 + \ldots} \tag{11}$$

so, under the reality conditions in remark 3 the normal behavior of the invariant tori is determined –for $|J_1|$, $|J_2|$ sufficiently small–, just by the first term $2dJ_1$ inside the square root. Hence, suppose first that the coefficient a is positive. Then, under the reality condition dJ_1 must be negative, so by (11) the invariant tori will be elliptic and the periodic orbit with $I_1 = J_1$ is –linear–, unstable (see figure 1). Otherwise, let a be negative. Now the reality condition forces $dJ_1 > 0$, which implies hyperbolic tori and the periodic orbit surrounded (in the sense specified above) by the invariant tori is stable for $I_1 = J_1$. These considerations motivate the proposition 1.29 at the end of chapter 1, which we also write here.

Proposition 4. *Under the assumptions of theorem 2 –including the reality conditions of remark 3–, the type of the bifurcation is determined by the sign of the coefficient a. More precisely:*

Case 1. $a > 0$ *then, we say that the bifurcation is "direct": there appear elliptic tori around complex-unstable periodic orbits; and if*

Case 2. $a < 0$ *the bifurcation is "inverse": hyperbolic invariant tori unfold around stable periodic orbits. In this case, the family contains also parabolic and elliptic tori.*

In this way, we have determined the parameter, a, that governs both, the existence and the type of quasi-periodic bifurcation.

Chapter 2. Thus far the *formal* part of the memoir. Indeed, it is worth studying the whole Hamiltonian (normal form the plus remainder) and to inquire whether quasi-periodic solutions still persist (or not) after the remainder is added. So, one has to mess around with some type of KAM perturbation methods, but this implies some knowledge –or, at least, some hypothesis–, about how large the perturbation could be (because is this size which determines the order of the measure of the gaps). For us, two approaches are possible at this point: one can deal with the normal form as an integrable Hamiltonian and then add a generic perturbation –considered as small as one might need– or work in a more quantitative direction, asking, if R is the distance to the resonant periodic orbit, what could be the order, $r_{opt} = r_{opt}(R)$, of the normal form leading to the smallest remainder[(4)] in this R-neighborhood. Thus, one surely will be able to derive R-dependent "optimal" effective bounds, say $M(R)$, such that $\|\mathfrak{R}^{(r_{opt})}\| \leq M(R)$. In the analytical context, and assuming Diophantine frequencies, one typically expects to obtain exponentially small bounds for the remainder as a function of R. Something like:

$$M(R) \sim \mathrm{O}\left(\exp\left(-\frac{c_1}{R^{c_2}}\right)\right),$$

for certain $c_1, c_2 > 0$ (c_2 typically depends on the exponent of the Diophantine conditions). However, this works for semi-simple homological equations (even when resonant periodic orbits are dealt). In our case, though, the Jordan structure of the homological equations causes that when one solves for the coefficients associated to a monomial of certain degree, n, the associated small divisors appear not just raised up to the first power (as happens in the semi-simple case), but up to a power which depends on the degree n. Heuristically speaking, the amplification factor n^τ, ($\tau > 1$ independent of n) of the terms of degree n of the solutions of the homological equations, is replaced by a factor that can be guessed to be close to $n^{\tau n}$. The bounds obtained in the present resonant non semi-simple case are given in proposition 2.19 on page 81. We summarize below –omitting technicalities–, the main results described there:

Theorem 5. *If the frequencies* $\boldsymbol{\omega}^* = (\omega_1, \omega_2)$ *in the Hamiltonian H of theorem 1 satisfy the Diophantine conditions,*

$$|\langle \boldsymbol{k}, \boldsymbol{\omega}\rangle| \geq \frac{\gamma}{|\boldsymbol{k}|^\tau}, \qquad \boldsymbol{k} \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\},$$

for $\tau > 1$ *and for a certain* $\gamma > 0$*. Then:*

(4) or, at least, that optimizes some suitable bound of the remainder.

(i) There is $R_0 > 0$ such that, for a given distance R to the critical periodic orbit with $0 < R < R_0$, the optimal normalizing order –according to our bounds–, is given by $r_{opt} = \lfloor \check{r} \rfloor$, with $\check{r}$ depending on R through,

$$\check{r} = \mathrm{e}^{W_0\left(\ln\left(c_3 \frac{R}{R_0}\right)^{-1/c_4}\right)},$$

where $\mathrm{e} = \exp(1)$; c_3, c_4 are two constants independent of R and W_0 denotes the principal branch of the Lambert-W function –see section 2.6.1 and references therein[5]*–.*

(ii) The remainder of the normal form (theorem 1) is bounded by,

$$|||\mathfrak{R}^{(r_{opt})}||| \leq c_5 \left(1 - \frac{R}{R_0}\right)^{-1} \left(c_3 \frac{R}{R_0}\right)^{\frac{r_{opt}(R)}{2} - 1}, \qquad (12)$$

in this R-neighborhood of the critical periodic orbit. Here, c_5 is a constant independent of R.

(iii) Moreover, $\mathfrak{R}^{(r_{opt})}$ goes to zero faster than any analytic order, i. e.,

$$\mathfrak{R}^{(r_{opt})} = \mathrm{o}\left(\left(\frac{R}{R_0}\right)^n\right), \qquad (R/R_0 \to 0^+),$$

for any positive integer n.

Therefore, in a small enough neighborhood of the critical periodic orbit, the remainder, $\mathfrak{R}^{(r_{opt})}$, can be thought of as a perturbation of the normal form, and justifies the application of KAM methods in the next chapter. Of course, the setting of a very explicit constructive scheme to compute the normal form plays an essential rôle in the derivation of the bounds given above.

Chapter 3 is devoted to the discussion of the persistence of the bifurcated invariant tori derived in chapter 1. First, we note that we have different possible situations where persistence can be investigated: we can consider the case of direct bifurcation ($a > 0$) or the inverse ($a < 0$). Furthermore (and depending on the case) we can study the persistence of (real) *elliptic*, *parabolic* or *hyperbolic* tori. What we have done here is to study in detail the case of elliptic tori in the direct bifurcation. We have chosen the *elliptic* case because the context of elliptic tori is always the most difficult to deal with, and contains almost all the difficulties inherent to this (degenerate) problem. Likewise, we think that is important to stress here the main difficulties (and differences) of this problem with respect to others results of persistence of invariant tori (see Sevryuk, 1997; Pöschel, 1989) which make interesting by itself the methodology that we have followed in chapter 3. To discuss this, let us start giving a new parametrization of the unperturbed (i. e., those coming from the normal form) bifurcated families of 2D-tori. This parametrization will be more suitable if –as in our case– one wants to control the real character of the tori (see theorem 3.1 o page 98):

[5]Here, it is enough to known that for any $z \in \mathbb{R}$, $z > -1/\mathrm{e}$, is $w = W_0(z) \Leftrightarrow w\mathrm{e}^w = z$ with $w > -1$.

Theorem 6. *If the coefficient d in (2) is $d \neq 0$, then the function $\mathcal{T} : \mathbb{R} \times \Gamma \times \mathbb{T}^2 \to \mathbb{R}^6$, defined by*

$$\mathcal{T}(t; \xi, \eta; \boldsymbol{\theta}^{(0)}) = \big(\boldsymbol{\Omega}(\xi, \eta)t + \boldsymbol{\theta}^{(0)}, \xi, \mathcal{I}(\xi, \eta), 2\xi\eta, 0\big) \tag{13}$$

with: $\Gamma \subset \mathbb{R}^2$ a neighborhood of $(0,0)$, $\mathcal{I}$ an analytic function on Γ, defined implicitly by the equation

$$\eta^2 = \partial_1 \mathcal{Z}_r \left(\xi, \mathcal{I}(\xi, \eta), 2\xi\eta\right), \tag{14}$$

and the vector of frequencies, $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$ given by the components,

$$\begin{aligned} \Omega_1(\xi, \eta) &= \omega_1 + \partial_2 \mathcal{Z}_r(\xi, \mathcal{I}(\xi, \eta), 2\xi\eta), \\ \Omega_2(\xi, \eta) &= \omega_2 + \eta + \partial_3 \mathcal{Z}_r(\xi, \mathcal{I}(\xi, \eta), 2\xi\eta); \end{aligned} \tag{15}$$

constitutes a two-parameter family of solutions of (6) winding the corresponding family of two-dimensional invariant tori.

In terms of these new parameters, the characteristic exponents of the unperturbed family are given by,

$$\lambda_\pm(\xi, \eta) = \pm\sqrt{-4\eta^2 - 2a\xi - 2\xi\partial^2_{1,1}\mathcal{Z}_3(\xi, \mathcal{I}(\xi, \eta), 2\xi\eta)}, \quad (\xi, \eta) \in \Gamma, \ \xi > 0.$$

Then, the first difficulty arises when we have to choose the suitable set of parameters to characterize the tori of the family along the iterative KAM process. Let us mention that we have three frequencies to control: the two intrinsic ones, Ω_1, Ω_2 of the quasi-periodic motion and the normal one (the real part of λ_+), but just two parameters to control them ξ, η. So, we have to face the so called "lack of parameters" problem (see Moser, 1967; Sevryuk, 1999, or chapter 3 of this memoir for a more detailed explanation).

Typically, on applying KAM techniques for low-dimensional tori, one sets a diffeomorphism between some neighborhood of the origin in the parameter space (ξ, η) and a vicinity of (ω_1, ω_2) in the space of intrinsic frequencies (Ω_1, Ω_2) (similarly as stated in item (ii), theorem 2). Hence, the characteristic exponents $\lambda_\pm$ may be put also as a function of the intrinsic frequencies. For elliptic tori, besides the non-degeneracy of the these frequencies, one needs to ask the normal frequencies to "move" as a function of $\boldsymbol{\Omega}$, this forces to impose suitable "transversal" conditions in the denominators of the KAM process (see Sevryuk, 1999; Jorba and Villanueva, 1997a). In our case, for $\boldsymbol{\zeta} = (\xi, \eta)$ in a small neighborhood of the origin the invariant tori will be elliptic when $a > 0$ (and $\xi > 0$), as follows easily form the expression for $\lambda_\pm$. However, the typical transversal conditions,

$$\mathrm{Im}\left(\mathrm{grad}_{\boldsymbol{\Omega}}\langle \boldsymbol{\ell}, \boldsymbol{\lambda}(\boldsymbol{\Omega})\rangle|_{\boldsymbol{\Omega}=\boldsymbol{\omega}}\right) \notin \mathbb{Z}^2, \qquad \text{for any } \boldsymbol{\ell} \in \mathbb{Z}^2 \text{ with } 0 \le |\ell_1| + |\ell_2| \le 2, \ \ell_1 \neq \ell_2,$$

(where $\boldsymbol{\lambda}^*(\boldsymbol{\Omega}) = (\lambda_+(\boldsymbol{\Omega}), \lambda_-(\boldsymbol{\Omega})))$, does not work, because the derivatives of $\boldsymbol{\lambda}(\boldsymbol{\Omega})$ are not defined for $\boldsymbol{\Omega} = \boldsymbol{\omega}$ (the elliptic invariant tori are too close to parabolic). We have overcome this situation taking as *basic frequencies* for the unperturbed tori not $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$, the intrinsic frequencies, but $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$ with $\mu = |\lambda_+|$ and then, the first component of the intrinsic frequencies, Ω_1, as a function of $\boldsymbol{\Lambda}$, i. e.: $\Omega_1 = \Omega_1(\boldsymbol{\Lambda})$. In other words: we "label" the (elliptic) invariant tori with their normal frequency and second intrinsic frequency. It is checked that, with this parametrization, the small divisors do change in the normal directions, so one can proceed with to the KAM iterative scheme, which –due to the forementioned proximity of parabolic tori–, involves a more tricky control on the different terms of the Hamiltonian appearing at each successive step.

As we have implicitly mentioned, we will look for reducible elliptic tori (see J. Puig, 2002, for a survey on quasi-periodic reducibility), and hence the normal frequency will be well defined. So, in our approach we have not taken into account the machinery allowing to work with non-reducible elliptic tori (see Bourgain, 1997). A non-reducible approach complicates strongly the formulation, with no significative gain in the measure estimates. However, reducibility is a very important trick in order to simplify the resolution of the homological equations, which, in our case (and in spite of these simplifications) have some additional difficulties. To be more precise: the way we choose the basic frequencies, forced by the nondegeneracy conditions yields a coupling between some of the equations. This bind us to a very careful determination of the compatibility term which allows to keep the basic frequencies fixed at any step of the iterative process. On page 106 we state theorem 3.9, gathering the results related with the preservation of the (direct) bifurcated elliptic invariant tori. Here we present a shortened version:

Theorem 7. *Consider the Hamiltonian in theorem 1, assuming that the coefficient a of the quadratic part of the normal form (see (2)) is positive. Moreover, we define $\mu = |\lambda_+|$, $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$. Then, there exists a symplectic change:*

$$\Psi : \mathcal{D} \times \mathfrak{A} \subseteq \mathbb{T}^2 \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R} \times \mathcal{A} \longrightarrow \mathbb{T}^1 \times \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^2,$$

where $\mathfrak{A}$ is a Cantorian of the initial set of basic frequencies $\mathcal{A}$, such that:

(1) The measure of $\mathfrak{A}$ is plenty in the following sense: let $M(R)$ denote the bound for $\mathfrak{R}^{(r_{opt})}$ given in theorem 5 and $\mathcal{A}_R$ denote the subset of basic frequencies of $\mathcal{A}$ in a R-neighborhood of $\boldsymbol{\omega}^* = (0, \omega_2)$, then the Lebesgue measure:

$$meas\,(\mathcal{A}_R \setminus \mathfrak{A}) \sim (M(R))^{\frac{\alpha}{2}},$$

being $0 < \alpha < 1$, a fixed constant.

(2) The transformed Hamiltonian $\mathcal{H} = H \circ \Psi_{\boldsymbol{\Lambda}}$ can be cast into:

$$\mathcal{H}(\boldsymbol{\theta}, q, \boldsymbol{I}, p; \boldsymbol{\Lambda}) = \phi(\boldsymbol{\Lambda}) + \langle \boldsymbol{\Omega}(\boldsymbol{\Lambda}), \boldsymbol{I}\rangle + \frac{1}{2}\langle \boldsymbol{z}, \mathcal{B}(\boldsymbol{\Lambda})\boldsymbol{z}\rangle + \\ + \frac{1}{2}\langle \boldsymbol{I}, \mathcal{C}(\boldsymbol{\theta}; \boldsymbol{\Lambda})\boldsymbol{I}\rangle + \langle \boldsymbol{z}, \mathcal{E}(\boldsymbol{\Lambda})\boldsymbol{I}\rangle + \mathcal{H}_*(\boldsymbol{\theta}, q, \boldsymbol{I}, p; \boldsymbol{\Lambda}), \quad \boldsymbol{\Lambda} \in \mathfrak{A}, \quad (16)$$

where $\boldsymbol{\Omega}^*(\boldsymbol{\Lambda}) = (\Omega_1(\boldsymbol{\Lambda}), \Omega_2)$; $\mathcal{B}$, $\mathcal{E}$, $\mathcal{C}$ are 2×2 matrices ($\mathcal{B}$, $\mathcal{E}$ depending only on $\boldsymbol{\Lambda}$ whereas the matrix $\mathcal{C}$ depends also on $\boldsymbol{\theta}$), and $\mathcal{H}_*$ holds the terms of order greater than two in $\boldsymbol{z}$, $\boldsymbol{I}$.

(3) For every $\boldsymbol{\Lambda}$, the corresponding Hamiltonian (16) has a (reducible) invariant torus at $\boldsymbol{z} = 0$, $\boldsymbol{I} = 0$, with vector of intrinsic frequencies $\boldsymbol{\Omega}(\boldsymbol{\Lambda})^* = (\Omega_1(\boldsymbol{\Lambda}), \Omega_2)$, and normal frequency given by μ.

(4) $(\boldsymbol{\theta}, \boldsymbol{\Lambda}) \in \mathbb{T}^2 \times \mathfrak{A} \mapsto \Psi(\boldsymbol{\theta}, 0, \mathbf{0}, 0, \boldsymbol{\Lambda})$ is a parametrization of a Whitney regular Cantorian manifold holding the family of invariant tori, which can be embedded in a C^∞ regular manifold, in such a way that the measure of the extension of the Cantorian manifold to this regular manifold, is of the same order than the measure of the gaps coming from the elimination of frequencies in the KAM process.

In appendix A, we gather some auxiliary lemmas used to prove the different results spread along the text and finally, appendix B includes some basic background on Hamiltonian systems, stability of periodic orbits and transformation theory (Lie method series). Longer than our initial purpose was, it is included only to make this work as self contained as possible.

Chapter 1

An analysis using normal forms

In this chapter, we issue the formal part of our study. The objective is to analyze the dynamics around the $1 : -1$ resonance which appears when a family of periodic orbits changes its stability from elliptic to a complex hyperbolic saddle passing through degenerate elliptic.

Our analytical approach consists of computing, up to some given arbitrary order r, the *formal* normal form around that resonant (or *critical*) periodic orbit. This involves the following steps.

Step 1. We change the system of coordinates to a suitable one: the *adapted coordinates* around the resonant periodic orbit by means of a symplectic change.

Step 2. We apply a canonical Floquet transformation to reduce the normal variational equations of the orbit to constant coefficients.

Step 3. We complexify the Hamiltonian, making a linear complex change plus an extension of the transformed Hamiltonian to a complex domain.

Step 4. Finally, we proceed with the nonlinear reduction and describe, in some tricky and constructive way, the normal form.

Dealing with the truncated normal form itself and the differential equations related to it, the following results arise.

Result 1. We derive the existence of two families of invariant 2D tori which bifurcate from the critical orbit.

Result 2. We identify the coefficient of the normal form that determines the linear stability of the bifurcated tori. This allows us to show the Hopf-like character of the unfolding: normal hyperbolic tori unfold "around" elliptic orbits while elliptic tori appear around hyperbolic orbits.

The (formal) analysis of this chapter is followed by: (i) As the normal form at all order is not convergent in general, if we want to apply some perturbative technique, it is necessary to have some account of the smallness of that part of the Hamiltonian which has not been reduced to normal form. This is done in chapter 2. Our goal there is to prove that, if the order r of the normal form is conveniently taken as a function of the distance R to the resonant periodic orbit, then the norm of the remainder goes to zero as R does, but faster than any analytic order. (ii) in chapter 3, an application of KAM method –adapted to low dimension tori–, shows that most (in measure sense) of the tori of the 2D family still persists when the whole Hamiltonian is considered.

1.1 Overview of the chapter

Let us sketch, more detailed, the contents of this chapter 1. After the statement of the problem and two examples of transitions stable complex-unstable (see section below and section 1.3), we introduce (local) adapted coordinates around the resonant (degenerated, transition, critical) periodic orbit. The purpose of this change is to separate the dynamics along the periodic orbit, described now by and angle θ and its conjugate action I from the movement in the normal directions, which will be accounted by the positions ξ_1, ξ_2 and their corresponding momenta η_1, η_2. In the next section Floquet or *linear* reduction is applied. The final goal is thus to arrive –through a linear symplectic 2π-periodic change– to a "clean" Hamiltonian in the sense that its quadratic part $\mathcal{H}_2$ (see (1.5.21)) does not depend on the angle variable θ. Moreover the transformations are chosen to put this quadratic part in Williamson normal form (with respect to the normal directions). See the appendix 6 of Arnol'd (1974) and the references to the works of Galine and Williamson given there [(1)]. Albeit not strictly necessary, the linearly reduced Hamiltonian is still complexified (section 1.6) to simplify the structure of the homological equations arising in the nonlinear normalization (normal form) process, which begins in section 1.7. There, we introduce the Giorgilli-Galgani algorithm (definition 1.7). This will be canonical transformation device used to derive the normal form. From its recursive structure, we determine the form of the homological equations and, heuristically, the reduction algorithm. Nevertheless (and hence the "heuristicness"), it is *assumed* for such algorithm to success that, at any order $s = 3, 4, \ldots$, one can solve the corresponding homological equations in both the generating function G_s and its corresponding compatibility (complementary, resonant) term Z_s (see equations (1.7.11)). Such "solvability" question is answered affirmatively in theorem 1.20, which gives the form of the resonant terms Z_s. Our goal is to derive its proof in a constructive way, we mean, giving explicit (computable order by order) expressions for the coefficients of G_s and Z_s. In fact, this discussion constitutes the stuff of sections 1.7.1 and 1.7.2 whereas theorem 1.20 is presented as a way of conclusion at the end. Therefore, the formerly guessed reduction algorithm is fully justified and stated in section 1.7.3 by proposition 1.21 which, as noted in the text, is formally identical to the corresponding ones in Giorgilli et al. (1989) and Simó (1989). However, and for the sake of completeness, its proof has been also included in this memoir. At the end of the same section, theorem 1.24 gives the versal normal form at which the (initial) Hamiltonian can be reduced in a neighborhood of the resonant periodic orbit. Actually, it can be casted into this normal form and a remainder –see (1.7.67) and (1.7.68) for the real and complex normal form respectively–. So far the (formal) normalization computations. From this point, and up to the end of the chapter, we study the dynamics of the normal form itself, i. e., without taking the remainder into account. In particular one sees that, outside a zero

(1) In fact, in the classification for the normal forms of quadratic Hamiltonians given in the appendix of the Arnol'd book, the one corresponding to our case should be $H = \pm\frac{1}{2}(\frac{1}{b^2}q_1^2 + q_2^2) - b^2 p_1 q_2 + p_2 q_1$, but the trivial symplectic change:

$$q_1 = -by_1, \qquad p_1 = \frac{1}{b}x_1,$$
$$q_2 = \epsilon y_2, \qquad p_2 = -\epsilon x_2,$$

(with $\epsilon = \pm 1$) will –after identifying $b \equiv \omega_2$ and changing, if necessary, the sign of the time $t \mapsto \epsilon t$–, yield the desired form $\mathcal{H}_2$ given in (1.5.21).

measure set, the normal form is integrable, for three functionally independent integrals –equations (1.8.4)– are found. Further, it can be seen that (locally around the critical periodic orbit) the family of periodic orbits are, in a quite natural way, parametrized by the action I_1. With this parametrization, we set up the variational equations around the periodic orbits and derive explicit expressions for their characteristic exponents. The transition from stable to complex instable may then be easily described in terms of the expansion of these exponents with respect I_1 (see figure 1.7). This is done in section 1.8.1. In section 1.9 we seek for quasi-periodic solutions. In fact one realizes that these solutions can be more easily seen if the polar symplectic change (1.9.1), which introduces a new angle and a new conjugated action, is first applied. From the Hamiltonian differential equations in this new coordinates, the existence of a two-parametric 2D-family of invariant tori are derived and stated in theorem 1.27, which also gives a parametrization of the family. The next step is, with that parametrization, to set up the variational equations around the tori, compute their characteristic exponents and discuss their normal linear behavior. This discussion is formalized in proposition 1.29, which establish the linear character (elliptic, hyperbolic) of the bifurcated tori. The similarities from the present unfolding of invariant tori "around" (see figure 1.9) the family of periodic orbits and the classical Andronov-Hopf bifurcation can be appreciated from this last result. Finally a study of the dynamics of the fourth order normal form closes the chapter.

1.2 Formulation of the problem

Let $H(\boldsymbol{\zeta})$ with $\boldsymbol{\zeta}^* = (\xi_1, \xi_2, \xi_3, \eta_1, \eta_2, \eta_3)$, be a real three degree of freedom analytic Hamiltonian, with its associated Hamiltonian system

$$\dot{\boldsymbol{\zeta}} = J_3 \operatorname{grad} H(\boldsymbol{\zeta}), \tag{1.2.1}$$

where J_3 is the matrix of the standard canonical 2-form in $\mathbb{R}^6$ (see appendix B, section B.1). Suppose that this system has a nondegenerate family of periodic orbits, $\{\mathcal{M}_\sigma\}_{\sigma\in\mathbb{R}}$, such that for some value of the parameter of the family σ, say $\sigma = 0$, the corresponding orbit $\mathcal{M}_0$ (from now on, the *critical* or *resonant* periodic orbit), has an irrational (in the sense to be fixed later) collision of its nontrivial Floquet (characteristic) multipliers.

To be more precise, we suppose that, for $\sigma < 0$, those nontrivial characteristic multipliers of $\mathcal{M}_\sigma$ lie on the unit circle, they approach pairwise as σ goes to $\sigma = 0$, for this value they collide and separate towards the complex plane when $\sigma > 0$. This evolution is plotted in figure 1.2.

It remains to precise what is meant when we say that the collision of characteristic multipliers on the unit circle is irrational. In few words: if $2\pi\nu_0$ is the characteristic exponent corresponding to the characteristic multiplier λ_0 of the resonant periodic orbit $\mathcal{M}_0$ (and so $\lambda_0 = e^{2\pi\, i\nu_0}$), the collision is irrational if $\nu_0 \notin \mathbb{Q}$.

Remark 1.1. When this is not so and $\nu_0 = p/q \in \mathbb{Q}$ (and hence the characteristic exponent is commensurable with 2π) then, generically, phenomena of multiple-period bifurcation take place. It can be easily described using symplectic maps. For instance, if we consider a one parameter symplectic map, say $F_\sigma : \mathbb{R}^4 \to \mathbb{R}^4$; q-periodic points $\boldsymbol{x}_0$,

$$F_\sigma^q(\boldsymbol{x}_0) = \boldsymbol{x}_0$$

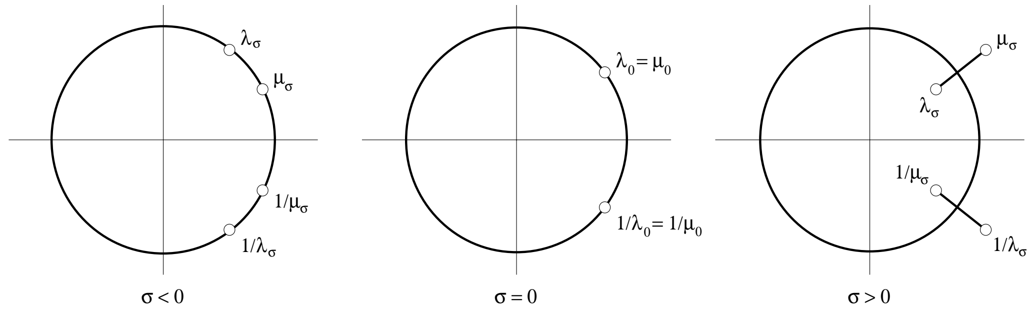


Figure 1.1: The transition from linear stability to complex instability for the family of periodic orbits $\{\mathcal{M}_\sigma\}_{\sigma\in\mathbb{R}}$, takes place through a collision of the nontrivial (i. e. different from 1) eigenvalues of the monodromy matrix corresponding to $\mathcal{M}_0$.

unfold. A general theory for the bifurcation of period-q points can be found in Bridges and Furter (1993). As an example, we mention Pfenniger's paper (1985a), where families of q-periodic points bifurcating from a resonant fixed point are found numerically for two generalizations of the standard Froeschlé map (see the example of section 1.3.2)[2].

Remark 1.2. Assuming genericity of the collision, the monodromy matrix of the orbit $\mathcal{M}_0$, $Y(\mathcal{M}_0)$, has the following Jordan normal form $\widetilde{Y}$,

$$\widetilde{Y} = \left(\begin{array}{cc|cc|cc} 1 & 0 & & & & \\ 1 & 1 & & & & \\ \hline & & \lambda_0 & 0 & & \\ & & 1 & \lambda_0 & & \\ \hline & & & & 1/\lambda_0 & 0 \\ & & & & 1 & 1/\lambda_0 \end{array}\right) \qquad (1.2.2)$$

Thus, none of the Jordan blocks of the monodromy matrix at the resonance is trivial (e. g. diagonal). In particular, the nontrivial character of the first block –corresponding to the eigenvalue equal to 1–, follows from the nondegeneracy of the family of periodic orbits (this is equivalent to the "twist" condition on the family of periodic orbits. If $\omega(\sigma)$ is the frequency of the orbit $\mathcal{M}_\sigma$, then we have that $\omega'(0) \neq 0$).

1.3 Complex instability: two examples

The instabilization process just described is known in the literature as the transition form *stability* to *complex instability*, and there are several Hamiltonian systems and symplectic maps where it takes place. In the present section we shall account for two representative examples.

1.3.1 The Restricted three body problem

We consider two bodies –called *primaries*–, moving on circular Newtonian orbits around their common center of masses. To this two body problem, we add a massless particle

[2] The one parameter symplectic map could be a Poincaré map around the critical periodic orbit using the energy as parameter.

whose movement is influenced by, but does not change the orbit of the primaries. The motion of this third particle is the widely study *Restricted Three Body Problem* (see Szebehely, 1967). To simplify the equations, the units of length, time and mass are chosen such that the sum of their masses, their mutual distance and the gravitational constant are equal to one. With these units,the period of the orbit of the primaries is 2π (so the angular velocity is equal to one). Moreover, in the RTBP it is common the use of rotating or *synodical* coordinates, where the origin is fixed at the center of mass of the two primaries, the x axis is given by the line defined by the two primaries and oriented from the smaller of the primaries to the bigger one. The z axis has the direction of the angular momentum of the motion of the primaries and the y axis is chosen to form a positively oriented system of reference.

If μ and $1-\mu$ with $0 < \mu \leq 1/2$ (called the *mass parameter*), are the masses of the primaries, in the synodic system, their positions are given respectively by $(\mu-1,0,0)$ and $(\mu,0,0)$. In this system, the Hamiltonian for the motion of the third particle is

$$H(x,y,z,p_x,p_y,p_z) = \frac{1}{2}(p_x^2+p_y^2+p_z^2) + yp_x - xp_y - \frac{1-\mu}{r_1} - \frac{\mu}{r_2}, \tag{1.3.1}$$

where the momenta $p_x = \dot{x} - y$, $p_y = \dot{y} + x$ and $p_z = \dot{z}$ and r_1, r_2 are the distances from the particle to the primaries: $r_1^2 = (x-\mu)^2 + y^2 + z^2$, $r_2^2 = (x-\mu+1)^2 + y^2 + z^2$. Note that then, the (x,y) plane is invariant by the flow, so we can restrict the Hamiltonian to this plane. This restriction is known as the *planar* RTBP. Through all this section only the spatial case is considered, so we shall refer (1.3.1) simply as the RTBP.

Its associated Hamiltonian system has five equilibrium points. Three of them located on the x axis (the *collinear* or Euler L_1, L_2 and L_3). The other two (the *triangular* or Lagrange equilibrium points L_4, L_5) form two equilateral triangles with the primaries; their coordinates in the phase space are $L_4(-\frac{1}{2}+\mu, -\frac{\sqrt{3}}{2}, 0, \frac{\sqrt{3}}{2}, -\frac{1}{2}+\mu, 0)$, and $L_5(-\frac{1}{2}+\mu, \frac{\sqrt{3}}{2}, 0, -\frac{\sqrt{3}}{2}, -\frac{1}{2}+\mu, 0)$.

In what follows we shall study the stability of the Lyapunov vertical family of orbits around L_5 (see below), but due to the symmetries of the problem, the same results will be valid for L_4. It is convenient, first, to bring the origin to L_5. This motivates the translation,

$$\begin{aligned} x &= X - \frac{1}{2} + \mu, & y &= Y + \frac{\sqrt{3}}{2}, & z &= Z \\ p_x &= P_X - \frac{\sqrt{3}}{2}, & p_y &= P_Y - \frac{1}{2} + \mu, & p_z &= P_Z. \end{aligned}$$

With this translation, the coordinates of L_5 are $(0,0,0,0,0,0)$ and the transformed Hamiltonian (1.3.1) is

$$\begin{aligned} \widehat{H}(X,Y,Z,P_X,P_Y,P_Z) &= \frac{1}{2}\left(P_X^2+P_Y^2+P_Z^2\right) + \\ &+ YP_X - XP_Y + \left(\frac{1}{2}-\mu\right)X - \frac{\sqrt{3}}{2}Y - \frac{1-\mu}{R_1} - \frac{\mu}{R_2}, \end{aligned} \tag{1.3.2}$$

now with,

$$R_1^2 = \left(X - \frac{1}{2}\right)^2 + \left(Y + \frac{\sqrt{3}}{2}\right)^2 + Z^2,$$
$$R_2^2 = \left(X + \frac{1}{2}\right)^2 + \left(Y + \frac{\sqrt{3}}{2}\right)^2 + Z^2.$$

The quadratic part of the Hamiltonian (1.3.2) is:

$$\widehat{H}_2 = \frac{1}{2}\left(P_X^2 + P_Y^2 + P_Z^2\right) + YP_X - XP_Y + \frac{1}{8}X^2 - \\ - \frac{5}{8}Y^2 + \frac{1}{2}Z^2 + \frac{\sqrt{27}}{2}\left(\frac{1}{2} - \mu\right)XY + \frac{1}{2}\mu(1-\mu) - \frac{3}{2}, \quad (1.3.3)$$

and the corresponding linearized system is $\dot{Z} = \mathfrak{G}Z$, being $\mathfrak{G}$ the constant matrix:

$$\mathfrak{G} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1/4 & \gamma & 0 & 0 & 1 & 0 \\ \gamma & 5/4 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

with $\gamma = -\frac{\sqrt{27}}{2}\left(\frac{1}{2} - \mu\right)$. Computation of $\operatorname{Spec}\mathfrak{G}$ gives the eigenvalues,

$$\lambda_1 = \mathrm{i}, \qquad \lambda_{2,3} = \sqrt{-\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - 27\mu(1-\mu)}},$$

and $\lambda_{j+3} = -\lambda_j, j = 1, 2, 3$; since $\mathfrak{G}$ is an infinitesimal symplectic matrix.

The pair $\pm\mathrm{i}$ gives rise to vertical oscillations with angular frequency equal to 1. The others are the ones of the planar RTBP. All the eigenvalues are purely imaginary and different if $0 < \mu < \mu_R = \frac{1}{2}(1 - \sqrt{23/27}) \approx 0.03852$, (the *Routh's mass parameter*). For $\mu = \mu_R$ the planar frequencies collide on the imaginary axis. This produces a change in the linear stability and for $\mu_R < \mu \leq 1/2$, L_5 becomes unstable.

Since λ_1 is purely imaginary and since for $0 < \mu < 1/2$ the quotients $\frac{\lambda_2}{\lambda_1}, \frac{\lambda_3}{\lambda_1}$ are never integers, the Liapunov's center theorem (see Siegel and Moser, 1971, chapter 2, §16), assures the existence of a family of periodic orbits, depending analytically on a real parameter: the so called vertical family of L_5 (and the orbits of this family can be locally parametrized by the vertical amplitudes).

A point to note is that if $\mu \neq \mu_R$, and at least for small vertical amplitudes, the linear stability of the family is the same as L_5. Nevertheless, the linear character may change for large enough amplitudes. Actually, we can identify L_5 with the periodic orbit of period 2π and zero amplitude so the monodromy matrix of this orbit is given by $\mathfrak{H} = \exp(2\pi\mathfrak{G})$. Of course, this matrix has two (trivial) eigenvalues equal to one (those which come from the pair $\pm i$ of $\mathfrak{G}$). By the considerations above, when $0 < \mu < \mu_R$, the four nontrivial eigenvalues of the corresponding $\mathfrak{H}$ lie on the unit circle and collide when $\mu = \mu_R$. We can thus continue (numerically) this resonance with respect to μ and the amplitude of the orbit. In fact, it be can continued with respect to any regular parameter of the family. Figure 1.2 shows the curve corresponding to this resonance. The parameters plotted are μ and the vertical velocity $\dot{z}$ of the periodic orbit when it cuts the hyperplane $z = 0$ in the positive direction (see Jorba and Villanueva, 1998).

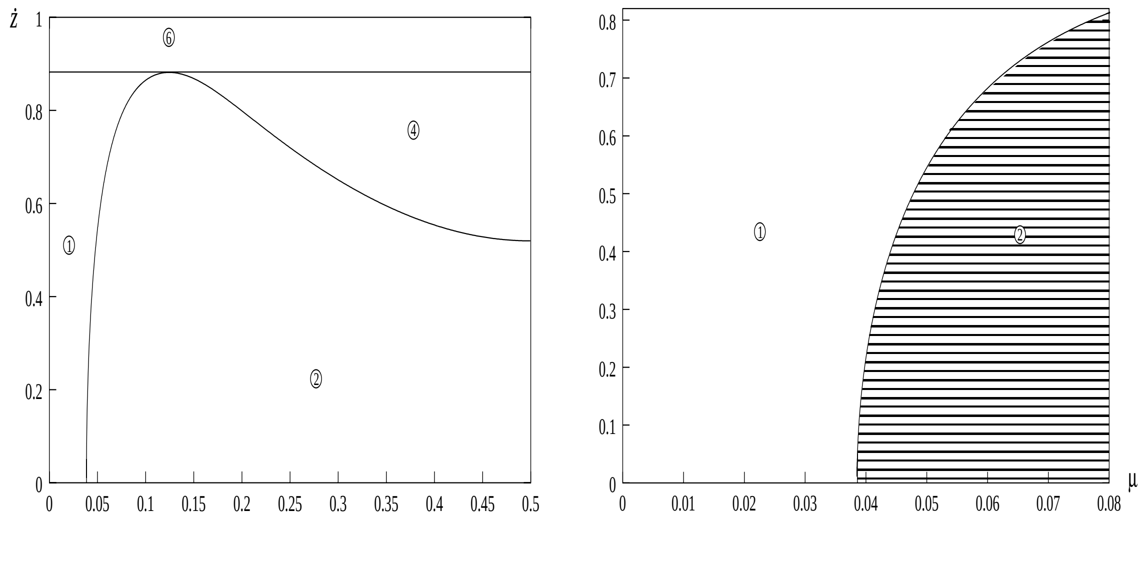
As an example, we fix $\mu = 0.05$ ($> \mu_R$, so L_5 is complex unstable) and compute the vertical family for this particular value. To do this, we take the hyperplane $z = 0$ as the surface of section, fix an energy level corresponding to a small amplitude and search for a fixed point of the reduced Poincaré map (see section B.2.1, appendix B). This can be done first approximating by the solutions of the linearized system (which are known explicitly) and then refining by a modified Newton method. The arc step method (see Gómez et al., 1993) is applied to continue numerically the fixed point with respect to the energy.

For each orbit in the family, we have computed also the associated stability indices b_1, b_2 (see section B.2.2 of appendix B) and represented them with respect to $\dot{z}$ (as before, the positive vertical velocity at $z = 0$). This is done for a range $0 \leq \dot{z} \leq 1$ in figure 1.3(a). Note: the index c_1 which appears therein is defined from b_1, b_2 through $c_1 = \frac{1}{2}(b_1 + b_2)$, (see remark B.19, appendix B).

On the other hand, figure 1.3(b) shows the path traced by the family in the Broucke diagram (appendix B, section B.2.2). Therein, we select six orbits –numbered from 1 to 6–, and represent them in figures 1.4, 1.5 and 1.6. Their initial conditions and period are specified at table 1.1. From both figures 1.3(a) and 1.3(b), it can be seen that a change complex-instability to stability takes place (the critical orbit is marked with number '3'). Another transition to even semi-instability can also be appreciated for higher values of $\dot{z}$.

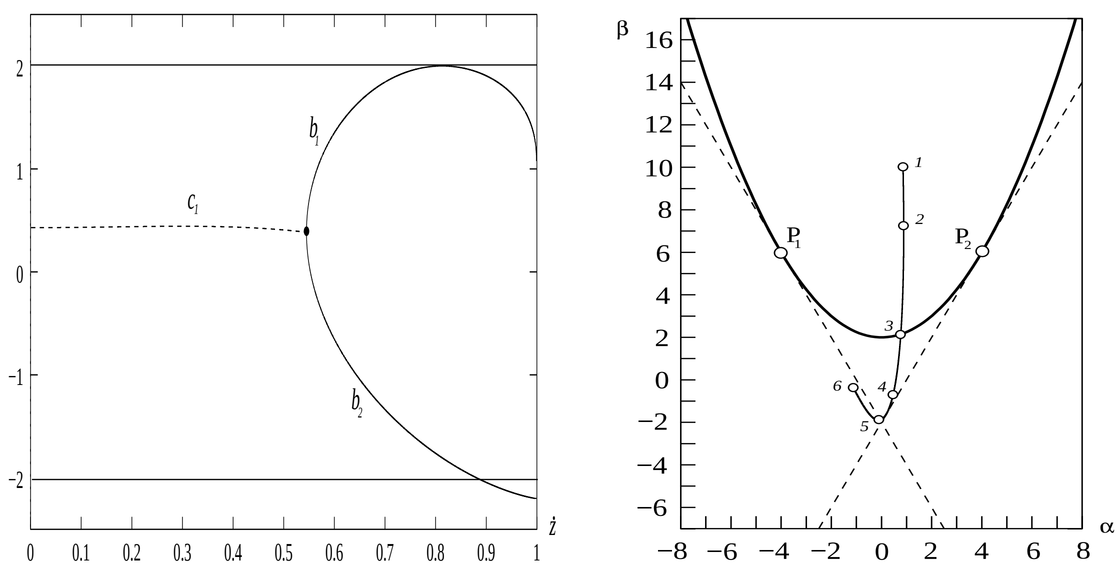
#	T	x	y
1	0,628 318 536 168 +01	−0,450 000 109 572 +00	0,866 025 333 494 +00
2	0,628 777 272 368 +01	−0,458 924 515 931 +00	0,860 261 746 326 +00
3	0,630 089 232 495 +01	−0,481 245 305 933 +00	0,845 561 510 888 +00
4	0,631 599 927 537 +01	−0,501 046 769 737 +00	0,832 415 959 636 +00
5	0,633 840 714 636 +01	−0,516 359 489 344 +00	0,823 417 455 682 +00
6	0,637 047 835 699 +01	−0,469 443 622 028 +00	0,860 308 810 893 +00
#	p_x	p_y	p_z
1	−0,866 024 911 750 +00	−0,449 999 890 428 +00	0,999 999 457 892 −03
2	−0,825 060 390 352 +00	−0,440 138 145 163 +00	0,286 536 941 211 +00
3	−0,711 590 543 355 +00	−0,404 832 774 245 +00	0,544 945 235 936 +00
4	−0,583 483 428 082 +00	−0,350 428 498 869 +00	0,717 405 223 842 +00
5	−0,383 984 492 470 +00	−0,237 527 444 586 +00	0,886 984 840 039 +00
6	0,161 132 747 195 −03	0,136 748 575 021 −01	0,999 044 666 550 +00

Table 1.1: Initial conditions and period of the periodic orbits marked by numbers $1, \ldots, 6$ in the Broucke diagram of figure 1.3(b). T is the period and the initial value of the vertical position is $z = 0$, so it is not listed in the table



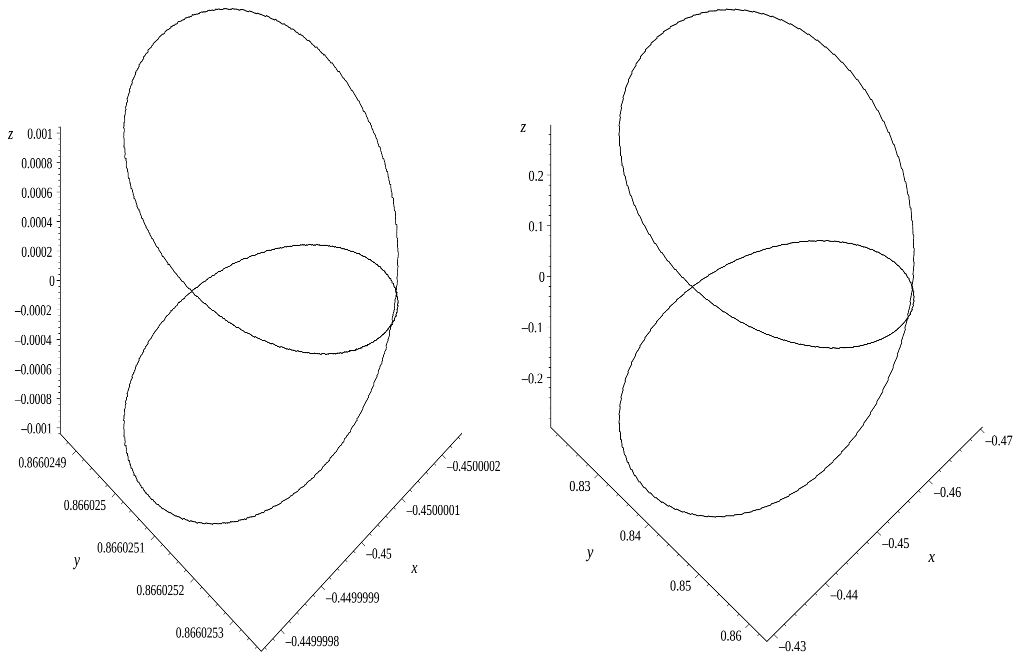
(a) (b)

Figure 1.2: 1.2(a) Change of the linear character of the orbits of the vertical family at L_5. Figure 1.2(b) represents in more detail the neighborhood of the critical value $\mu = \mu_R$. In both figures, the mass parameter is plotted on the horizontal axis, while on the vertical axis $\dot{z}$ (the positive vertical velocity when $z = 0$) is represented. The numbers enclosed by circles denote the stability types according to the section B.2.2 of appendix B (figures B.3 and B.4).



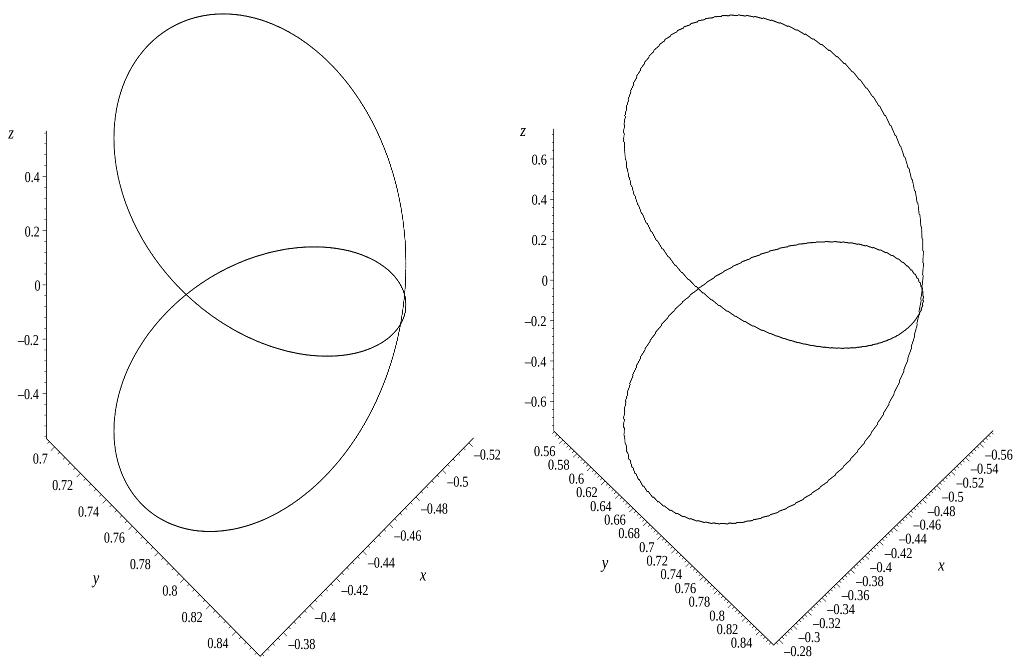
(a) (b)

Figure 1.3: 1.3(a) Stability indices of the Liapunov vertical family of L_5 and corresponding to a value of $\mu = 0.05$ of the mass parameter. In 1.3(b) the path of the family in the Broucke diagram (appendix B, section B.2.2) is shown. See explanation at the text.



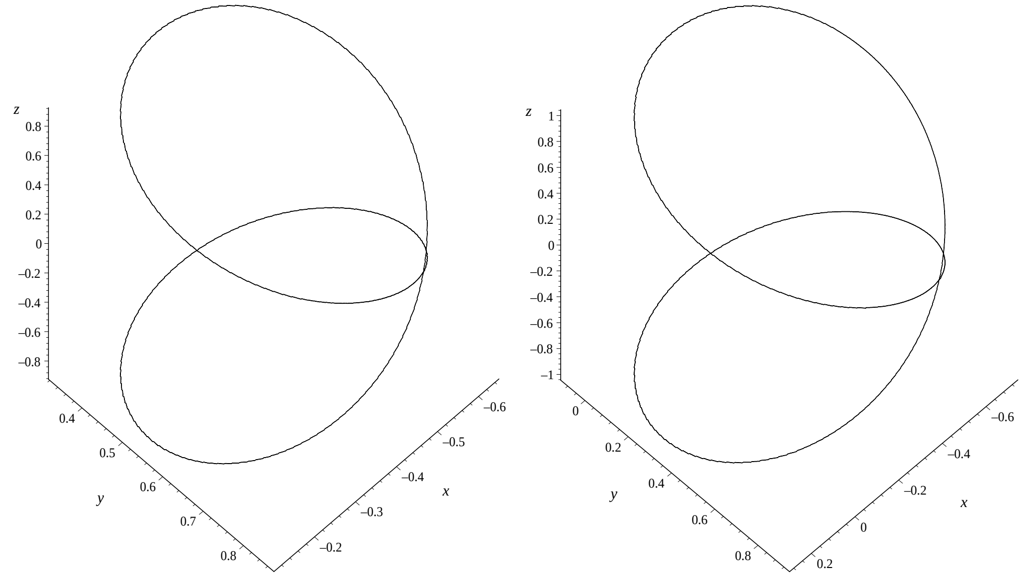
(a) Orbit # 1 (b) Orbit # 2

Figure 1.4: Two complex unstable periodic orbits. They correspond to the initial conditions 1 and 2 listed at the table above. In the Broucke diagram of figure 1.3(b) they are pointed with their number besides white dots (◯).



(a) Orbit # 3 (b) Orbit # 4

Figure 1.5: Orbits 3, 4, in the Broucke diagram (figure 1.3(b)). In particular 3 is the critical orbit, so the coefficients (α, β) of the characteristic polynomial of its monodromy matrix represented there lie on the arch $P_1 P_2$.



(a) Orbit # 5 (b) Orbit # 6

Figure 1.6: Orbits 5, and 6 in the Broucke diagram (figure 1.3(b)). Note, by comparison with the two previous figures (1.4 and 1.5) that the shape of the orbits remains the same, but their amplitude in the vertical direction increases notably along the family.

1.3.2 The Froeschlé's generalized mappings

Pfenniger (1985a), introduces the maps

$$T_s\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}x_1 + K\sin(x_1+x_2) + L\sin(x_1+x_2+x_3+x_4)\\ x_1+x_2\\ x_3 - L\sin(x_1+x_2+x_3+x_4)\\ x_3+x_4\end{pmatrix} \mod (2\pi)$$

and

$$T_t\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}x_1 + K\sin(x_1+x_2) + L\tan(x_1+x_2+x_3+x_4)\\ x_1+x_2\\ x_3 - L\tan(x_1+x_2+x_3+x_4)\\ x_3+x_4\end{pmatrix} \mod (2\pi),$$

which are symplectic, with respect to the two-form $\alpha^2 = dx_2 \wedge dx_1 + dx_3 \wedge dx_4$. By conjugation we can define two new mappings T_s' and T_t' as

$$T_s'(x) = S^{-1}T_s(Sx), \quad T_t'(x) = S^{-1}T_t(Sx);$$

being $S = \left(\begin{smallmatrix}0&0&0&1\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{smallmatrix}\right)$. These are symplectic mappings, now with respect to the standard two-form $\omega^2 = dx_1 \wedge dx_3 + dx_2 \wedge dx_4$. Explicitly, T'_s and T'_t are given by,

$$T'_s \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 - L\sin(x_1 + x_2 + x_3 + x_4) \\ x_2 + x_4 \\ x_1 + x_3 \\ x_4 + K\sin(x_2 + x_4) + L\sin(x_1 + x_2 + x_3 + x_4) \end{pmatrix} \mod (2\pi)$$

and

$$T'_t \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 - L\tan(x_1 + x_2 + x_3 + x_4) \\ x_2 + x_4 \\ x_1 + x_3 \\ x_4 + K\sin(x_2 + x_4) + L\tan(x_1 + x_2 + x_3 + x_4) \end{pmatrix} \mod (2\pi).$$

Both maps T'_s and T'_t have the same Jacobian matrix around the fixed point $x = 0$. We put $DT'_s(0) = DT'_t(0) = A$, and making the computations

$$A = \begin{pmatrix} 1-L & -L & -L & -L \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ L & K+L & L & 1+K+L \end{pmatrix}. \tag{1.3.4}$$

Nevertheless, their non-linear behavior are different (see the referenced paper of Pfenniger, and also the work of Ollé and Pfenniger, 1999).

From the characteristic polynomial of the matrix A,

$$p(\lambda) = \lambda^4 - (4+K)\lambda^3 + (6 + 2K - KL)\lambda^2 - (4+K)\lambda + 1,$$

we identify (section B.2.2) the coefficients α, β and the discriminant Δ,

$$\alpha = -(4+K), \quad \beta = 6 + K(2-L), \quad \Delta = K(K+4L).$$

The transition to complex instability requires $\Delta = 0$, so if we keep fixed the value of K, the critical value of the parameter L, $L_{crit} = -K/8$. Moreover, the transition stable-complex unstable corresponds to a point on the arch P_1P_2 in the Broucke diagram (figure B.3). This restricts K to the interval $-8 < K < 0$.

After these two examples, we go again to the general problem, as described in section 1.2.

1.4 An analytic approach

As a first step to proceed with our approach, we shall introduce (see Bruno, 1989, 1994, and references therein) local coordinates around the *critical* periodic orbit $\mathcal{M}_0$ through an analytic 2π-periodic change of variables,

$$\xi_i = \xi_i(\widetilde{\theta}, \widetilde{\boldsymbol{\xi}}, \widetilde{I}, \widetilde{\boldsymbol{\eta}}), \qquad \eta_i = \eta_i(\widetilde{\theta}, \widetilde{\boldsymbol{\xi}}, \widetilde{I}, \widetilde{\boldsymbol{\eta}}), \tag{1.4.1}$$

$i = 1, 2, 3$ and with $\widetilde{\boldsymbol{\xi}} = (\widetilde{\xi}_1, \widetilde{\xi}_2)$, $\widetilde{\boldsymbol{\eta}} = (\widetilde{\eta}_1, \widetilde{\eta}_2)$. Furthermore, we shall ask the change (1.4.1) to satisfy the following properties:

P1.4- 1. It maps the product set $\mathfrak{V} = \mathbb{T}^1 \times \Omega$, where Ω is a five-dimensional open set around the origin, onto some (possible small) neighborhood, $\mathfrak{U}$, of $\mathcal{M}_0$.

P1.4- 2. The orbit $\mathcal{M}_0$ is given by $\widetilde{\boldsymbol{\xi}} = \widetilde{\boldsymbol{\eta}} = 0$ and $\widetilde{I} = 0$ (and parametrized by $\widetilde{\theta}$).

P1.4- 3. The change (1.4.1) is symplectic with $\widetilde{\theta}$, $\widetilde{\boldsymbol{\xi}}$ and $\widetilde{I}$, $\widetilde{\boldsymbol{\eta}}$ the new conjugate positions and momenta respectively. So in this coordinates, the system (1.2.1) is transformed into another Hamiltonian system,

$$\begin{aligned} \dot{\widetilde{\theta}} &= \frac{\partial \widetilde{H}}{\partial \widetilde{I}}, & \dot{\widetilde{I}} &= -\frac{\partial \widetilde{H}}{\partial \widetilde{\theta}}. \\ \dot{\widetilde{\xi}}_i &= \frac{\partial \widetilde{H}}{\partial \widetilde{\eta}_i}, & \dot{\widetilde{\eta}}_i &= -\frac{\partial \widetilde{H}}{\partial \widetilde{\xi}_i}, \qquad i = 1, 2. \end{aligned} \tag{1.4.2}$$

For an explicit construction of *local* canonical coordinates such like the ones just described, (see Jorba and Villanueva, 1998).

The transformed Hamiltonian, $\widetilde{H}$, defined in $\mathfrak{V}$, is analytic and 2π-periodic in $\widetilde{\theta}$, so it can be expanded in a convergent Taylor series,

$$\widetilde{H}(\widetilde{\theta}, \widetilde{\boldsymbol{\xi}}, \widetilde{I}, \widetilde{\boldsymbol{\eta}}) = \sum_{k,\boldsymbol{l},\boldsymbol{m}} \widetilde{h}_{k,\boldsymbol{l},\boldsymbol{m}}(\widetilde{\theta}) \widetilde{I}^k \widetilde{\boldsymbol{\xi}}^{\boldsymbol{l}} \widetilde{\boldsymbol{\eta}}^{\boldsymbol{m}}, \tag{1.4.3}$$

with the standard multi-index notation $\widetilde{\boldsymbol{\xi}}^{\boldsymbol{l}} \widetilde{\boldsymbol{\eta}}^{\boldsymbol{m}} = \widetilde{\xi}_1^{l_1} \widetilde{\xi}_2^{l_2} \widetilde{\eta}_1^{m_1} \widetilde{\eta}_2^{m_2}$, which we shall use throughout the text. The index k, and the components of $\boldsymbol{l}^* = (l_1, l_2)$, $\boldsymbol{m}^* = (m_1, m_2)$ range over the nonnegative integers, while the coefficients $\widetilde{h}_{k,\boldsymbol{l},\boldsymbol{m}}(\widetilde{\theta})$ are analytic 2π-periodic functions and can be expanded in Fourier series. If we restrict the system (1.4.2) to the periodic orbit $\mathcal{M}_0$, and take into account the development (1.4.3), we get

$$0 = \widetilde{h}'_{0,\mathbf{0}}(\widetilde{\theta}), \quad \dot{\widetilde{\theta}} = \widetilde{h}_{1,\mathbf{0}}(\widetilde{\theta}), \quad 0 = \widetilde{h}_{0,\boldsymbol{e}_i}(\widetilde{\theta}), \quad i = 1, 2, 3, 4, \tag{1.4.4}$$

($\boldsymbol{e}_i$ is the i-th unit vector in $\mathbb{R}^4$ and $\mathbf{0}$ is the zero of $\mathbb{R}^4$), since by the condition *P1.4-2*, $\dot{\widetilde{\boldsymbol{\xi}}} = \dot{\widetilde{\boldsymbol{\eta}}} = \mathbf{0}$ and $\dot{\widetilde{I}} = 0$, on the periodic orbit $\mathcal{M}_0$.

From equations above, it follows that $\widetilde{h}_{0,\mathbf{0}}(\widetilde{\theta}) = \text{const.}$, so we can set $\widetilde{h}_{0,\mathbf{0}} = 0$.

Now, let $1/\omega_1$ be the mean value of $1/\widetilde{h}_{1,\mathbf{0}}$, therefore,

$$\frac{\omega_1}{\widetilde{h}_{1,\mathbf{0}}(\widetilde{\theta})} = 1 + \sum_{n=1}^{\infty} (a_n \cos n\widetilde{\theta} + b_n \sin n\widetilde{\theta}).$$

Note that $\widetilde{h}_{1,\mathbf{0}}(\widetilde{\theta}) \neq 0$. Since $\mathcal{M}_0$ is a periodic solution, there is no stationary points on it.

We introduce a new change, which involves only $\widetilde{I}$ and $\widetilde{\theta}$

$$I = \frac{\widetilde{I}}{\omega_1} \widetilde{h}_{1,\mathbf{0}}(\widetilde{\theta}), \quad \theta = f(\widetilde{\theta}), \tag{1.4.5}$$

where $f(\widetilde{\theta})$ is defined by integration of the quotient $\omega_1/\widetilde{h}_{1,\mathbf{0}}(\widetilde{\theta})$, i. e.

$$f(\widetilde{\theta}) = \int \frac{\omega_1}{\widetilde{h}_{1,\mathbf{0}}(\widetilde{\theta})} \, d\widetilde{\theta} = \widetilde{\theta} + \sum_{k=1}^{\infty} \frac{1}{k} (a_k \sin k\widetilde{\theta} - b_k \cos k\widetilde{\theta}).$$

It is straightforward to see that $dI \wedge d\theta = d\widetilde{I} \wedge d\widetilde{\theta}$, so the change (1.4.5) is canonical. Moreover, on $\mathcal{M}_0$, $\dot{\theta} = \omega_1$, so $\theta = \omega_1 t + \text{const.}$; but $\mathcal{M}_0$ is a periodic orbit with period $T_0 = T(\mathcal{M}_0)$, therefore we have $\omega_1 = 2\pi/T_0$. In other words, ω_1 is the angular frequency of the periodic orbit. Thus, if we expand the new transformed Hamiltonian, $\widehat{H}(\theta, I, \widetilde{\boldsymbol{\xi}}, \widetilde{\boldsymbol{\eta}})$, like in (1.4.3) and as before, restrict the corresponding Hamiltonian equations to the orbit $\mathcal{M}_0$, we shall obtain the following relations

$$\widehat{h}_{0,\mathbf{0}}(\theta) = 0, \quad \widehat{h}_{1,\mathbf{0}} = \frac{2\pi}{T_0}, \quad \widehat{h}_{0,\boldsymbol{e}_i}(\theta) = 0, \quad i = 1,2,3,4. \tag{1.4.6}$$

Here, $\widehat{h}_{k,\boldsymbol{l},\boldsymbol{m}}(\theta)$ are the 2π-periodic coefficients of the Taylor expansion of the Hamiltonian transformed by the change (1.4.5). In this way, we have determined the constant and linear terms of the Hamiltonian.

1.5 Linear normalization

Consider now the following low order terms of the expansion of the Hamiltonian:

$$\widehat{h}_{1,\mathbf{0}}(\theta)I + \sum_{|\boldsymbol{m}|_1+|\boldsymbol{l}|_1=2} \widehat{h}_{0,\boldsymbol{l},\boldsymbol{m}}(\theta)\widetilde{\boldsymbol{\xi}}^{\boldsymbol{l}}\widetilde{\boldsymbol{\eta}}^{\boldsymbol{m}} = \omega_1 \left(I + \frac{1}{2} < \widetilde{\boldsymbol{\zeta}}, \Gamma(\theta)\, \widetilde{\boldsymbol{\zeta}} > \right), \tag{1.5.1}$$

being $\widetilde{\boldsymbol{\zeta}} = (\widetilde{\boldsymbol{\xi}}, \widetilde{\boldsymbol{\eta}})$ and $\Gamma(\theta)$ a 4×4 matrix whose coefficients are real analytic 2π-periodic functions. Furthermore, the norm $|\boldsymbol{v}|_1 = \sum_i |v_i|$ has already been introduced.

To achieve the linear normalization, we skip the higher order terms in the Taylor development of the Hamiltonian and consider the normal variational equations around the orbit, which are given by the following linear Hamiltonian system,

$$\dot{\theta} = \omega_1, \quad \dot{I} = 0, \tag{1.5.2}$$

$$\dot{\widetilde{\boldsymbol{\zeta}}} = \omega_1 J_2 \Gamma(\theta)\, \widetilde{\boldsymbol{\zeta}}. \tag{1.5.3}$$

Combining the first of the equations (1.5.2), and the *normal* linear system (1.5.3), one obtains

$$\frac{d\widetilde{\boldsymbol{\zeta}}}{d\theta} = J_2 \Gamma(\theta)\, \widetilde{\boldsymbol{\zeta}}. \tag{1.5.4}$$

Let $X(\theta)$, be a fundamental matrix of the solutions of (1.5.4). Then

$$X(\theta + 2\pi) = X(\theta) M_0,$$

where M_0 is a constant nonsingular matrix which, apart from the block $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ of (1.2.2), has the same Jordan block structure than the monodromy matrix of the periodic orbit $\mathcal{M}_0$. Since (1.5.4) is a Hamiltonian system, M_0 is a (2π-periodic) canonical matrix to be fixed later. We introduce the following linear substitution,

$$\widetilde{\boldsymbol{\zeta}} = B(\theta)\boldsymbol{z},$$

$\boldsymbol{z}^* = (\boldsymbol{x}^*, \boldsymbol{y}^*)$, with $B(\theta)$ a canonical and 2π-periodic on θ matrix to be determined later. Therefore, the system (1.5.4) transforms to another linear Hamiltonian system:

$$\frac{d\boldsymbol{z}}{d\theta} = J_2 \left(B^* \Gamma B + B^* J_2 B' \right) \boldsymbol{z}, \tag{1.5.5}$$

here, the asterisk, (*), denotes the transposed matrix. Now, let us assume the following hypothesis (discussed below):

H1.5- 1. A constant symmetric matrix, N_0, exists such that M_0 admits an exponential representation of type,

$$M_0 = \exp(2\pi J_2 N_0), \tag{1.5.6}$$

H1.5- 2. There exists a constant canonical matrix D, and a matricial normal form G, to be chosen later such that the matrix N_0 can be expressed as a product like,

$$N_0 = D^* G D.$$

Assuming *H1.5-1* and *H1.5-2*, we take for $B(\theta)$,

$$B(\theta) = X(\theta) D^{-1} \exp(-\theta J_2 G), \tag{1.5.7}$$

and it is easy to check both, the 2π-periodicity of $B(\theta)$ and that the transformed system (1.5.5) turns out to be,

$$\frac{d\boldsymbol{z}}{d\theta} = J_2 G \boldsymbol{z}.$$

Next, we come back to the Hamiltonian $\widehat{H}$ and perform the transformation,

$$\theta = \theta_1, \quad I = I_1 + \frac{1}{2}\langle \boldsymbol{z}, B^* J_2 B' \boldsymbol{z}\rangle,$$
$$\widetilde{\boldsymbol{\zeta}} = B(\theta_1)\boldsymbol{z},$$

which can be checked out to be canonical and 2π-periodic in θ. Here, $B(\theta)$ is the matrix given by (1.5.7). Direct substitution shows that the quadratic part (1.5.1) of $\widetilde{H}$ transforms to

$$\begin{aligned} \widetilde{\mathcal{H}}_2(\theta_1, I_1, \boldsymbol{z}) &= \omega_1 I_1 + \frac{\omega_1}{2}\langle \boldsymbol{z}, (B^* \varGamma B + B^* J_2 B')\,\boldsymbol{z}\rangle \\ &= \omega_1 I_1 + \frac{\omega_1}{2}\langle \boldsymbol{z}, G\boldsymbol{z}\rangle . \end{aligned} \tag{1.5.8}$$

It remains to fix the matricial normal form G. If λ_0 and $1/\lambda_0$ are the two double eigenvalues of the monodromy matrix of $\mathcal{M}_0$ (as shown in figure 1.2), let ν be a real number such that $\lambda = e^{i2\pi\nu}$; then the matrix G can be written as,

$$G = \begin{pmatrix} 0 & 0 & 0 & -\nu \\ 0 & 0 & \nu & 0 \\ 0 & \nu & \epsilon/2\pi & 0 \\ -\nu & 0 & 0 & \epsilon/2\pi \end{pmatrix}, \tag{1.5.9}$$

(with $\epsilon = \pm 1$). Then

$$2\pi J_2 G = \begin{pmatrix} 0 & 2\pi\nu & \epsilon & 0 \\ -2\pi\nu & 0 & 0 & \epsilon \\ 0 & 0 & 0 & 2\pi\nu \\ 0 & 0 & -2\pi\nu & 0 \end{pmatrix}.$$

Remark 1.3. Actually, we take G such that the infinitesimal symplectic matrix $2\pi J_2 G$ is in normal form with respect conjugation by elements of $\mathrm{Sp}(2,\mathbb{R})$. For purely imaginary eigenvalues –see Van der Meer (1985), and also the reference of Burgoyne and Cushman (1974) quoted therein–, we get in $\mathrm{sp}(2,\mathbb{R})$ the two normal forms,

$$\begin{pmatrix} 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \\ -\lambda_1 & 0 & 0 & 0 \\ 0 & -\lambda_2 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & -\alpha & \epsilon & 0 \\ \alpha & 0 & 0 & \epsilon \\ 0 & 0 & 0 & -\alpha \\ 0 & 0 & \alpha & 0 \end{pmatrix} \tag{1.5.10}$$

(λ_1, λ_2, α may be positive or negative and $\epsilon = \pm 1$) correspond to the semisimple (diagonalizable) and non-semisimple (and also non-nilpotent) cases respectively. ♠

It is possible to establish the following.

Proposition 1.4. *With the choice of G above, the two hypothesis H1.5-1 and H1.5-2 are fulfilled.*

To prove this result we shall first introduce the next (structural) lemma –see also Bridges and Furter (1993)–.

Lemma 1.5. *Let $A \in Sp(2,\mathbb{R})$, with*

$$\mathrm{Spec}(A) = \left\{ \mathrm{e}^{\pm \mathrm{i}\theta} \right\},\ \theta \in (0,\pi) \quad \text{and} \quad \dim \mathrm{Ker}\left(A - \mathrm{e}^{\pm \mathrm{i}\theta} I_4\right) = 1.$$

Then, there exists $C \in Sp(2,\mathbb{R})$, such that,

$$C^{-1} A C = \left(\begin{array}{c|c} \mathcal{R}_\theta & \epsilon \mathcal{R}_\theta \\ \hline 0 & \mathcal{R}_\theta \end{array} \right), \tag{1.5.11}$$

with $\mathcal{R}_\theta = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix} \right)$, and, as before, $\epsilon = \pm 1$, which is an invariant of the collision.

Proof of lemma 1.5. We define $\lambda_\pm = \mathrm{e}^{\pm \mathrm{i}\theta}$; and let $\boldsymbol{z}$ and $\boldsymbol{w}$ be the geometric and generalized eigenvector of λ_+ respectively, so

$$A\boldsymbol{z} = \lambda_+ \boldsymbol{z}, \tag{1.5.12}$$

$$(A - \lambda_+ I_4)\, \boldsymbol{w} = \lambda_+ \boldsymbol{z}, \tag{1.5.13}$$

we note that the existence of such generalized eigenvector $\boldsymbol{w}$ follows immediately from the canonical normal form of matrix A which, from the hypothesis of the lemma, must be composed by the two Jordan blocks $\left(\begin{smallmatrix} \lambda & 0 \\ 1 & \lambda \end{smallmatrix} \right)$ and its complex conjugate, $\left(\begin{smallmatrix} \overline{\lambda} & 0 \\ 1 & \overline{\lambda} \end{smallmatrix} \right)$.

Let us now introduce the 2-form $\omega^2(\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{u}^* J_2 \boldsymbol{v}$ (i. e., the standard canonical two form in $\mathbb{C}^{2n}$, see example B.5 of appendix B). As A is a symplectic matrix, taking into account the definitions (1.5.12) and (1.5.13), we have,

$$\omega^2(\boldsymbol{z},\boldsymbol{w}) = \omega^2(A\boldsymbol{z}, A\boldsymbol{w}) = \omega^2(\lambda_+\boldsymbol{z}, \lambda_+\boldsymbol{z} + \lambda_+\boldsymbol{w}) = \lambda_+^2 \omega^2(\boldsymbol{z},\boldsymbol{\omega}),$$

so $(1-\lambda_+^2)\,\omega^2(\boldsymbol{z},\boldsymbol{w}) = 0$ and then, $\omega^2(\boldsymbol{z},\boldsymbol{w}) = 0$, since $|\lambda_\pm| = 1$ and $\lambda_+ \neq \pm 1$. In the same way,

$$\omega^2(\boldsymbol{z},\overline{\boldsymbol{w}}) = \omega^2(A\boldsymbol{z}, A\overline{\boldsymbol{w}}) = \omega^2(\lambda_+\boldsymbol{z}, \overline{\lambda}_+\overline{\boldsymbol{w}} + \overline{\lambda}_+\overline{\boldsymbol{z}}) = \omega^2(z,\overline{\boldsymbol{w}}) + \omega^2(\boldsymbol{z},\overline{\boldsymbol{z}}),$$

(here, the bar denotes complex conjugation), and it follows $\omega^2(\boldsymbol{z},\overline{\boldsymbol{z}}) = 0$. On the other hand

$$\begin{aligned}\omega^2(\boldsymbol{w},\overline{\boldsymbol{w}}) &= \omega^2(A\boldsymbol{w},A\overline{\boldsymbol{w}}) = \omega^2(\lambda_+\boldsymbol{z}+\lambda_+\boldsymbol{w},\overline{\lambda}_+\overline{\boldsymbol{z}}+\overline{\lambda}_+\overline{\boldsymbol{w}}) = \\ &= \omega^2(\boldsymbol{z},\overline{\boldsymbol{z}})+\omega^2(\boldsymbol{z},\overline{\boldsymbol{w}})+\omega^2(\boldsymbol{w},\overline{\boldsymbol{w}})+\omega^2(\boldsymbol{w},\overline{\boldsymbol{z}}).\end{aligned}$$

Then: $\omega^2(\boldsymbol{z},\overline{\boldsymbol{w}}) = -\omega^2(\boldsymbol{w},\overline{\boldsymbol{z}})$ but, due to the skew-symmetry of ω^2, this last product is equal to $\overline{\omega^2(\overline{\boldsymbol{z}},\boldsymbol{w})}$, and hence $\omega^2(\boldsymbol{z},\overline{\boldsymbol{w}}) = \overline{\omega^2(\boldsymbol{z},\overline{\boldsymbol{w}})}$, so $\omega^2(\boldsymbol{z},\overline{\boldsymbol{w}})$ is a real number. To summarize,

$$\omega^2(\boldsymbol{z},\boldsymbol{w}) = 0, \quad \omega^2(\boldsymbol{z},\overline{\boldsymbol{z}}) = 0 \quad \text{and} \quad \omega^2(\boldsymbol{z},\overline{\boldsymbol{w}}) \in \mathbb{R}. \tag{1.5.14}$$

Precisely, we shall define Δ as,

$$\Delta := \omega^2(\boldsymbol{z},\overline{\boldsymbol{w}}) \tag{1.5.15}$$

and the new scaled vectors, $\boldsymbol{z}', \boldsymbol{w}'$ as

$$\boldsymbol{z}' = \frac{\boldsymbol{z}}{\sqrt{|\Delta|}}, \quad \boldsymbol{w}' = \frac{\boldsymbol{w}}{\sqrt{|\Delta|}}, \tag{1.5.16}$$

as $\omega^2(\boldsymbol{z}',\overline{\boldsymbol{w}}') = \Delta/|\Delta| = \operatorname{sign}(\Delta)$, it is convenient to introduce the quantity,

$$\epsilon := \frac{\Delta}{|\Delta|}.$$

Then, $\epsilon = \pm 1$, and it is easy to realize that it does not depend on the chosen $\boldsymbol{z}$, so it is a sign characteristic of the collision. To continue, we substitute $\boldsymbol{w}'$ by $\boldsymbol{w}'' = \boldsymbol{w}' + \mathrm{i}\alpha\boldsymbol{z}'$, with $\alpha \in \mathbb{R}$ taken to make $\omega^2(\boldsymbol{w}'',\overline{\boldsymbol{w}}'') = 0$. As

$$\begin{aligned}\omega^2(\boldsymbol{w}'',\overline{\boldsymbol{w}}'') &= \omega^2(\boldsymbol{w}'+\mathrm{i}\alpha\boldsymbol{z}',\overline{\boldsymbol{w}}'-\mathrm{i}\alpha\overline{\boldsymbol{z}}') \\ &= \omega^2(\boldsymbol{w}',\overline{\boldsymbol{w}}') - \mathrm{i}\alpha\omega^2(\boldsymbol{w}',\overline{\boldsymbol{z}}') + \mathrm{i}\alpha\omega^2(\boldsymbol{z}',\overline{\boldsymbol{w}}') + \alpha^2\omega^2(\boldsymbol{z}',\overline{\boldsymbol{z}}') \\ &= \omega^2(\boldsymbol{w}',\overline{\boldsymbol{w}}') + 2\mathrm{i}\alpha\epsilon,\end{aligned}$$

we can take,

$$\alpha = -\frac{\omega^2(\boldsymbol{w}',\overline{\boldsymbol{w}}')}{2\mathrm{i}\epsilon}. \tag{1.5.17}$$

(Note that $\omega^2(\boldsymbol{w}',\overline{\boldsymbol{w}}') = -\overline{\omega^2(\boldsymbol{w}',\overline{\boldsymbol{w}}')}$, and thus $\omega^2(\boldsymbol{w}',\overline{\boldsymbol{w}}')$ is purely imaginary). At this point it is important to remark that neither the scaling (1.5.16), nor the choice of $\boldsymbol{w}''$ alter the two first properties of (1.5.14); thus, it is checked that

$$\omega^2(\boldsymbol{z}',\boldsymbol{w}'') = 0, \quad \omega^2(\boldsymbol{z}',\overline{\boldsymbol{z}}') = 0,$$

and moreover:

$$\omega^2(\boldsymbol{z}',\overline{\boldsymbol{w}}'') = \omega^2(\boldsymbol{z}',\overline{\boldsymbol{w}}'-\mathrm{i}\alpha\overline{\boldsymbol{z}}') = \omega^2(\boldsymbol{z}',\overline{\boldsymbol{w}}') - \mathrm{i}\alpha\omega^2(\boldsymbol{z}',\overline{\boldsymbol{z}}') = \epsilon.$$

We have,

$$\begin{aligned}A\boldsymbol{z}' &= \lambda_+\boldsymbol{z}', \\ A(\epsilon\boldsymbol{w}'') &= \epsilon A(\boldsymbol{w}'+\mathrm{i}\alpha\boldsymbol{z}') \\ &= \epsilon\lambda_+\boldsymbol{z}' + \epsilon\lambda_+\boldsymbol{w}' + \mathrm{i}\alpha\epsilon\lambda_+\boldsymbol{z}' \\ &= (\epsilon\lambda_+)\boldsymbol{z}' + \lambda_+\epsilon(\boldsymbol{w}'+\mathrm{i}\alpha\boldsymbol{z}') \\ &= (\epsilon\lambda_+)\boldsymbol{z}' + \lambda_+\epsilon\boldsymbol{w}'', \\ A(\epsilon\overline{\boldsymbol{w}}'') &= \epsilon\overline{\lambda}_+\overline{\boldsymbol{w}}'' - \epsilon\overline{\lambda}_+(-\overline{\boldsymbol{z}}'), \\ A(-\overline{\boldsymbol{z}}') &= \overline{\lambda}_+(-\overline{\boldsymbol{z}}').\end{aligned}$$

So, in the (complex) basis $\mathfrak{B} = \left\{\frac{1}{\sqrt{2}}\boldsymbol{z}', \frac{\epsilon}{\sqrt{2}}\boldsymbol{w}'', \frac{\epsilon}{\sqrt{2}}\overline{\boldsymbol{w}}'', \frac{-1}{\sqrt{2}}\overline{\boldsymbol{z}}'\right\}$, the matrix A takes the form

$$\widehat{A} = \left(\begin{array}{cc|cc} \lambda_+ & \epsilon\lambda_+ & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ 0 & \lambda_+ & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{0}} & \lambda_- & 0 \\ & & -\epsilon\lambda_- & \lambda_- \end{array}\right),$$

($\overline{\lambda}_+ = \lambda_-$, by definition). $\mathfrak{B}$ is a symplectic basis, for

$$\begin{aligned} \omega^2(\boldsymbol{z}', \epsilon\, \boldsymbol{w}'') &= \epsilon\, \omega^2(\boldsymbol{z}', \boldsymbol{w}') = 0, \\ \omega^2(\boldsymbol{z}', \epsilon\, \overline{\boldsymbol{w}}'') &= \epsilon\, \omega^2(\boldsymbol{z}', \overline{\boldsymbol{w}}'') = \epsilon^2 = 1, \\ \omega^2(\boldsymbol{z}', \overline{\boldsymbol{z}}') &= 0, \\ \omega^2(\epsilon\, \boldsymbol{w}'', \epsilon\, \overline{\boldsymbol{w}}'') &= \epsilon^2\omega^2(\boldsymbol{w}'', \overline{\boldsymbol{w}}'') = 0, \\ \omega^2(\epsilon\boldsymbol{w}'', -\overline{\boldsymbol{z}}') &= -\epsilon\, \omega^2(\boldsymbol{w}'', \overline{\boldsymbol{z}}') = \epsilon\, \overline{\omega^2(\boldsymbol{z}', \overline{\boldsymbol{w}}'')} = \epsilon^2 = 1, \\ \omega^2(\epsilon\overline{\boldsymbol{w}}'', \overline{\boldsymbol{z}}') &= \overline{\omega^2(\epsilon\boldsymbol{w}'', \boldsymbol{z}')} = 0. \end{aligned}$$

Now, if we consider the (real) basis $\mathfrak{B}' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$, with $\boldsymbol{u}_i$, $i = 1, 2, 3, 4$ given by,

$$\boldsymbol{u}_1 = \frac{\boldsymbol{z}' + \overline{\boldsymbol{z}}'}{2}, \quad \boldsymbol{u}_2 = \frac{\boldsymbol{z}' - \overline{\boldsymbol{z}}'}{2\mathrm{i}}, \quad \boldsymbol{u}_3 = \epsilon\frac{\boldsymbol{w}'' + \overline{\boldsymbol{w}}''}{2}, \quad \boldsymbol{u}_4 = \epsilon\frac{\boldsymbol{w}'' - \overline{\boldsymbol{w}}''}{2\mathrm{i}}, \tag{1.5.18}$$

so the matrix of the change from $\mathfrak{B}$ to $\mathfrak{B}'$ is

$$\widehat{C} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\mathrm{i}\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\mathrm{i}\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\mathrm{i}\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\mathrm{i}\sqrt{2}} & 0 & 0 \end{pmatrix}.$$

It can be checked out that $\widehat{C}^* J_2 \widehat{C} = J_2$ (i. e., $\widehat{C}$ is a canonical matrix) and also,

$$A' = \widehat{C}^{-1}\widehat{A}\widehat{C} = \left(\begin{array}{cc|cc} \cos\theta & \sin\theta & \epsilon\cos\theta & \epsilon\sin\theta \\ -\sin\theta & \cos\theta & -\epsilon\sin\theta & \epsilon\cos\theta \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{0}} & \cos\theta & \sin\theta \\ & & -\sin\theta & \cos\theta \end{array}\right),$$

but this is just the matrix (1.5.11) of the statement of the lemma. This ends the proof. $\square$

The basis, $\mathfrak{B}'$ can be written as, $\mathfrak{B}' = \{\mathrm{Re}\,\boldsymbol{z}', \mathrm{Im}\,\boldsymbol{z}', \epsilon\,\mathrm{Re}\,\boldsymbol{w}'', \epsilon\,\mathrm{Im}\,\boldsymbol{w}''\}$. So, the columns of the matrix C of the global change are formed by the vectors of this basis. We shall write,

$$C = \left(\mathrm{Re}\,\boldsymbol{z}' | \mathrm{Im}\,\boldsymbol{z}' | \epsilon \mathrm{Re}\,\boldsymbol{w}'' | \epsilon \mathrm{Im}\,\boldsymbol{w}''\right), \tag{1.5.19}$$

In the next example, we construct the basis $\mathfrak{B}'$ and the matrix C for a given specific matrix.

Example 1.6. For the matrix (1.3.4) of section 1.3.2, the critical values of the parameter, L, were $L = L_{crit} = -K/4$. This produces a degenerate (or critical) matrix A_{crit}, of the form

$$A_{crit} = \begin{pmatrix} 1+\dfrac{1}{4}K & \dfrac{1}{4}K & \dfrac{1}{4}K & \dfrac{1}{4}K \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ -\dfrac{1}{4}K & \dfrac{3}{4}K & -\dfrac{1}{4}K & 1+\dfrac{3}{4}K \end{pmatrix},$$

with $-8 < K < 0$. It turns out that $\mathrm{Spec}(A_{crit}) = \{\lambda_\pm\}$, being

$$\lambda_\pm = 1 + \frac{1}{4}K \pm \mathrm{i}\frac{1}{4}\sqrt{8|K| - K^2}.$$

If we put $\lambda_+ = \cos\theta + \mathrm{i}\sin\theta$, we can identify

$$\cos\theta = 1 + \tfrac{1}{4}K \quad \text{and} \quad \sin\theta = \tfrac{1}{4}\sqrt{8|K| - K^2}, \tag{1.5.20}$$

and compute the geometric and generalized eigenvectors:

$$\begin{aligned} z^* &= \left(\tfrac{K+\mathrm{i}\sqrt{8|K|-K^2}}{4}, 1, 1, \tfrac{K+\mathrm{i}\sqrt{8|K|-K^2}}{4} \right) \\ w^* &= \left(\tfrac{4+K+\mathrm{i}\sqrt{8|K|-K^2}}{4}, \tfrac{2\mathrm{i}\sqrt{8|K|-K^2}}{K}, 0, \tfrac{20+3K+3\mathrm{i}\sqrt{8|K|-K^2}}{4} \right), \end{aligned}$$

so, in this case, $\Delta = \omega^2(\boldsymbol{z}, \overline{\boldsymbol{w}}) = \boldsymbol{z}^* J_2 \overline{\boldsymbol{w}} = K + 8 > 0$; and thus, $\epsilon = \mathrm{sign}(\Delta) = 1$. From here, we can normalize to compute $\boldsymbol{z}'$ and $\boldsymbol{w}'$. Then, the corresponding α, as defined by (1.5.17) is,

$$\alpha = -\frac{20 + 3K}{2\sqrt{K+8}}|K|^{1/2}$$

and the matrix C of the change (see (1.5.19) and consider $\epsilon = 1$), can be taken as:

$$C = \begin{pmatrix} \frac{K}{4\sqrt{K/2+4}} & \frac{1}{4}\sqrt{\frac{8|K|-K^2}{K/2+4}} & -\frac{\sqrt{2}(K+12)}{8\sqrt{K+8}} & -\frac{\sqrt{2|K|}(K+4)}{8(K+8)} \\ \frac{1}{\sqrt{K/2+4}} & 0 & 0 & \frac{\sqrt{2|K|}(K+12)}{2K(K+8)} \\ \frac{1}{\sqrt{K/2+4}} & 0 & 0 & -\frac{\sqrt{2|K|}(3K+20)}{2K(K+8)} \\ \frac{K}{4\sqrt{K/2+4}} & \frac{1}{4}\sqrt{\frac{8|K|-K^2}{K/2+4}} & \frac{\sqrt{2}(3K+20)}{8\sqrt{K+8}} & \frac{\sqrt{2|K|}(3K+28)}{8(K+8)} \end{pmatrix}.$$

So, the conjugate matrix is a block matrix, of the expected type, i. e.: $C^{-1}A_{crit}C = \left(\begin{smallmatrix} c_K & c_K \\ 0 & c_K \end{smallmatrix}\right)$, with the block c_K defined by the 2×2 matrix,

$$c_K = \begin{pmatrix} 1 + \frac{1}{4}K & \frac{1}{4}\sqrt{8|K| - K^2} \\ -\frac{1}{4}\sqrt{8|K| - K^2} & 1 + \frac{1}{4}K \end{pmatrix}.$$

Thus, taking into account the former identification (1.5.20), it is clear that this is a rotation matrix (i. e., $c_K = \mathcal{R}_\theta$), then it has been checked that the matrix C is the transformation matrix we were looking for. $\diamond$

Applying now the lemma 1.5, it is straightforward to prove the proposition 1.4.

Proof of proposition 1.4. If we put $S_0 = \exp(2\pi J_2 G)$, it can be seen that:

$$S_0 = \left(\begin{array}{c|c} \mathcal{R}_{2\pi\nu} & \epsilon\mathcal{R}_{2\pi\nu} \\ \hline 0 & \mathcal{R}_{2\pi\nu} \end{array}\right).$$

Applying lemma 1.5, a (real) canonical matrix, C, exists such that $C^{-1}M_0C = S_0$, so

$$M_0 = D^{-1}S_0D = D^{-1}\exp(2\pi J_2 G)D = \exp(2\pi J_2 D^* G D),$$

with $D = C^{-1}$ (and therefore a canonical matrix). Then, $N_0 = D^*GD$, and the proposition is proved. □

Hence, the quadratic part of the Hamiltonian, as expressed in (1.5.8), and with the normalized matrix G of (1.5.9) is,

$$\widetilde{\mathcal{H}}_2(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}) = \omega_1 I_1 + \omega_1\nu(y_1x_2 - y_2x_1) + \epsilon\frac{\omega_1}{4\pi}(y_1^2 + y_2^2).$$

Next, we make the symplectic change,

$$\begin{aligned} x_1 &= \epsilon\sqrt{\frac{\omega_1}{2\pi}}x_1', \qquad & y_1 &= \epsilon\sqrt{\frac{2\pi}{\omega_1}}y_1', \\ x_2 &= \sqrt{\frac{\omega_1}{2\pi}}x_2', \qquad & y_2 &= \sqrt{\frac{2\pi}{\omega_1}}y_2', \end{aligned}$$

and

$$I_1 = \epsilon I_1', \qquad \theta_1 = \epsilon\theta_1'.$$

With this last change, the quadratic part of the Hamiltonian takes the form (tildes and primes have been dropped),

$$\mathcal{H}_2(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}) = \epsilon\omega_1 I_1 + \epsilon\nu\omega_1(y_1x_2 - y_2x_1) + \frac{\epsilon}{2}(y_1^2 + y_2^2).$$

Now, we define $\omega_2 = \nu\omega_1$; and changing the sign of the time, if necessary, $t \mapsto \epsilon t$, it is obtained finally a new linearly reduced Hamiltonian, whose quadratic part is:

$$\mathcal{H}_2(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}) = \omega_1 I_1 + \omega_2(y_1x_2 - y_2x_1) + \frac{1}{2}(y_1^2 + y_2^2). \tag{1.5.21}$$

This will be the quadratic part of the Hamiltonian we shall consider henceforth, and whose complete expansion can now be written as,

$$\mathcal{H}(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}) = \mathcal{H}_2 + \sum_{2l+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1\geq 3} h_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1)I_1^l\boldsymbol{x}^{\boldsymbol{m}}\boldsymbol{y}^{\boldsymbol{n}}. \tag{1.5.22}$$

1.6 Complexification of the Hamiltonian

Before continuing with the nonlinear normalization, and in order to get the homological equations in a simpler form, it is convenient to introduce the following (complex) coordinates,

$$x_1 = \frac{q_1 - p_2}{\sqrt{2}} \quad x_2 = -\frac{q_1 + p_2}{\mathrm{i}\sqrt{2}}, \quad y_1 = \frac{q_2 + p_1}{\sqrt{2}}, \quad y_2 = -\frac{q_2 - p_1}{\mathrm{i}\sqrt{2}}. \tag{1.6.1}$$

These last relations define a canonical change, $(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{S}(\boldsymbol{q}, \boldsymbol{p})$ which transforms the Hamiltonian (1.5.22) into

$$H(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) = H_2 + \sum_{2l+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1 \geq 3} h_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}}, \tag{1.6.2}$$

where H_2 is the quadratic part (see (1.6.4)). As usual, we have put $\boldsymbol{q}^* = (q_1, q_2)$, $\boldsymbol{p}^* = (p_1, p_2)$ and $h_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1)$ are analytic 2π-periodic functions. For $\rho > 0$, $R > 0$, we define the set

$$\mathcal{D}(\rho, R) = \{(\theta_1, I_1, \boldsymbol{q}, \boldsymbol{p}) \in \mathbb{C}^6 : |\mathrm{Im}\,\theta_1| \leq \rho, \ |I_1| \leq R^2, \ |\boldsymbol{z}| \leq R\} \tag{1.6.3}$$

with $\boldsymbol{z}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*) \in \mathbb{C}^4$ and $|\cdot|$ denotes the supremum norm of a complex vector. We shall consider the Hamiltonian H defined in $\mathcal{D}(\rho^*, R^*)$. Therefore, $\mathcal{S}(\mathcal{D}(\rho^*, R^*)) \subseteq \mathcal{D}(\rho^*, \sqrt{2}R^*)$, and we shall ask ρ^*, R^* small enough such that this last domain is included in the complexification of the transformed –through all the changes described in sections 1.4, 1.5– neighborhood, $\mathfrak{U}$, of the initial periodic orbit $\mathcal{M}_0$.

Also, by direct substitution of (1.6.1) in the Hamiltonian (1.5.22), it can be seen that the quadratic part in (1.6.2) is,

$$H_2 = \omega_1 I_1 + \mathrm{i}\omega_2(q_1 p_1 + q_2 p_2) + q_2 p_1. \tag{1.6.4}$$

This will be the lowest-order term in our normal form. Note that, in the change (1.6.1) $\boldsymbol{x}$ and $\boldsymbol{y}$ will be real, provided

$$\overline{q}_1 = -p_2, \quad \overline{q}_2 = p_1, \quad \overline{p}_1 = q_2, \quad \overline{p}_2 = -q_1, \tag{1.6.5}$$

If the above relations are assumed to hold and, as the complex Hamiltonian H is the transformed of a real Hamiltonian $\mathcal{H}$, it must be $\overline{H} = H$. So, if we consider the expansion of H in Poisson (Taylor-Fourier) series,

$$H(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) = \sum_{k \in \mathbb{Z}} \sum_{l,\boldsymbol{m},\boldsymbol{n}} h_{k,l,\boldsymbol{m},\boldsymbol{n}} I_1^l \, \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \exp(\mathrm{i}k\theta_1), \tag{1.6.6}$$

it is readily checked that the coefficients h_{k,l,m_1,m_2,n_1,n_2} must satisfy the following symmetries,

$$\overline{h}_{k,l,m_1,m_2,n_1,n_2} = (-1)^{m_1+n_2} h_{-k,l,n_2,n_1,m_2,m_1}. \tag{1.6.7}$$

Reciprocally, if the complex coefficients of an expansion of type (1.6.6) satisfy the symmetries (1.6.7) above, then composition with the inverse change, $\mathcal{S}^{-1}$, transforms it back to a real Poisson series.

1.7 Nonlinear normalization

Here, we shall apply a normal form process to reduce the higher (greater than two) degree terms of the Hamiltonian (1.6.2). As we are interested in bounding the remainder, it is advisable to use some "closed" transformation algorithm in the following sense: instead of applying to the function $f = \sum_{l\geq 1} f_l$ a sequence of transformations $\{\phi_1^{G_s}\}_{s\geq 3}$ obtained as time unit flow of the Hamiltonians G_s for $s = 3, 4, \dots$ (and hence the transformed function is given by $f \circ \phi_1^{G_3} \circ \cdots \circ \phi_1^{G_s} \circ \cdots$), one looks for a generating function $G = \sum_{s\geq 3} G_s$ of a *single* canonical change: ϕ^G such that $T_G f = f \circ \phi^G = \sum_{s\geq 1} F_s$. This kind of algorithms provide recursive formulas to compute the terms F_s of degree $s = 1, 2, \dots$. In particular, throughout this work, we shall use the Giorgilli-Galgani algorithm (see Giorgilli and Galgani, 1978, 1985; Giorgilli et al., 1989) applied to formal series of type,

$$f = \sum_{(l,\boldsymbol{m},\boldsymbol{n})\in\mathbb{Z}_+\times\mathbb{Z}_+^2\times\mathbb{Z}_+^2} f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \tag{1.7.1}$$

with the multi-index notation: $\boldsymbol{q}^{\boldsymbol{m}} = q_1^{m_1} q_2^{m_2}$, $\boldsymbol{p}^{\boldsymbol{n}} = p_1^{n_1} p_2^{n_2}$ and with $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. On the other hand, the coefficients $f_{l,\boldsymbol{m},\boldsymbol{n}}$ are 2π-periodic functions of its (complex) argument θ_1 and we suppose that they can be expanded in Fourier series

$$f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) = \sum_{k\in Z} f_{k,l,\boldsymbol{m},\boldsymbol{n}} \exp(\mathrm{i}k\theta_1). \tag{1.7.2}$$

Now, given a monomial $h_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}}$ in (1.7.1), we define its adapted degree (so its degree from now on) as

$$\deg\left(h_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}}\right) = 2l + |\boldsymbol{m}|_1 + |\boldsymbol{n}|_1, \tag{1.7.3}$$

where $|\cdot|_1$ denotes the moduli-summation norm, i. e., if $\boldsymbol{u} \in \mathbb{R}^n$ (or $\mathbb{C}^n$) it is $|\boldsymbol{u}|_1 = \sum_{i=1}^n |u_i|$. Thus, the degree of the action variable is counted twice with respect to the degree of coordinates and momenta. Furthermore, throughout this section, $\mathfrak{E}$ will denote the space of the formal Taylor-Fourier series –in $\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}$– of type (1.7.1) and $\mathfrak{E}_s$ the subspace of those ones with (adapted) degree s; so given $f \in \mathfrak{E}$, it can be expanded degreewise as $f = \sum_{s\geq 1} f_s$, with $f_s \in \mathfrak{E}_s$ for all $s = 1, 2, \dots$ With this notation we can define now (see Giorgilli et al., 1989).

Definition 1.7 (The Giorgilli-Galgani algorithm). *Let $G \in \mathfrak{E}$ be given by the sum $G = \sum_{s\geq 3} G_s$, we define the map $T_G : \mathfrak{E} \to \mathfrak{E}$ in the following way: if $f = \sum_{l\geq 1} f_l \in \mathfrak{E}$, then*

$$T_G f = \sum_{s\geq 1} F_s, \tag{1.7.4}$$

where

$$F_s = \sum_{l=1}^{s} f_{l,s-l}, \tag{1.7.5}$$

and the terms $f_{l,s}$ can be computed recursively by the formulas,

$$f_{l,0} = f_l, \quad f_{l,s} = \sum_{j=1}^{s} \frac{j}{s} L_{G_{2+j}} f_{l,s-j}. \tag{1.7.6}$$

Here $L_u = \{\cdot, u\}$ (see section B.3 of appendix B). In our case, when $f, g \in \mathfrak{E}$, their Poisson bracket can be written as,

$$\{f,g\} = \frac{\partial f}{\partial \theta_1}\left(\frac{\partial g}{\partial I_1}\right)^* - \frac{\partial f}{\partial I_1}\left(\frac{\partial g}{\partial \theta_1}\right)^* + \frac{\partial f}{\partial \boldsymbol{z}} J_2 \left(\frac{\partial g}{\partial \boldsymbol{z}}\right)^*,$$

with this definition, $\deg(\{f,g\}) = \deg(f) + \deg(g) - 2$, and by induction, it is easily seen that $\deg(f_{l,s}) = l + s$. The adapted degree is here the key to extend the Giorgilli-Galgani formulas to functions in the space $\mathfrak{E}$.

In the first and the third of the references quoted at the beginning of the section, the algorithm was applied to polynomial series for computing normal forms around equilibrium points, while in the second –i. e., in Giorgilli and Galgani (1985)–, the method is used to transform a Hamiltonian of type $H(\boldsymbol{q}, \boldsymbol{p}, \varepsilon) = h(\boldsymbol{p}) + \varepsilon f(\boldsymbol{q}, \boldsymbol{p})$, with $\boldsymbol{p}^* = (p_1, \ldots, p_n) \in \mathcal{G} \subset \mathbb{R}^n$ and $\boldsymbol{q}^* = (q_1, \ldots, q_n) \in \mathbb{T}^n$, the n-dimensional torus. Nevertheless, therein, only a finite number of harmonics in the expansion of the perturbative term, $\varepsilon f(\boldsymbol{q}, \boldsymbol{p})$, are considered.

In addition, it can be shown that the coordinate transformation given by $\theta_1 = T_G \theta_1'$, $I_1 = T_G I_1'$, $q_i = T_G q_i'$ and $p_i = T_G p_i'$ with $(i = 1, 2)$ is canonical. Furthermore, if $f \in \mathfrak{E}$, and the function F is constructed as described in 1.7, then $F(\theta_1', I_1', \boldsymbol{q}', \boldsymbol{p}') = T_G f(\theta_1', I_1', \boldsymbol{q}', \boldsymbol{p}') = f(T_G \theta_1', T_G I_1', T_G \boldsymbol{q}', T_G \boldsymbol{p}')$. For an account of these properties, together with their corresponding proofs, see Giorgilli and Galgani (1978).

Remark 1.8. We will not use new names neither for the transformed functions, nor for the new coordinates so, to simplify the notation, the primes will be omitted. ♠

The idea is thus to take $f = H$, the Hamiltonian function, and employ the algorithm 1.7 to cast it into a simpler form (its "Normal form"), removing all the *nonresonant* terms. With this purpose, we construct an *ad hoc* generating function of the form $G = \sum_{s \geq 3} G_s$ and both, G_s and the resonant terms, Z_s, can be determined recursively. The first two steps of the reduction process are formally indicated below[3].

So we want the transformed Hamiltonian, $T_G H$, to have the form $T_G H = Z_1 + Z_2 + \cdots + Z_s + \ldots$, with $Z_1 = 0$ and $Z_2 = H_2$. Accordingly with the notation of definition 1.7, we shall put $H_{l,0} = H_l$. Therefore, by the formula (1.7.6): $H_{1,1} = L_{G_3} H_{1,0} = 0$. If we suppose that this is true also for $H_{1,i}$ with $i = 1, \ldots, s-1$, by (1.7.6), $H_{1,s} = \sum_{j=1}^{s} \frac{j}{s} L_{G_{2+j}} H_{1,s-j} = 0$. Thus, by induction $H_{1,s} = 0$, for all $s \in \mathbb{Z}_+$ and hence, joining (1.7.4) and (1.7.5),

$$T_G H = \sum_{s \geq 2} \sum_{l=0}^{s-2} H_{l+2, s-l-2}$$

with $H_{l,s}$ given by (1.7.6). From this last expression, we can identify

$$Z_s = \sum_{l=0}^{s-2} H_{l+2, s-l-2}, \qquad (1.7.7)$$

for $s \geq 3$ and, as it has been already prompted, $Z_2 = H_2$. So, for example,

$$Z_3 = H_{2,1} + H_{3,0}.$$

[3] This is a formal construction, if one wants convergence, we have to truncate G at a finite order.

Here, using that $H_{2,1} = L_{G_3}H_2$ and $H_{3,0} = H_3$, the expression above can be arranged as,

$$L_{H_2}G_3 + Z_3 = H_3. \tag{1.7.8}$$

As H_3 is known from the development of the initial Hamiltonian; and, we accept (see section 1.7.1) that this *homological* equation can be solved in G_3 and Z_3. In the same way,

$$Z_4 = H_{2,2} + H_{3,1} + H_{4,0} = \frac{1}{2}L_{G_3}H_{2,1} + L_{G_4}H_{2,0} + L_{G_3}H_{3,0} + H_{4,0},$$

and, since $H_{2,1} = Z_3 - H_3$, Z_4 can be expressed as

$$Z_4 = \frac{1}{2}L_{G_3}Z_3 + \frac{1}{2}L_{G_3}H_3 + H_4 + L_{G_4}H_2.$$

Now, if we define

$$F_4 = \frac{1}{2}L_{G_3}Z_3 + \tfrac{1}{2}L_{G_3}H_3 + H_4, \tag{1.7.9}$$

the fourth-degree homological equations result

$$L_{H_2}G_4 + Z_4 = F_4, \tag{1.7.10}$$

from which, as before, we suppose that G_4 and Z_4 can be determined.

This process can be carried out recursively, provided we know an explicit expression for the right hand side term, F_s, as a function of the previous computed terms $Z_3, \ldots, Z_{s-1}; G_3, \ldots, G_{s-1}$. plus the components $H_3, \ldots, H_s$ of the initial Hamiltonian –as in (1.7.10) for $s = 4$–. But, indeed, it first requires (for $i = 3, \ldots, s$), the existence of G_i and Z_i as solutions of the corresponding homological equations (above, it was only *assumed* for $i = 3, 4$). So, before describing the "reduction algorithm" (see section 1.7.3) for computing the normal form; we need to investigate the solvavility of

$$L_{H_2}G_s + Z_s = F_s, \tag{1.7.11}$$

for a given function F_s. This constitutes the subject of the next section.

1.7.1 Algebraic properties of the homological equations

Here, we shall consider (1.7.11) with $F_s \in \mathfrak{E}_s$, being $\mathfrak{E}_s \subset \mathfrak{E}$ the subspace of the formal Poisson series of type (1.6.6) having all their monomials degree s. More precisely: $F \in \mathfrak{E}_s$, if and only if,

$$F = \sum_{\substack{2l+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1=s \\ k\in\mathbb{Z}}} f_{k,l,\boldsymbol{m},\boldsymbol{n}} I_1^l \, \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \exp(\mathrm{i}k\theta_1). \tag{1.7.12}$$

Furthermore, it is assumed that the coefficients of F_s satisfy the symmetry relations of (1.6.7). To account briefly for this fact, we introduce the following definition.

Definition 1.9. *We say that $h \in \mathfrak{E}$ satisfies the $\mathcal{S}$-symmetries, or that h is $\mathcal{S}$-symmetric, if and only if the relations*

$$\overline{h}_{k,l,m_1,m_2,n_1,n_2} = (-1)^{m_1+n_2} h_{-k,l,n_2,n_1,m_2,m_1}, \tag{1.7.13}$$

are fulfilled by all the coefficients of the monomials in h. The sets $\mathfrak{E}^{\mathcal{S}}$ ($\mathfrak{E}_s^{\mathcal{S}}$) are those subsets of $\mathfrak{E}$ ($\mathfrak{E}_s$) with all their elements satisfying the $\mathcal{S}$-symmetries.

Our target is to show that if $F_s \in \mathfrak{E}_s^{\mathcal{S}}$, there exist $G_s, Z_s \in \mathfrak{E}_s^{\mathcal{S}}$ verifying identically equation (1.7.11). What we shall do is to set $Z_s = 0$ and to investigate the solvability of the equations

$$L_{H_2} G_s = F_s. \tag{1.7.14}$$

In particular, we shall find out the possible resonant monomials in F_s, which will determine the form of Z_s in (1.7.11).

The keystone for the proof is the way how operator L_{H_2} acts on a monomial $g = I_1^l\, q_1^{m_1} q_2^{m_2} p_1^{n_1} p_2^{n_2} \exp(ik\theta_1) \in \mathfrak{E}_s$. It can be seen, by direct computation from the definition of L_{H_2}, that

$$L_{H_2} g = \{g, H_2\} = \left(\Omega + m_1 \frac{q_2}{q_1} - n_2 \frac{p_1}{p_2}\right) g, \tag{1.7.15}$$

where Ω is introduced as,

$$\Omega = \Omega_{k,|\boldsymbol{m}|_1,|\boldsymbol{n}|_1} = \mathrm{i}\omega_1 k + \mathrm{i}\omega_2(|\boldsymbol{m}|_1 - |\boldsymbol{n}|_1). \tag{1.7.16}$$

On the other hand, the quotient q_2/q_1 does not appear if the monomial g has $m_1 = 0$. Similarly for the quotient p_1/p_2. With this remark, the expression (1.7.15) is fully justified. Immediately from there, it follows that L_{H_2} preserves:

(i) The value of $k \in \mathbb{Z}$ of the monomial g.

(ii) The degree of the action, l.

(iii) The sums of the degrees $|\boldsymbol{m}|_1$ of $\boldsymbol{q}$, and $|\boldsymbol{n}|_1$ of $\boldsymbol{p}$, respectively.

(iv) As a consequence of the last two points, the global degree s is also preserved. This is consistent with our definition of the adapted degree.

In a natural way, we can consider the subspaces $\mathfrak{E}_{k,l,M,N}$, given by the Poisson series $F \in \mathfrak{E}_s$ of the form

$$F = \sum_{\substack{0\le m\le M\\ 0\le n\le N}} f_{k,l,m,M-m,N-n,n} I_1^l\, q_1^m q_2^{M-m} p_1^{N-n} p_2^n \exp(\mathrm{i}k\theta_1), \tag{1.7.17}$$

with $2l + M + N = s$. Therefore,

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

Hence, since every one of the subspaces $\mathfrak{E}_{k,l,M,N}$ is invariant under L_{H_2}, we can split the homological equations for the degree s into "boxes" with fixed k and fixed l, M, N verifying $2l + M + N = s$.

Another point to note is that the coefficients of a $F \in \mathfrak{E}_{k,l,M,N}$ are readily determined by just a pair of indices because, in view of (1.7.17), it is advisable to denote: $f_{m,n} = f_{k,l,m,M-m,N-n,n}$ (we skip both, k and l since they are held fixed).

With this notation, the homological equations (1.7.14) restricted to the different finite-dimensional subspaces, $\mathfrak{E}_{k,l,M,N}$, can be translated into an algebraic system of linear equations for the (complex) coefficients. Explicitly,

$$\Omega_{k,M,N}\, g_{m,n} + (m+1)\, \widetilde{\delta}_{m,M}\, g_{m+1,n} - (n+1)\, \widetilde{\delta}_{n,N}\, g_{m,n+1} = f_{m,n}, \tag{1.7.18}$$

where $\widetilde{\delta}_{i,j} = 1 - \delta_{i,j}$, being $\delta_{i,j}$ the Kronecker's delta, and $g_{m,n}, f_{m,n}$ (for $0 \le m \le M$, $0 \le n \le N$) are the coefficients of the projections (on $\mathfrak{E}_{k,l,M,N}$) of the generating function G_s and the right hand side term F_s of (1.7.14), respectively.

To investigate the existence of solutions of (1.7.18), we first write them in a convenient matrix form of type,

$$\Lambda \boldsymbol{g} = \boldsymbol{f}. \tag{1.7.19}$$

Here, $\boldsymbol{g} = \boldsymbol{g}_{l,M,N,k}$ and $\boldsymbol{f} = \boldsymbol{f}_{l,M,N,k}$ are the arrays holding $g_{m,n}$ and $f_{m,n}$ respectively, for $0 \le m \le M$ and $0 \le n \le N$ ($M + N = s - 2l$, fixed). Suppose that these coefficients have been ordered 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, $\boldsymbol{g}$ and $\boldsymbol{f}$ will take the form,

$$\boldsymbol{g}^* = (\boldsymbol{g}^*_M, \boldsymbol{g}^*_{M-1}, \ldots, \boldsymbol{g}^*_0), \quad \boldsymbol{f}^* = (\boldsymbol{f}^*_M, \boldsymbol{f}^*_{M-1}, \ldots, \boldsymbol{f}^*_0), \tag{1.7.20}$$

still with $\boldsymbol{g}^*_J = (g_{J,N}, g_{J,N-1}, g_{J,N-2}, \ldots, g_{J,1}, g_{J,0})$, for $J = 0, \ldots, M$ (and identically for $\boldsymbol{f}$). It is then straightforward to check that the system (1.7.19) is written by blocks as,

$$\begin{pmatrix} D_N & & & & & \\ E_M & D_N & & & & \\ & E_{M-1} & D_N & & & \\ & & \ddots & \ddots & & \\ & & & E_2 & D_N & \\ & & & & E_1 & D_N \end{pmatrix} \begin{pmatrix} \boldsymbol{g}_M \\ \boldsymbol{g}_{M-1} \\ \boldsymbol{g}_{M-2} \\ \vdots \\ \boldsymbol{g}_1 \\ \boldsymbol{g}_0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_M \\ \boldsymbol{f}_{M-1} \\ \boldsymbol{f}_{M-2} \\ \vdots \\ \boldsymbol{f}_1 \\ \boldsymbol{f}_0 \end{pmatrix}. \tag{1.7.21}$$

The different blocks stand for $E_j = j \cdot I_{N+1}$ (i. e., the product of the integer j with the identity $(N+1)\times(N+1)$ matrix, for $j = 1, 2, \ldots, M$), whereas $D_N = \Omega \cdot I_{N+1} - P_N$, with $\Omega = \Omega_{k,M,N}$ and being P_N the $(N+1)\times(N+1)$ nilpotent matrix,

$$P_N = \begin{pmatrix} 0 & & & & \\ N & 0 & & & \\ & N-1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}. \tag{1.7.22}$$

It follows from this description, that the matrix Λ in (1.7.19) –now identified with the one of the linear system (1.7.21)–, is a band 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 them equal to Ω. Hence, if $\Omega \ne 0$, this specific part of the homological equations (we mean their components in the subspace $\mathfrak{E}_{k,l,M,N}$), has an unique solution. Also, the next lemma is readily deduced.

Lemma 1.10. *A necessary condition for $f_{k,l,\boldsymbol{m},\boldsymbol{n}} I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \exp(\mathrm{i}k\theta_1)$, a monomial of F_s (see equation (1.7.11)) to be resonant (i. e., non removable by the choice of an appropriate generating function, G_s) is that $k = 0$ and $|\boldsymbol{m}|_1 = |\boldsymbol{n}|_1$.*

This results from: (i) the definition, at the end of section 1.5 of $\omega_2 = \nu\omega_1$, where $2\pi\nu$ was the characteristic exponent of the critical periodic orbit, (ii) our assumption $\nu \notin \mathbb{Q}$, corresponding to the hypothesis of *irrational* collision (see page 3), and (iii) the definition in (1.7.16) of Ω, as a linear combination of the two frequencies ω_1 and ω_2 with integer coefficients. Clearly, the result works for all $s = 3, 4, \ldots$

Definition 1.11. *A monomial* $f = f_{k,l,\boldsymbol{m},\boldsymbol{n}} I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \exp(\mathrm{i}k\theta_1) \in \mathfrak{E}_s$ *is said to be of* $\mathcal{M}$*-type if* $k = 0$ *and* $|\boldsymbol{m}|_1 = |\boldsymbol{n}|_1$.

Remark 1.12. Before continuing with the study of the (possible) resonant terms –those of $\mathcal{M}$-type just defined–, it is important to stress that if, as assumed, the right hand side F_s in (1.7.11), satisfies the $\mathcal{S}$-symmetries of definition 1.9 and does not contain $\mathcal{M}$-type monomials, then the solution of the system –the generating function G_s–, verifies also the S-symmetries. This is a consequence of their preservation under the Poisson bracket, and can be checked directly from the solutions of the linear equations (1.7.21):

$$g_{k,M,N} = \tfrac{1}{\Omega_{k,M,N}} f_{k,M,N}, \tag{1.7.23a}$$

$$g_{k,M,N-n} = \tfrac{N-n+1}{\Omega_{k,M,N}} g_{k,M,N-n+1} + \tfrac{1}{\Omega_{k,M,N}} f_{k,M,N-n}, \tag{1.7.23b}$$

$$g_{k,M-m,N} = \tfrac{1}{\Omega_{k,M,N}} f_{k,M-m,N} - \tfrac{M-m+1}{\Omega_{k,M,N}} g_{k,M-m+1,N}, \tag{1.7.23c}$$

$$\begin{aligned} g_{k,M-m,N-n} = {} & \tfrac{1}{\Omega_{k,M,N}} f_{k,M-m,N-n} - \tfrac{M-m+1}{\Omega_{k,M,N}} g_{k,M-m+1,N-n} \\ & + \tfrac{N-n+1}{\Omega_{k,M,N}} g_{k,M-m,N-n+1} \end{aligned} \tag{1.7.23d}$$

($m = 1, \ldots, M$; $n = 1, \ldots, N$); here, the index k has been explicitly written. By complex conjugation of the first equation, $\overline{g}_{k,M,N} = \overline{f}_{k,M,N}/\overline{\Omega}_{k,M,N}$; but $\overline{f}_{k,M,N} = (-1)^{M+N} f_{-k,N,M}$, and $\overline{\Omega}_{k,M,N} = \Omega_{-k,N.M}$, therefore $\overline{g}_{k,M,N} = (-1)^{M+N} g_{-k,N,M}$. Thus, (1.7.23b) is true for $n = 1$, since,

$$\begin{aligned} \overline{g}_{k,M,N-1} &= \frac{1}{\overline{\Omega}_{k,M,N}} \left(N\,\overline{g}_{k,M,N} + \overline{f}_{k,M,N-1}\right) \\ &= \frac{(-1)^{M+N-1}}{\Omega_{-k,N,M}} \left(f_{-k,N-n,M} - N\,g_{-k,N,M}\right) \end{aligned}$$

but by (1.7.23c), interchanging M and N, this last expression equals to $g_{-k,N-1,M}$. Now, suppose all the previous $g_{k,M,N-2}, g_{k,M,N-3}, \ldots, g_{k,M,N-n+1}$ satisfy the $\mathcal{S}$-symmetry. Then, for $g_{k,M,N-n}$, using (1.7.23b),

$$\begin{aligned} \overline{g}_{k,M,N-n} &= \frac{1}{\overline{\Omega}_{k,M,N-n}} \left((N-n+1)\overline{g}_{k,M,N-n+1} + \overline{f}_{k,M,N}\right) \\ &= \frac{(-1)^{M+N-n}}{\Omega_{-k,N,M}} \left(f_{-k,N-n,M} - (N-n+1)\,g_{-k,N-n+1,M}\right), \end{aligned}$$

which equals –again by (1.7.23b)–, to $(-1)^{M+N-n} g_{-k,N-n,M}$. The reader can apply the same induction arguments to (1.7.23b) and (1.7.23c) to prove that also:

$$\begin{aligned} \overline{g}_{k,M-m,N} &= (-1)^{M+N-m} g_{-k,N,M-m}, \\ \text{and } \overline{g}_{k,M-m,N-n} &= (-1)^{M-m+N-n} g_{-k,N-n,M-m}. \end{aligned}$$

♠

With this remark, we can complete the lemma 1.10 adding.

Lemma 1.13. *If the monomials of $F_s \in \mathfrak{E}_s$ satisfy the $\mathcal{S}$-symmetries and none of them are of $\mathcal{M}$-type; then, the equations* (1.7.11) *have an unique solution, $G_s \in \mathfrak{E}_s$, which fulfills also the $\mathcal{S}$-symmetries.*

1.7.2 An study of the $\mathcal{M}$-type resonant monomials

Now we are going to investigate the solvability of the homological equations when $\mathcal{M}$-type monomials are taken into account. Thus, we restrict the linear operator L_{H_2} to the space $\mathfrak{E}_{0,l,M,M} \subset \mathfrak{E}_s$, $(s = 2l + 2M)$. Even more, consider the space $\mathcal{P}_M$ of homogeneous polynomials of degree $2M$ in q_1, q_2, p_1, p_2, with complex coefficients,

$$\widehat{F} \in \mathcal{P}_M \Leftrightarrow \widehat{F} = \sum_{j,\nu=0}^{M} \widehat{F}_{j,\nu}\, q_1^j q_2^{M-j} p_1^{M-\nu} p_2^{\nu}. \tag{1.7.24}$$

Therefore, $\mathfrak{E}_{0,l,M,M} = \{I^l \widehat{F}$, with $\widehat{F} \in \mathcal{P}_M\}$, and we put symbolically $\mathfrak{E}_{0,l,M,M} = I^l \mathcal{P}_M$. Moreover, the operator $\widehat{L}$ is defined by,

$$\begin{aligned} \widehat{L} : \mathcal{P}_M \;&\rightarrow\; \mathcal{P}_M \\ \widehat{F} \;&\mapsto\; \widehat{L}\widehat{F} = \{\widehat{F}, H_2\}, \end{aligned}$$

with H_2 given in (1.6.4) and being $\{\widehat{F}, H_2\} = \sum_{i=1}^{2} \left(\frac{\partial \widehat{F}}{\partial q_i}\frac{\partial H_2}{\partial p_i} - \frac{\partial \widehat{F}}{\partial p_i}\frac{\partial H_2}{\partial q_i} \right)$. $\widehat{L}$ can be thought of as the restriction of L_{H_2} on $\mathcal{P}_M$. It has then full sense to consider the *reduced* homological equations defined on $\mathcal{P}_M$ by

$$\widehat{L}\widehat{G} + \widehat{Z} = \widehat{F}. \tag{1.7.25}$$

So, if $\widehat{G}$, $\widehat{Z} \in \mathcal{P}_M$ satisfy the above equation, then $I^l\widehat{G}$, $I^l\widehat{Z}$ will be a solution of (1.7.25). Moreover, the following decomposition works,

$$\mathcal{P}_M = \mathrm{Range}\,\widehat{L} \oplus \mathrm{Ker}\,\widehat{L}^{\dagger} \tag{1.7.26}$$

where $\widehat{L}^{\dagger}$ is the adjoint operator of $\widehat{L}$. In fact, $\mathrm{Ker}\,\widehat{L}^{\dagger} = (\mathrm{Range}\,\widehat{L})^{\perp}$ (the kernel of the adjoint operator of $\widehat{L}$ is the orthogonal complement of its image). Thus, the resonant terms will lie in this space.

If should $\widehat{L}_{H_2}$ (and hence $\widehat{L}$) be self adjoint, giving rise to diagonal homological equations, then $\mathrm{Ker}\,\widehat{L} = \mathrm{Ker}\,\widehat{L}^{\dagger}$. But in view of (1.7.15), this is clearly not the case, and we need to introduce a suitable inner product in $\mathcal{P}_M$ to determine easily the adjoint $\widehat{L}^{\dagger}$. According to Elphick et al. (1987), given two homogeneous polynomials $F = \sum_{\boldsymbol{m},\boldsymbol{n}} F_{\boldsymbol{m},\boldsymbol{n}} \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}}$ and $G = \sum_{\boldsymbol{m}',\boldsymbol{n}'} G_{\boldsymbol{m}',\boldsymbol{n}'} \boldsymbol{q}^{\boldsymbol{m}'} \boldsymbol{p}^{\boldsymbol{n}'}$, we define their bracket by,

$$\langle F|G\rangle = \sum_{\substack{m_1,m_2,n_1,n_2\\ m_1',m_2',n_1',n_2'}} F_{m_1,m_2,n_1,n_2} \overline{G}_{m_1',m_2',n_1',n_2'} \frac{\partial^{m_1+m_2+n_1+n_2}}{\partial q_1^{m_1}\partial q_2^{m_2}\partial p_1^{n_1}\partial p_2^{n_2}} \left(q_1^{m_1'} q_2^{m_2'} p_1^{n_1'} p_2^{n_2'} \right)\Big|_{(0,0,0,0)} \tag{1.7.27}$$

$$= \sum_{\substack{m_1,m_2,n_1,n_2\\ m_1',m_2',n_1',n_2'}} m_1!\,m_2!\,n_1!\,n_2!\, F_{m_1,m_2,n_1,n_2} \overline{G}_{m_1',m_2',n_1',n_2'} \delta_{m_1,m_1'} \delta_{m_2,m_2'} \delta_{n_1,n_1'} \delta_{n_2,n_2'} \tag{1.7.28}$$

$$= \sum_{m_1,m_2,n_1,n_2} m_1!\, m_2!\, n_1!\, n_2!\, F_{m_1,m_2,n_1,n_2} \overline{G}_{m_1,m_2,n_1,n_2}. \tag{1.7.29}$$

Obviously, in this last formula, $\frac{\partial^{m_1+m_2+n_1+n_2}}{\partial q_1^{m_1}\partial q_2^{m_2}\partial p_1^{n_1}\partial p_2^{n_2}}$ takes the partial derivatives m_1 times w. r. t. q_1, m_2 times w. r. t. to q_2 and so on. It is straightforward to check that this bracket satisfies all the properties of a (complex) Hermitian product: $\langle F|G\rangle = \overline{\langle G|F\rangle}$, and so on. In particular, for two polynomials of $\mathcal{P}_M$, say

$$\widehat{F} = \sum_{j,\nu=0}^{M} \widehat{F}_{j,\nu}\, q_1^j q_2^{M-j} p_1^{M-\nu} p_2^{\nu} \tag{1.7.30}$$

and

$$\widehat{G} = \sum_{j',\nu'=0}^{M} \widehat{G}_{j',\nu'}\, q_1^{j'} q_2^{M-j'} p_1^{M-\nu'} p_2^{\nu'}, \tag{1.7.31}$$

the bracket $\langle \widehat{F}|\widehat{G}\rangle$ results,

$$\langle \widehat{F}|\widehat{G}\rangle = \sum_{j,\nu=0}^{M} j!\,(M-j)!\,(M-\nu)!\,\nu!\,\widehat{F}_{j,\nu}\overline{\widehat{G}}_{j,\nu}, \tag{1.7.32}$$

and the next lemma determines the operator $\widehat{L}^{\dagger}$.

Lemma 1.14. *The operator* $\widehat{L}^{\dagger} = \{\cdot, H_2^{\dagger}\}$, *with* $H_2^{\dagger} = q_1 p_2$, *is the adjoint operator of* $\widehat{L}$ *with respect to the Hermitian product defined by* (1.7.32).

Proof. Let $\widehat{F}$ and $\widehat{G}$ be two polynomials in $\mathcal{P}_M$ expressed both as in (1.7.30) and (1.7.31). Direct computation shows,

$$\begin{aligned}\widehat{L}\widehat{F} = &\sum_{j=0}^{M-1}\sum_{\nu=0}^{M}(j+1)\,\widehat{F}_{j+1,\nu}\, q_1^j\, q_2^{M-j} p_1^{M-\nu} p_2^{\nu}\\ &-\sum_{j=0}^{M}\sum_{\nu=0}^{M-1}(\nu+1)\,\widehat{F}_{j,\nu+1}\, q_1^j\, q_2^{M-j} p_1^{M-\nu} p_2^{\nu},\end{aligned}$$

$$\begin{aligned}\widehat{L}^{\dagger}\widehat{G} = &\sum_{j=1}^{M}\sum_{\nu=0}^{M}(M-j+1)\,\widehat{G}_{j-1,\nu}\, q_1^j\, q_2^{M-j} p_1^{M-\nu} p_2^{\nu}\\ &-\sum_{j=0}^{M}\sum_{\nu=1}^{M}(M-\nu+1)\,\widehat{G}_{j,\nu-1}\, q_1^j\, q_2^{M-j} p_1^{M-\nu} p_2^{\nu}.\end{aligned}$$

Now, writing explicitly the product $\langle \widehat{L}\widehat{F}|\widehat{G}\rangle$

$$\begin{aligned}\langle \widehat{L}\widehat{F}|\widehat{G}\rangle = &\sum_{j=0}^{M-1}\sum_{\nu=0}^{M}(j+1)!\,(M-j)!\,(M-\nu)!\,\nu!\,\widehat{F}_{j+1,\nu}\,\overline{\widehat{G}}_{j,\nu}\\ &-\sum_{j=0}^{M}\sum_{\nu=0}^{M-1} j!\,(M-j)!\,(M-\nu)!\,(\nu+1)!\,\widehat{F}_{j,\nu+1}\,\overline{\widehat{G}}_{j,\nu}\end{aligned} \tag{1.7.33}$$

and, in the same way,

$$\langle \widehat{F}|\widehat{L}^{\dagger}\widehat{G}\rangle = \sum_{j=1}^{M}\sum_{\nu=0}^{M} j!\,(M-j+1)!\,(M-\nu)!\,\nu!\,\widehat{F}_{j,\nu}\,\overline{\widehat{G}}_{j-1,\nu}$$
$$-\sum_{j=0}^{M}\sum_{\nu=0}^{M} j!\,(M-j)!\,(M-\nu+1)!\,\nu!\,\widehat{F}_{j,\nu}\,\overline{\widehat{G}}_{j,\nu-1} \quad (1.7.34)$$

which equals (1.7.33), as a shift $j \mapsto j+1$ in the indices of the first sum and $\nu \mapsto \nu+1$ in the indices of the second shows. Then, $\langle \widehat{L}\widehat{F}|\widehat{G}\rangle = \langle \widehat{F}|\widehat{L}^{\dagger}\widehat{G}\rangle$, and the proof of the lemma is completed. $\square$

Now let $\widehat{F}$ be a polynomial in $\mathcal{P}_M$, then $\widehat{F} \in \operatorname{Ker} \widehat{L}^{\dagger}$ if and only if $\widehat{L}^{\dagger}\widehat{F} = \{\widehat{F}, H_2^{\dagger}\} = 0$. Thus, $\widehat{F}$ must be a solution of the partial differential equation (see Schmidt, 1994),

$$-p_2\frac{\partial u}{\partial p_1} + q_1\frac{\partial u}{\partial q_2} = 0. \quad (1.7.35)$$

A function u is a solution of a first order linear equation like this, if and only if it is a first integral of the associated characteristic system (see Arnol'd, 1978, Chap. 2, §7.B, or eventually, any textbook on differential equations).

For (1.7.35) the mentioned characteristic equations (i. e., the Hamiltonian equations of $H_2^{\dagger}$) are,

$$\dot{q}_1 = 0, \qquad \dot{q}_2 = q_1, \qquad \dot{p}_1 = -p_2, \qquad \dot{p}_2 = 0, \quad (1.7.36)$$

then, the functions, q_1, p_2, $q_1p_1 + q_2p_2$ are first integrals of the system (1.7.36). If we want solutions in $\mathcal{P}_M$, we have to consider the following first integrals

$$\xi_1 = \frac{\mathrm{i}}{2}\left(q_1p_1 + q_2p_2\right), \qquad \xi_2 = q_1p_2. \quad (1.7.37)$$

(Note that ξ_1, ξ_2 are real under the symmetries (1.6.5) introduced by the complexification). Therefore, any homogeneous polynomial $\widehat{S}$ of degree M in ξ_1, ξ_2

$$\widehat{S} = \sum_{j=0}^{M} a_j \xi_1^{M-j}\xi_2^{j}, \quad (1.7.38)$$

will be a solution of (1.7.36), and hence it must be in $\operatorname{Ker}\widehat{L}^{\dagger}$.

Remark 1.15. With the same aim of remark 1.12, we introduce here the subset $\mathcal{P}_M^{\mathcal{S}}$ constituted by those polynomials of $\mathcal{P}_M$ satisfying the $\mathcal{S}$-symmetries, i. e., for $\widehat{F} \in \mathcal{P}_M$, written as in (1.7.30),

$$\widehat{F} \in \mathcal{P}_M^{\mathcal{S}} \Leftrightarrow \overline{\widehat{F}}_{j,\nu} = (-1)^{j+\nu}\widehat{F}_{\nu,j}.$$

One can check that $\mathcal{P}_M^{\mathcal{S}}$ is a *real* linear subspace of $\mathcal{P}_M$ with $\dim \mathcal{P}_M^{\mathcal{S}} = (M+1)^2$, the number of independent real coefficients, taking into account the symmetries –see the proof of lemma 1.16–. On the other hand, the Hermitian product defined by (1.7.32) restricted to $\mathcal{P}_M^{\mathcal{S}}$ is a real inner product. To see this, take $\widehat{F}, \widehat{G} \in \mathcal{P}_M^{\mathcal{S}}$ and realize that,

$$\begin{aligned}\langle \widehat{F}|\widehat{G}\rangle = \ldots + j!\,(M-j)!\,(M-\nu)!\,\nu!\,\widehat{F}_{j,\nu}\overline{\widehat{G}}_{j,\nu} + \ldots \\ \ldots + \nu!\,(M-\nu)!\,(M-j)!\,j!\,\widehat{F}_{\nu,j}\overline{\widehat{G}}_{\nu,j} + \ldots \\ = \ldots + (-1)^{j+\nu} j!\,(M-j)!\,(M-\nu)!\,\nu!\,\widehat{F}_{j,\nu}\widehat{G}_{\nu,j} + \ldots \\ \ldots + (-1)^{j+\nu}\nu!\,(M-\nu)!\,(M-j)!\,j!\overline{\widehat{F}}_{j,\nu}\overline{\widehat{G}}_{\nu,j} + \ldots \\ = \ldots + 2(-1)^{j+\nu} j!\,(M-j)!\,(M-\nu)!\,\nu!\,\mathrm{Re}\,(\widehat{F}_{j,\nu}\widehat{G}_{j,\nu}) + \ldots\end{aligned}$$

which is a real number[4]. Thus, $\langle \widehat{F}|\widehat{G}\rangle = \langle \widehat{G}|\widehat{F}\rangle$ and the rest of the properties for the real inner product follow.

Moreover, as the $\mathcal{S}$-symmetries are preserved under L_{H_2}, and hence under $\widehat{L}$, we can consider the restriction of $\widehat{L}$ to $\mathcal{P}_M^{\mathcal{S}}$ (say, $\widehat{L}^{\mathcal{S}} = \widehat{L}|_{\mathcal{P}_M^{\mathcal{S}}}$) and apply there the same decomposition (1.7.26) to have $\mathcal{P}_M^{\mathcal{S}} = \mathrm{Range}\,(L^{\mathcal{S}}) \oplus \mathrm{Ker}\,\left(L^{\mathcal{S}}\right)^{\dagger}$. ♠

Lemma 1.16. *$\mathfrak{Q} = \{\xi_1^{M-j}\xi_2^j\}_{j=0,\ldots,M}$, with ξ_1, ξ_2 given by (1.7.37), is a basis of the kernel of the adjoint operator of $\widehat{L}^{\mathcal{S}}$, considered as a real subspace, i. e.,*

$$\mathrm{Ker}(\widehat{L}^{\mathcal{S}})^{\dagger} = \mathrm{Span}\mathfrak{Q}. \tag{1.7.39}$$

Therefore, any $\widehat{S}$ in $\mathrm{Ker}(\widehat{L}^{\mathcal{S}})^{\dagger}$ can be expressed by a linear combination as in (1.7.38), now with real coefficients.

Proof. First of all, we stress that $\xi_1^{M-j}\xi_2^j \in \mathcal{P}_M^{\mathcal{S}}$, for all $j = 0, \ldots, M$, since

$$\overline{q}_1 = -p_2,\ \overline{q}_2 = p_1,\ \overline{p}_1 = q_2,\ \overline{p}_2 = -q_1 \Rightarrow \overline{\xi_1^{M-j}\xi_2^j} = \xi_1^{M-j}\xi_2^j.$$

Next, let $\widehat{F} \in \mathcal{P}_M^{\mathcal{S}}$. As any monomial $q_1^j q_2^{M-j} p_1^{M-\nu} p_2^{\nu}$, can be written as,

$$q_1^j q_2^{M-j} p_1^{M-\nu} p_2^{\nu} = \begin{cases} (q_1p_1)^j (q_2p_2)^{\nu} (q_2p_1)^{M-j-\nu} & \text{if } j+\nu \le M, \\ (q_1p_1)^{M-\nu} (q_2p_2)^{M-j} (q_1p_2)^{j+\nu-M} & \text{if } j+\nu > M, \end{cases} \tag{1.7.40}$$

$\widehat{F}$ can be written into the form,

$$\widehat{F} = \sum_{\nu=1}^{M} \eta_2^{\nu} \sum_{\nu+m+n=M} \widehat{f}_{\nu,m,n}\,\eta_1^m\eta_3^n + \sum_{\nu=0}^{M} \eta_4^{\nu} \sum_{\nu+m+n=M} \widehat{g}_{\nu,m,n}\eta_1^m\eta_3^n, \tag{1.7.41}$$

where,

$$\eta_1 = q_1p_1, \quad \eta_2 = q_1p_2, \quad \eta_3 = q_2p_2, \quad \eta_4 = q_2p_1 \tag{1.7.42}$$

and with the symmetries,

$$\overline{\widehat{f}}_{l,m,n} = (-1)^{m+n}\widehat{f}_{l,n,m}, \quad \text{and} \quad \overline{\widehat{g}}_{l,m,n} = (-1)^{m+n}\widehat{g}_{l,n,m}. \tag{1.7.43}$$

[4] Observe that if $j = \nu$, then the coefficients $\widehat{F}_{j,\nu}$ and $\widehat{G}_{j,\nu}$ are real.

Then, it is clear that,

$$\mathfrak{M} = \{\eta_2^\nu \eta_1^m \eta_3^n : 1 \le \nu \le M \text{ and } \nu + m + n = M\} \\ \cup \{\eta_4^\nu \eta_1^m \eta_3^n : 0 \le \nu \le M \text{ and } \nu + m + n = M\},$$

are a basis of $\mathcal{P}_M$, hence, the dimension of $\mathcal{P}_M^{\mathcal{S}}$ is given by the number of real independent coefficients. It is not difficult to check that, for a fixed ν in the sums above, the quantity of real and imaginary parts –not related by the symmetries (1.7.43)–, necessary to form all the complex coefficients $\widehat{f}_{\nu,m,n}$ are $M - \nu + 1$ (and the same for $\widehat{g}_{\nu,m,n}$). Thus, summation over ν gives,

$$\dim \mathcal{P}_M^{\mathcal{S}} = \sum_{\nu=1}^{M} (M - \nu + 1) + \sum_{\nu=0}^{M} (M - \nu + 1) = (M+1)^2.$$

Let ξ_3 and ξ_4 be the two degree polynomials defined as

$$\xi_3 = \frac{1}{2}(q_1 p_1 - q_2 p_2), \qquad \xi_4 = q_2 p_1, \tag{1.7.44}$$

and we introduce the system of $(M+1)^2$ vectors of $\mathcal{P}_M^{\mathcal{S}}$ through

$$\mathfrak{T} = \{\xi_2^\nu \xi_1^m \xi_3^n : 1 \le \nu \le M \text{ and } \nu + m + n = M\} \\ \cup \{\xi_4^\nu \xi_1^m \xi_3^n : 0 \le \nu \le M \text{ and } \nu + m + n = M\}, \tag{1.7.45}$$

(with ξ_1, ξ_2 defined in (1.7.37)). It turns out that they are linearly independent in $\mathcal{P}_M$, because the map $(\xi_1, \xi_3) \mapsto (\eta_1, \eta_3)$ is linear and invertible, in fact $\left(\begin{smallmatrix} \xi_1 \\ \xi_3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} i/2 & i/2 \\ 1/2 & -1/2 \end{smallmatrix}\right)\left(\begin{smallmatrix} \eta_1 \\ \eta_3 \end{smallmatrix}\right)$; so $\mathfrak{T}$ constitutes a *real* basis of $\mathcal{P}_M^{\mathcal{S}}$ (in the sense that all the elements of $\mathcal{P}_M^{\mathcal{S}}$ can be put as linear combinations of elements of $\mathfrak{T}$ with *real* coefficients). We observe that,

(i) $\mathfrak{T}$ *completes* the basis $\mathfrak{Q}$ (see lemma 1.16) to a basis of $\mathcal{P}_M^{\mathcal{S}}$, since $\mathfrak{Q} \subset \mathfrak{T}$.

(ii) For a given $\widehat{F} \in \mathcal{P}_M^{\mathcal{S}} \setminus \operatorname{Span} \mathfrak{Q}$, there exists $\widehat{G} \in \mathcal{P}_M^{\mathcal{S}}$ (not unique) such that $\widehat{L}^{\mathcal{S}} \widehat{G} = \widehat{F}$.

To show this last point we take $\widehat{F}$ and $\widehat{G}$ and write them down in the basis $\mathfrak{T}$, so

$$\widehat{F} = \sum_{\substack{0 \le \nu < M \\ \nu+m+n=M \\ (\text{with } n \ne 0)}} \widehat{f}_{\nu,m,n}\, \xi_2^\nu \xi_1^m \xi_3^n + \sum_{\substack{1 \le \nu \le M \\ \nu+m+n=M}} \widehat{g}_{\nu,m,n}\, \xi_4^\nu \xi_1^m \xi_3^n \tag{1.7.46a}$$

$$\widehat{G} = \sum_{\substack{1 \le \nu \le M \\ \nu+m+n=M}} f_{\nu,m,n}\, \xi_2^\nu \xi_1^m \xi_3^n + \sum_{\substack{0 \le \nu \le M \\ \nu+m+n=M}} g_{\nu,m,n}\, \xi_4^\nu \xi_1^m \xi_3^n. \tag{1.7.46b}$$

Remark 1.17. We note here that $\widehat{F}$ does not hold resonant terms (these would be precisely the ones with $n = 0$ in the first sum, but this value for the index n has been explicitly excluded). Further, we stress that the sums defining $\widehat{F}$ and $\widehat{G}$ are arranged in a different way (note specially the ranges of the index ν). ♠

On the other hand, the operator $\widehat{L}^{\mathcal{S}}$ acting on the elements of the basis $\mathfrak{T}$ produces,

$$\begin{aligned}\widehat{L}^{\mathcal{S}}\left(\xi_2^{\nu}\xi_1^{m}\xi_3^{n}\right) &= \frac{\partial}{\partial\xi_2}\left(\xi_2^{\nu}\xi_1^{m}\xi_3^{n}\right)\{\xi_2,H_2\}+\frac{\partial}{\partial\xi_1}\left(\xi_2^{\nu}\xi_1^{m}\xi_3^{n}\right)\{\xi_1,H_2\}\\ &\quad+\frac{\partial}{\partial\xi_3}\left(\xi_2^{\nu}\xi_1^{m}\xi_3^{n}\right)\{\xi_3,H_2\},\\ \widehat{L}^{\mathcal{S}}\left(\xi_4^{\nu}\xi_1^{m}\xi_3^{n}\right) &= \frac{\partial}{\partial\xi_4}\left(\xi_4^{\nu}\xi_1^{m}\xi_3^{n}\right)\{\xi_4,H_2\}+\frac{\partial}{\partial\xi_1}\left(\xi_4^{\nu}\xi_1^{m}\xi_3^{n}\right)\{\xi_1,H_2\}\\ &\quad+\frac{\partial}{\partial\xi_3}\left(\xi_4^{\nu}\xi_1^{m}\xi_3^{n}\right)\{\xi_3,H_2\},\end{aligned}$$

but,

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

Therefore,

$$\widehat{L}^{\mathcal{S}}\left(\xi_2^{\nu}\xi_1^{m}\xi_3^{n}\right)=-2\nu\xi_2^{\nu-1}\xi_1^{m}\xi_3^{n+1}+n\xi_2^{\nu}\xi_1^{m}\xi_3^{n-1}\xi_4,$$

and when $1\le\nu\le M$, we can arrange the second term on the right using that $\xi_2\xi_4=-\xi_1^2-\xi_3^2$, so $n\xi_2^{\nu-1}\xi_1^{m}\xi_3^{n-1}\xi_2\xi_4=-n\xi_2^{\nu-1}\xi_1^{m+2}\xi_3^{n-1}-n\xi_2^{\nu-1}\xi_1^{m}\xi_3^{n+1}$, which when it is joined to the first term

$$\widehat{L}^{\mathcal{S}}\left(\xi_2^{\nu}\xi_1^{m}\xi_3^{n}\right)=-(2\nu+n)\xi_2^{\nu-1}\xi_1^{m}\xi_3^{n+1}-n\xi_2^{\nu-1}\xi_1^{m+2}\xi_3^{n-1}, \tag{1.7.47}$$

with $1\le\nu\le M$ and $\nu+m+n=M$. Similarly,

$$\widehat{L}^{\mathcal{S}}\left(\xi_4^{\nu}\xi_1^{m}\xi_3^{n}\right)=n\xi_4^{\nu+1}\xi_1^{m}\xi_3^{n-1} \tag{1.7.48}$$

for $0\le\nu\le M$, $\nu+m+n=M$. With (1.7.47) and (1.7.48) we can compute the action of $\widehat{L}^{\mathcal{S}}$ on $\widehat{G}$ and write down explicitly the equation $\widehat{L}^{\mathcal{S}}\widehat{G}=\widehat{F}$ in the unknown real coefficients $f_{\nu,m,n}$ ($1\le\nu\le M$, $\nu+m+n=M$), and $g_{\nu,m,n}$ ($0\le\nu\le M$, $\nu+m+n=M$). In this way, one gets,

$$\begin{aligned}-\sum_{\nu=1}^{M}\xi_2^{\nu-1}&\sum_{\nu+m+n=M}\left((2\nu+n)f_{\nu,m,n}\,\xi_1^{m}\xi_3^{n+1}+nf_{\nu,m,n}\,\xi_1^{m+2}\xi_3^{n-1}\right)\\ &+\sum_{\nu=0}^{M}\xi_4^{\nu+1}\sum_{\nu+m+n=M}ng_{\nu,m,n}\,\xi_1^{m}\xi_3^{n-1}\\ =\sum_{\nu=0}^{M-1}\xi_2^{\nu}&\sum_{\substack{\nu+m+n=M\\(n\neq0)}}\widehat{f}_{\nu,m,n}\,\xi_1^{m}\xi_3^{n}+\sum_{\nu=1}^{M}\xi_4^{\nu}\sum_{\nu+m+n=M}\widehat{g}_{\nu,m,n}\,\xi_1^{m}\xi_3^{n}.\end{aligned} \tag{1.7.49}$$

By comparison of coefficients in the second sums of both sides, we arrive to the relations:

$$g_{\nu-1,m,n+1}=\frac{\widehat{g}_{\nu,m,n}}{n+1}, \tag{1.7.50}$$

with $1\le\nu\le M$ ($\nu+m+n=M$). But (1.7.50) do not determine $g_{\nu,m,0}$ ($\nu+m=M$), so these can be chosen arbitrarily (for they play no rôle in the homological equations). In particular we shall set them to zero, i. e., we take: $g_{\nu,m,0}=0$ ($\nu+m=M$). Similarly, comparison of coefficients in the first sums on the left and on the right hand side of (1.7.49), with $0 \leq \nu < M$ held fixed, leads to the linear system,

$$\widehat{f}_{\nu,0,M-\nu} = -(M+\nu+1) f_{\nu+1,0,M-\nu-1}, \tag{1.7.51a}$$

$$\widehat{f}_{\nu,1,M-\nu-1} = -(M+\nu) f_{\nu+1,1,M-\nu-2}, \tag{1.7.51b}$$

$$\begin{aligned}\widehat{f}_{\nu,m,M-\nu-m} = & -(M+\nu-m+1) f_{\nu+1,m,M-\nu-m-1} \\ & -(M-\nu-m+1) f_{\nu+1,m-2,M-\nu-m+1}\end{aligned} \tag{1.7.51c}$$

with $2 \leq m \leq M-\nu-1$ in the last equation. If we introduce the vectors $\boldsymbol{f}, \widehat{\boldsymbol{f}} \in \mathbb{R}^{M-\nu}$ by,

$$\boldsymbol{f} = \begin{pmatrix} f_{\nu+1,0,M-\nu-1} \\ f_{\nu+1,1,M-\nu-2} \\ f_{\nu+1,2,M-\nu-3} \\ \vdots \\ f_{\nu+1,M-\nu-2,1} \\ f_{\nu+1,M-\nu-1,0} \end{pmatrix}, \qquad \widehat{\boldsymbol{f}} = \begin{pmatrix} \widehat{f}_{\nu,0,M-\nu} \\ \widehat{f}_{\nu,1,M-\nu-1} \\ \widehat{f}_{\nu,2,M-\nu-2} \\ \vdots \\ \widehat{f}_{\nu,M-\nu-2,2} \\ \widehat{f}_{\nu,M-\nu-1,1} \end{pmatrix}, \tag{1.7.52}$$

for $\nu = 0, \ldots, M-1$. Then, the equations (1.7.51a)–(1.7.51c) can be expressed in vector notation as,

$$\Lambda' \boldsymbol{f} = \widehat{\boldsymbol{f}} \tag{1.7.53}$$

with the $(M-\nu) \times (M-\nu)$ dimensional matrix $\Lambda' = \Lambda'_\nu$ given by,

$$\Lambda' = \begin{pmatrix} -(M+\nu+1) & & & & & & \\ 0 & -(M+\nu) & & & & & \\ -(M-\nu-1) & 0 & -(M+\nu-1) & & & & \\ & -(M-\nu-2) & 0 & -(M+\nu-2) & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -3 & 0 & -(2\nu+3) \\ & & & & & -2 \quad 0 & -(2\nu+2) \end{pmatrix} \tag{1.7.54}$$

which is nonsingular for every $0 \leq \nu < M$. We conclude, then, that every $\widehat{F}$ in the complementary space of Span $\mathfrak{Q}$ belongs to Range $\widehat{L}^{\mathcal{S}}$, but Span $\mathfrak{Q} \subseteq \ker \widehat{L}^{\mathcal{S}}$, so this implies Span $\mathfrak{Q} \equiv \ker \widehat{L}^{\mathcal{S}}$. These considerations close the proof of lemma 1.16. $\square$

Remark 1.18. Observe that in the proof of the last lemma, we have set up the homological equations for the $\mathcal{M}$-type monomials, but to solve them and compute the corresponding terms of the generating function, we need to write the elements $F \in \mathcal{P}_M^{\mathcal{S}}$ from its natural form $F = \sum_{j,\alpha=0}^{M} F_{j,\alpha}\, q_1^j q_2^{M-j} p_1^{M-\alpha} p_2^{\alpha}$ to a polynomial in ξ_1, ξ_2, ξ_3 and ξ_4 of type,

$$F = \sum_{\nu=0}^{M} \xi_2^{\nu} \sum_{\nu+\vartheta+j=M} f_{\nu,\vartheta,j}\, \xi_1^{\vartheta}\xi_3^{j} + \sum_{\nu=1}^{M} \xi_4^{\nu} \sum_{\nu+\vartheta+j=M} g_{\nu,\vartheta,j}\, \xi_1^{\vartheta}\xi_3^{j},$$

to easily identify the projections of F onto Range $\widehat{L}^{\mathcal{S}}$ and $\ker(\widehat{L}^{\mathcal{S}})^{\dagger}$ (see the next lemma below). By direct substitution it can be seen that the relation between the coefficients of the two expressions for F are,

$$f_{\nu,\vartheta,j} = \mathrm{i}^{\vartheta} \sum_{m+n=M-\nu} (-1)^n C(m,n,\vartheta) F_{M-n,M-m}, \qquad \text{if } 0 \le \nu \le M \tag{1.7.55}$$

$$g_{\nu,\vartheta,j} = \mathrm{i}^{\vartheta} \sum_{m+n=M-\nu} (-1)^n C(m,n,\vartheta) F_{m,n}, \qquad \text{if } 1 \le \nu \le M \tag{1.7.56}$$

and the form factors $C(m,n,\vartheta)$ defined by,

$$C(m,n,\vartheta) = \sum_{\beta=\max(0,\vartheta-m)}^{\min(\vartheta,n)} \binom{m}{\vartheta-\beta}\binom{n}{\beta}(-1)^{\vartheta-\beta}. \tag{1.7.57}$$

Satisfying, in addition, the symmetries,

$$C(n,m,\vartheta) = (-1)^{\vartheta} C(m,n,\vartheta). \tag{1.7.58}$$

Then we can check immediately that the coefficients (1.7.55) and (1.7.56) are real. For example,

$$\begin{aligned}\overline{f}_{\nu,\vartheta,j} &= (-\mathrm{i})^{\vartheta} \sum_{m+n=M-\nu} (-1)^n C(m,n,\vartheta) \overline{F}_{M-n,M-m} \\ &= (-\mathrm{i})^{\vartheta} \sum_{m+n=M-\nu} (-1)^n (-1)^{\vartheta} C(n,m,\vartheta)(-1)^{2M-m-n} F_{M-m,M-n} \\ &= \mathrm{i}^{\vartheta} \sum_{m+n=M-\nu} (-1)^m C(n,m,\vartheta) F_{M-m,M-n}\end{aligned}$$

which equals to $f_{\nu,\vartheta,j}$ after the interchange of the dumb indices m and n. Identical computations show $\overline{g}_{\nu,\vartheta,j} = g_{\nu,\vartheta,j}$. ♠

Lemma 1.16 clearly proves the next one.

Lemma 1.19. *For any $\widehat{F} \in \mathcal{P}_M^{\mathcal{S}}$, there exist homogeneous polynomials $\widehat{G}, \widehat{Z} \in \mathcal{P}_M^{\mathcal{S}}$, satisfying the reduced homological equations* (1.7.25), *with*

$$\widehat{L}^{\mathcal{S}}\widehat{G} = \sqcap_1(\widehat{F}) \quad \text{and} \quad \widehat{Z} = \sqcap_2(\widehat{F}),$$

being $\sqcap_1$, $\sqcap_2$ the projections of the space $\mathcal{P}_M^{\mathcal{S}}$ onto Range $\widehat{L}^{\mathcal{S}}$ *and* $\ker(\widehat{L}^{\mathcal{S}})^{\dagger}$, *respectively.*

Furthermore, The complementary term $\widehat{Z}$ can be expressed as an homogeneous polynomial of degree M in ξ_1, ξ_2 –see (1.7.37)*–, with real coefficients.*

Lemmas 1.13 and 1.19, considered together, answer the question of the (formal) solvability of the homological equations (1.7.11), and allow us to end the present section with the theorem.

Theorem 1.20. *The homological equations* (1.7.11)*, with* $F_s \in \mathfrak{E}_{\mathfrak{s}}^{\mathcal{S}}$*,* $s \geq 3$*, are identically fulfilled with* $G_s, Z_s \in \mathfrak{E}_{\mathfrak{s}}^{\mathcal{S}}$*, where* $Z_s = 0$ *when* s *is odd or, when* s *is even,* Z_s *can be written as an homogeneous polynomial of degree* $s/2$ *in* $I_1, \frac{\mathrm{i}}{2}(q_1p_1 + q_2p_2), q_1p_2$*,*

$$Z_s = \sum_{l=0}^{s/2} \sum_{j=0}^{s/2-l} Z_{s,l,j} I_1^l \left(\frac{\mathrm{i}}{2}(q_1p_1 + q_2p_2)\right)^j (q_1p_2)^{s/2-l-j}, \tag{1.7.59}$$

with real coefficients $Z_{s,l,j}$*.*

1.7.3 The reduction algorithm

Once the solvability of the homological equations has been determined, we can go back to the application of the transformation algorithm described at definition 1.7 and continue with the study on how it can be used to reduce *formally* the initial Hamiltonian (1.6.2). In this sense, our early discussion (in the last paragraph before section 1.7.1) can be revisited and we may adapt a classical –for semisimple elliptic equilibrium[5] point– result which furnishes the necessary recursive formulas.

Proposition 1.21. *Consider the Hamiltonian* $H = \sum_{s\geq 2} H_s$ *defined in the complex domain* $\mathcal{D}(\rho^*, R^*)$ *given by* (1.6.3)*, with*

$$H_2 = \omega_1 I_1 + \mathrm{i}\,\omega_2(q_1p_1 + q_2p_2) + q_2p_1$$

and

$$H_s = \sum_{\substack{k\in\mathbb{Z} \\ 2l+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1=s}} h_{k,l,\boldsymbol{m},\boldsymbol{n}} I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \exp{(\mathrm{i}k\theta_1)},$$

for $s > 2$*. We shall assume, in addition, that the frequencies* ω_1, ω_2 *appearing in the quadratic part* H_2 *are rational independent, so* $\omega_1/\omega_2 \notin \mathbb{Q}$*. Then,*

(i) There exists a generating function $G^{(r)} = \sum_{s\geq 3}^{r} G_s$ *such that* $T_{G^{(r)}} H = Z^{(r)} + \mathfrak{R}^{(r)}$*. Here,* $Z^{(r)}$ *is the normal form up to order* r*, whereas* $\mathfrak{R}^{(r)}$ *stands for the remainder.*

(ii) The normal form can be expanded as $Z^{(r)} = \sum_{s=2}^{r} Z_s$*, with* $Z_s \equiv 0$ *when* s *is odd, and the following relations are satisfied:*

$$\begin{aligned} Z_2 &= H_2 \\ L_{H_2} G_s + Z_s &= F_s, \quad s \geq 3, \end{aligned} \tag{1.7.60}$$

where,

$$\begin{aligned} F_3 &= H_3, \\ F_s &= \sum_{j=1}^{s-3} \frac{j}{s-2} L_{G_{2+j}} Z_{s-j} + \sum_{j=1}^{s-2} \frac{j}{s-2} H_{2+j,s-j-2}, \quad s \geq 4. \end{aligned} \tag{1.7.61}$$

[5] Generically, we shall talk about semisimple equilibrium when the linearized differential equation has a diagonalizable coefficient matrix' (as, for example, in (Haller, 1999)).

Here, the quantities $H_{l,k}$ may be computed recursively from the formulas (1.7.6) *of the Giorgilli-Galgani algorithm 1.7, i. e.,*

$$\begin{aligned} H_{l,0} &= H_l, \\ H_{l,k} &= \sum_{j=1}^{k} \frac{j}{k} L_{G_{2+j}} H_{l,k-j}. \end{aligned} \tag{1.7.62}$$

(iii) The remainder $\mathfrak{R}^{(r)}$ is given by the sum:

$$\mathfrak{R}^{(r)} = \sum_{s \geq r+1} F_s. \tag{1.7.63}$$

with F_s those computed from (1.7.61) *taking $Z_{s-j} = 0$ when $s - j \geq r + 1$ ($s \geq r + 1$ and $1 \leq j \leq s - 3$).*

Remark 1.22. Diophantine conditions are not required to ensure convergence of $G^{(r)}$ for a given (and fixed) r, despite the presence of divisors of the form $k\omega_1 + (M - N)\omega_2$, because $M - N$ ranges between $-r$ and r and therefore, for all $k \in \mathbb{Z}$, these divisors are bounded from below whenever ω_1/ω_2 is not rational. ♠

Remark 1.23. Note that F_s depends on G_ν, Z_ν for $\nu = 3, \ldots, s - 1$ through the recursive application of the Poisson bracket; but this operation preserves the $\mathcal{S}$-symmetries. Thus, it is straightforward to realize –accepting the results of the proposition 1.21–, that the generating function $G^{(r)}$, the normal form $Z^{(r)}$ and the remainder $\mathfrak{R}^{(r)}$, will satisfy the reality conditions if H does. ♠

The proof of proposition 1.21 is formally identical to the proof of the corresponding classical one in Giorgilli et al. (1989) and in the memoir presented in the same year by Simó. There, the authors considered a n-degree of freedom polynomial Hamiltonian with quadratic part $H_2 = \mathrm{i} \sum_{j=1}^{n} \omega_j q_j p_j$; but the homological equations deduced from the transformation algorithm in these papers are exactly those at the second item of the preceding proposition (see also below). Thus, the proof of 1.21 is included here only for the sake of completeness.

Proof of proposition 1.21. By proposition 1.20, the analysis done at the beginning of section 1.7, shows that Z_3, G_3, Z_4, G_4 can be found satisfying (1.7.60) and (1.7.61) identically. Assume the same is true for $\nu \leq s - 1$. From (1.7.7), we must have, for $s \geq 3$

$$\begin{aligned} Z_s &= H_{2,s-2} + \sum_{j=1}^{s-2} H_{2+j,s-j-2} \\ &= \sum_{j=1}^{s-3} \frac{j}{s-2} L_{G_{2+j}} H_{2,s-j-2} + L_{G_s} H_2 + \sum_{j=1}^{s-2} H_{2+j,s-j-2}. \end{aligned} \tag{1.7.64}$$

Hence,

$$L_{H_2} G_s + Z_s = \sum_{j=1}^{s-3} \frac{j}{s-2} L_{G_{2+j}} H_{2,s-j-2} + \sum_{j=1}^{s-2} H_{2+j,s-j-2}, \tag{1.7.65}$$

and one wants to prove that the right hand side of this equation equals to F_s in (1.7.61). To show this, we substitute (1.7.7) and (1.7.6) into (1.7.61) to obtain,

$$F_s = \sum_{j=1}^{s-3} \frac{j}{s-2} L_{G_{2+j}} \left(\sum_{m=0}^{s-j-2} H_{2+m,s-j-m-2} \right) + \sum_{j=1}^{s-3} \frac{j}{s-2} \left(\sum_{m=1}^{s-j-2} \frac{m}{s-j-2} L_{G_{2+m}} H_{2+j,s-j-m-2} \right) + H_s, \quad (1.7.66)$$

We split the first term in two pieces, the first holding only the term with $m = 0$. In the latter part, we permute the sums $\sum_{j=1}^{s-3} \sum_{m=1}^{s-j-2} = \sum_{j,m \geq 1, j+m \leq s-2} = \sum_{m=1}^{s-3} \sum_{j=1}^{s-m-2}$ and then exchange the names of the summation ("dummy") indices m, j. After these manipulations such part is similar to the second term and they can be grouped together adding their corresponding coefficients: $\frac{m}{s-2} + \frac{mj}{(s-j-2)(s-2)} = \frac{m}{s-j-2}$. Therefore,

$$\begin{aligned} F_s &= \sum_{j=1}^{s-3} \frac{j}{s-2} L_{G_{2+j}} H_{2,s-j-2} + \sum_{j=1}^{s-3} \sum_{m=1}^{s-j-2} \frac{m}{s-j-2} L_{G_{2+m}} H_{2+j,s-j-m-2} + H_s \\ &= \sum_{j=1}^{s-3} \frac{j}{s-2} L_{G_{2+j}} H_{2,s-j-2} + \sum_{j=1}^{s-2} H_{2+j,s-j-2}. \end{aligned}$$

Finally, proposition 1.20 assures that now G_s and Z_s can be found from the equation $L_{H_2} G_s + Z_s = F_s$. □

The following result is readily deduced from propositions 1.20, 1.21 and will be the starting point of the next section.

Theorem 1.24. *Consider the Hamiltonian system* (1.2.1). *As described in section 1.2, we suppose it has a family of periodic orbits depending on a parameter σ, $\{\mathcal{M}_\sigma\}_{\sigma \in \mathbb{R}}$, such that for some value, say $\sigma = 0$, the monodromy matrix of the (critical, resonant) periodic orbit $\mathcal{M}_0$, can be put in the Jordan form* (1.2.2). *Moreover, let $\vartheta = 2\pi\nu$ be the nontrivial characteristic exponent of $\mathcal{M}_0$. We assume $\nu \notin \mathbb{Q}$.*

Then, the initial Hamiltonian H, can be reduced by means of a symplectic $\mathcal{S}$-symmetric change defined in a complexified neighborhood of the periodic orbit $\mathcal{M}_0$, to a (complex) Hamiltonian given by the sum

$$H(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) = Z^{(r)}(\boldsymbol{q}, I_1, \boldsymbol{p}) + \mathfrak{R}^{(r)}(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}), \quad (1.7.67)$$

(we keep the same name for the new Hamiltonian), where the normal form $Z^{(r)}$ is given by the sum

$$Z^{(r)} = \sum_{s=2}^{r} Z_s,$$

with

$$Z_2 = \omega_1 I_1 + i\omega_2(q_1 p_1 + q_2 p_2) + q_2 p_1,$$

being ω_1 the frequency of the periodic orbit $\mathcal{M}_0$, $\omega_2 = \nu\omega_1$ and Z_s for $s \geq 3$ are the ones described in proposition 1.20; that is, $Z_s = 0$ for s odd or an homogeneous polynomial of degree $s/2$ in

$$I_1, \quad \frac{\mathrm{i}}{2}(q_1p_1 + q_2p_2), \quad q_1p_2,$$

with real coefficients, when s is even.

As the generating function verifies the reality conditions (1.7.13)*, it is advisable to apply additionally $\mathcal{S}^{-1}$ –the inverse of the complexifying transformation* (1.6.1)*–, to get a real transformed Hamiltonian (again without using new names for the functions),*

$$H(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}) = Z^{(r)}(\boldsymbol{x}, I_1, \boldsymbol{y}) + \mathfrak{R}^{(r)}(\theta_1, \boldsymbol{x}, I_1, \boldsymbol{y}). \tag{1.7.68}$$

Now, with $\theta_1, I_1, \boldsymbol{x}, \boldsymbol{y}$ real and the terms in $Z^{(r)}$ are given similarly by

$$Z_2 = \omega_1 I_1 + \omega_2(y_1x_2 - y_2x_1) + \frac{1}{2}(y_1^2 + y_2^2), \tag{1.7.69}$$

and, for $s \geq 3$, $Z_s = 0$ when s is odd or an homogeneous polynomial in

$$\tfrac{1}{2}(x_1^2 + x_2^2), \quad I_1, \quad \frac{1}{2}(y_1x_2 - y_2x_1),$$

with degree $s/2$, if s is even.

The sums $Z^{(r)} = \sum_{s=2}^{r} Z_s$ which appear in (1.7.68) and (1.7.69), are referred as the *complex* and the *real* normal form of the initial Hamiltonian respectively. Usually it is said that the initial Hamiltonian has been reduced to normal form up to order r. Under the hypotheses stated in theorem 1.24, it is clear that the described normal form process can be carried out up to any arbitrary order r.

1.8 Dynamics of the normal form

In this section we concentrate on the study of the normal form itself, so, from now on and up to the end of the present chapter, the Hamiltonian to consider will be the *real* normal form, $Z^{(r)}$ (i. e., we skip the remainder $\mathfrak{R}^{(r)}$ off). In view of theorem 1.24, $Z^{(r)}$ may be put into the form,

$$Z^{(r)} = \omega_1 I_1 + \omega_2 \boldsymbol{y} \times \boldsymbol{x} + \tfrac{1}{2}|\boldsymbol{y}|_2^2 + \mathcal{Z}_r\left(\tfrac{1}{2}|\boldsymbol{x}|_2^2, I_1, \boldsymbol{y} \times \boldsymbol{x}\right), \tag{1.8.1}$$

with the notation,

$$|\boldsymbol{x}|_2 = (x_1^2 + x_2^2)^{1/2}, \quad |\boldsymbol{y}|_2 = (y_1^2 + y_2^2)^{1/2}, \quad \boldsymbol{x} \times \boldsymbol{y} = x_1y_2 - x_2y_1$$

and $\mathcal{Z}_r(u, v, w)$ being a polynomial of degree $\lfloor r/2 \rfloor$, beginning with quadratic terms. We shall express it in the form,

$$\begin{aligned} \mathcal{Z}_r(u, v, w) = \frac{1}{2}(au^2 + bv^2 + cw^2) + duv + euw + fvw + \\ + \begin{cases} \sum_{3 \leq j+m+n \leq \lfloor r/2 \rfloor} f_{j,m,n} u^j v^m w^n, & \text{if } r \geq 6 \\ 0, & \text{if } r < 6 \end{cases} \end{aligned} \tag{1.8.2}$$

writing apart the terms of degree two, because their coefficients will play an essential role in the dynamics of the normal form (see below), also we define $\boldsymbol{\eta}(I_1, \boldsymbol{x}, \boldsymbol{y}) = \left(\frac{1}{2}|\boldsymbol{x}|_2^2, I_1, \boldsymbol{y} \times \boldsymbol{x}\right)$, hence

$$\mathcal{Z}_r \circ \boldsymbol{\eta} = \mathcal{Z}_r \left(\tfrac{1}{2}|\boldsymbol{x}|_2^2, I_1, \boldsymbol{y} \times \boldsymbol{x}\right)$$

and the Hamiltonian equations corresponding to the normal form can be written as,

$$\begin{aligned} \dot{\theta}_1 &= \omega_1 + \partial_2 \mathcal{Z}_r \circ \boldsymbol{\eta}, \\ \dot{I}_1 &= 0, \\ \dot{x}_1 &= \omega_2 x_2 + y_1 + x_2 \partial_3 \mathcal{Z}_r \circ \boldsymbol{\eta}, \\ \dot{x}_2 &= -\omega_2 x_1 + y_2 - x_1 \partial_3 \mathcal{Z}_r \circ \boldsymbol{\eta}, \\ \dot{y}_1 &= \omega_2 y_2 - x_1 \partial_1 \mathcal{Z}_r \circ \boldsymbol{\eta} + y_2 \partial_3 \mathcal{Z}_r \circ \boldsymbol{\eta}, \\ \dot{y}_2 &= -\omega_2 y_1 - x_2 \partial_1 \mathcal{Z}_r \circ \boldsymbol{\eta} - y_1 \partial_3 \mathcal{Z}_r \circ \boldsymbol{\eta}. \end{aligned} \tag{1.8.3}$$

Moreover, the normal form (1.8.1) is integrable, since it can be seen that the three quantities,

$$\mathcal{I}_1 = I_1, \quad \mathcal{I}_2 = \boldsymbol{x} \times \boldsymbol{y} \quad \text{and} \quad \mathcal{I}_3 = \tfrac{1}{2}|\boldsymbol{y}|_2^2 + \mathcal{Z}_r \left(\tfrac{1}{2}|\boldsymbol{x}|_2^2, I_1, \boldsymbol{x} \times \boldsymbol{y}\right) \tag{1.8.4}$$

are functionally independent first integrals outside the zero measure set defined by:

$$y_1 = 0, \qquad y_2 = 0, \qquad \partial_1 \mathcal{Z}_r = 0,$$

and invariant under the flow of (1.8.3).

1.8.1 Parametrization of the family of periodic orbits

It is straightforward to check that these Hamiltonian equations have a one-parameter family of periodic orbits given by

$$\mathcal{M}_{I_1} : \begin{cases} \theta_1 = (\omega_1 + \partial_2 \mathcal{Z}_r(0, I_1, 0))t + \theta_1^0, \\ I_1 = \text{const.}, \\ x_1 = x_2 = y_1 = y_2 = 0, \end{cases} \tag{1.8.5}$$

This implies that the action I_1 is a good parameter for the (local) description of the initial family of periodic orbits. So it can be denoted by $\{\mathcal{M}_{I_1}\}_{I_1 \in \mathbb{R}}$.

Remark 1.25. One may wonder if such parametrization is preserved when the remainder is added to the normal form and the complete transformed Hamiltonian is considered. In fact, the only monomials in $\mathfrak{R}^{(r)}$ which could destroy the given parametrization are: $x_m^j I_1^l \left\{ \begin{smallmatrix} \sin k\theta_1 \\ \cos k\theta_1 \end{smallmatrix} \right\}$ and $y_m^j I_1^l \left\{ \begin{smallmatrix} \sin k\theta_1 \\ \cos k\theta_1 \end{smallmatrix} \right\}$, or –when complex remainder is considered–, $q_m^j I_1^l \exp(ik\theta_1)$ and $p_m^j I_1^l \exp(ik\theta_1)$; with $m = 1, 2$, $k \in \mathbb{Z}$, $j = 0, 1$ and $l \in \mathbb{N}$ such that $j + 2l > r$. It can be readily seen then, that the denominators associated to these monomials are $(k \pm j\nu)\omega_1$ with $\nu \notin \mathbb{Q}$ (and k, j in the same range than before). So, small divisors do not appear here and the proposed (semi) normal form additional transformation –constructed as the limit of the successive canonical changes removing those monomials–, will be convergent in the appropriate domain. ♠

The linearized equations around these periodic orbits in the normal directions $(\boldsymbol{x}, \boldsymbol{y})$, as obtained from (1.8.3), are

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_2 & 1 & 0 \\ -\sigma_2 & 0 & 0 & 1 \\ -\sigma_1 & 0 & 0 & \sigma_2 \\ 0 & -\sigma_1 & -\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix}, \tag{1.8.6}$$

where σ_1, σ_2 are defined as:

$$\sigma_1 := \partial_1 \mathcal{Z}_r(0, I_1, 0) = dI_1 + \mathrm{O}(I_1^2), \tag{1.8.7}$$

and

$$\sigma_2 := \omega_2 + \partial_3 \mathcal{Z}_r(0, I_1, 0) = \omega_2 + fI_1 + \mathrm{O}(I_1^2), \tag{1.8.8}$$

being coefficients d and f those appearing in the quadratic part of the polynomial $\mathcal{Z}_r$ (see (1.8.2)).

From here, the characteristic multipliers associated to the normal directions of the periodic orbits are,

$$\begin{aligned} \alpha_{I_1}^{\pm} &= \mathrm{i}\sigma_2 \pm \sqrt{-\sigma_1} \\ &= \mathrm{i}(\omega_2 + fI_1) \pm \sqrt{-dI_1 + \mathrm{O}(I_1^2)} + \mathrm{O}(I_1^2), \end{aligned} \tag{1.8.9a}$$

$$\begin{aligned} \beta_{I_1}^{\pm} &= -\mathrm{i}\sigma_2 \pm \sqrt{-\sigma_1} \\ &= -\mathrm{i}(\omega_2 + fI_1) \pm \sqrt{-dI_1 + \mathrm{O}(I_1^2)} + \mathrm{O}(I_1^2). \end{aligned} \tag{1.8.9b}$$

Thus, if $|I_1|$ is small enough, the sign of the terms inside the square roots at the developments for $\alpha_{I_1}^{\pm}$ and $\beta_{I_1}^{\pm}$ in the above formulas, depends mainly on the sign of $-dI_1$. Therefore, we must distinguish two cases.

Case 1. $d > 0$, then the family of periodic orbits (1.8.5) is complex unstable for $I_1 < 0$, and (linearly) stable for $I_1 > 0$. See figure 1.7(a).

Case 2. $d < 0$. In this case the family turns out to be (linearly) stable for $I_1 < 0$ and complex unstable for $I_1 > 0$, as it can be appreciated in figure 1.7(b).

In the next section and going on with the analysis of the normal form $Z^{(r)}$, we shall see that, –under certain conditions which depend intrinsically on the Hamiltonian–, around the just studied periodic orbits there unfolds a two-parameter family of two dimensional invariant tori. Furthermore, a study of the normal behavior of such bifurcating tori is done. The global result resembles the classical Andronov-Hopf bifurcation, in the sense that unfolded *stable* objects (2D-invariant tori in our case) appear around lower dimensional *unstable* ones –here, the periodic orbits of the family–, whereas conversely, *unstable* 2D-invariant tori may unfold around *stable* periodic orbits. Whether the former or the latter phenomenon takes place, depends again upon the nature of the Hamiltonian. In the literature –see Van der Meer (1985)–, this kind of bifurcation is known as the *Hamiltonian Andronov-Hopf bifurcation.*

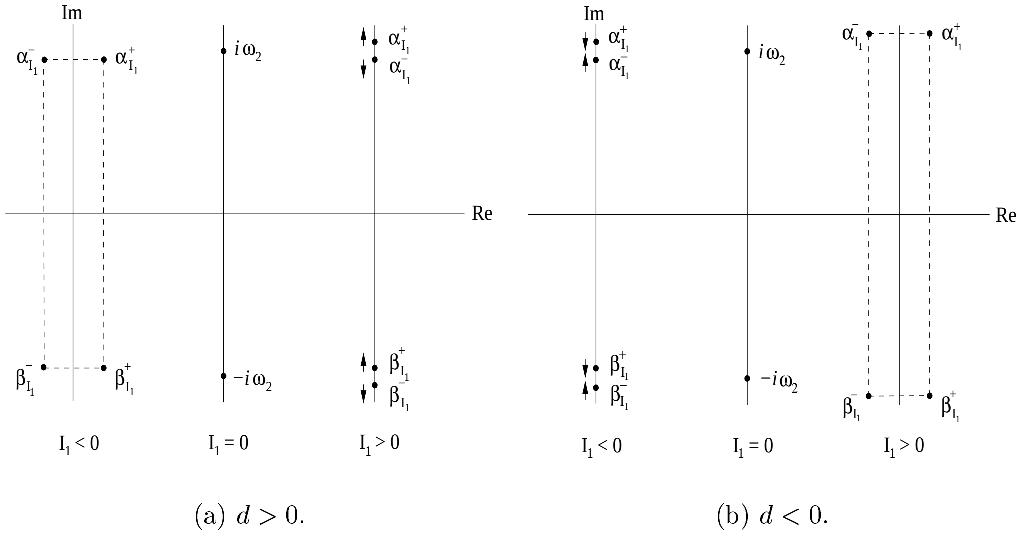


Figure 1.7: We note that when $I_1 = 0$, then $\alpha_0^- = \alpha_0^+ = i\omega_2$ and $\beta_0^- = \beta_0^+ = -i\omega_2$ (collision of characteristics exponents). Therefore, the family changes its linear character from complex-unstable to stable (when $d > 0$, fig. 1.7(a)), or vice-versa (when $d < 0$, fig.1.7(b)).

1.9 An unfolding 2D-invariant tori family

We are interested in the quasi-periodic bifurcation phenomena linked with the transition stable-complex unstable. We will describe such phenomena using the r-order normal form $Z^{(r)}$.

It turns out that the quasi-periodic solutions we are looking for are more easily seen if we first change to (canonical) polar coordinates,

$$\begin{aligned} x_1 &= \sqrt{2q}\cos\theta_2, \quad & y_1 &= -\frac{I_2}{\sqrt{2q}}\sin\theta_2 + p\sqrt{2q}\cos\theta_2, \\ x_2 &= -\sqrt{2q}\sin\theta_2, \quad & y_2 &= -\frac{I_2}{\sqrt{2q}}\cos\theta_2 - p\sqrt{2q}\sin\theta_2, \end{aligned} \tag{1.9.1}$$

with $q > 0$. This introduces a second action I_2, together with its conjugate angle θ_2 while q and p are the new normal coordinate and its conjugate momentum respectively. With these polar coordinates, the Hamiltonian $Z^{(r)}$ takes the form,

$$Z^{(r)}(\theta_1, \theta_2, q, I_1, I_2, p) = \omega_1 I_1 + \omega_2 I_2 + qp^2 + \frac{I_2^2}{4q} + \mathcal{Z}_r(q, I_1, I_2), \tag{1.9.2}$$

(following our convention, we keep the same name for the normal form expressed in the new coordinates). From its corresponding Hamiltonian equations,

$$\begin{aligned}
\dot{\theta}_1 &= \omega_1 + \partial_2 \mathcal{Z}_r(q, I_1, I_2), \\
\dot{\theta}_2 &= \omega_2 + \frac{I_2}{2q} + \partial_3 \mathcal{Z}_r(q, I_1, I_2), \\
\dot{q} &= 2qp, \\
\dot{I}_1 &= 0, \\
\dot{I}_2 &= 0, \\
\dot{p} &= -p^2 + \frac{I_2^2}{4q^2} - \partial_1 \mathcal{Z}_r(q, I_1, I_2),
\end{aligned} \tag{1.9.3}$$

it can be seen that the system generically presents a bifurcating two parameter family of quasi-periodic solutions. We formalize this result in theorem 1.27 below, but a previous general definition (borrowed, verbatim, from the book of Broer, Huitema and Sevryuk, 1996) is still needed.

Definition 1.26 (torus with parallel dynamics). *Consider a smooth vector field X on a manifold M with an invariant n-torus $\mathcal{T}$. We say that X on $\mathcal{T}$ induces parallel (or conditionally periodic, or Kronecker, or linear) motion, evolution, dynamics, or flow, if there exists a diffeomorphism from $\mathcal{T}$ to $\mathbb{T}^n$ transforming the restriction $X|_{\mathcal{T}}$ to a constant vector field $\sum_{j=1}^n \omega_j \partial/\partial x_j$ on the standard n-torus $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$ with angular coordinates $x_1, x_2, \ldots, x_n$ modulo 2π. In a more familiar notation, this vector field determines the system $\dot{x}_j = \omega_j$, $1 \leq j \leq n$, of differential equations. The numbers $\omega_1, \omega_2, \ldots, \omega_n$ are called (intrinsic) frequencies of the motion (evolution, dynamics or flow) on $\mathcal{T}$, but also of the invariant torus $\mathcal{T}$ itself.*

Theorem 1.27. *If the coefficient a in the quadratic part of $\mathcal{Z}_r$ in (1.8.2) is different from zero, there exists a real analytic function $\Upsilon : \mathfrak{U} \subset \mathbb{R}^2 \to \mathbb{R}$, $(0,0) \in \mathfrak{U}$, such that for $(J_1, J_2) \in \mathfrak{U}$, the two-dimensional tours:*

$$\begin{aligned}
\mathcal{T}_{J_1,J_2} = \{(\boldsymbol{\theta}, q, \boldsymbol{I}, p) \in \mathbb{T}^2 \times \mathbb{C} \times \mathbb{C}^2 \times \mathbb{C} :& \\
q = \Upsilon(J_1, J_2),\ I_1 = J_1,\ I_2 = 2J_2\Upsilon(J_1, J_2),\ p = 0\}&
\end{aligned} \tag{1.9.4}$$

is invariant under the flow of (1.9.3) *with parallel dynamics determined by the intrinsic frequencies:*

$$\Omega_1(J_1, J_2) = \omega_1 + \partial_2 \mathcal{Z}_r\left(\Upsilon(J_1, J_2), J_1, 2J_2\Upsilon(J_1, J_2)\right), \tag{1.9.5a}$$

$$\Omega_2(J_1, J_2) = \omega_2 + J_2 + \partial_3 \mathcal{Z}_r\left(\Upsilon(J_1, J_2), J_1, 2J_2\Upsilon(J_1, J_2)\right). \tag{1.9.5b}$$

Furthermore, let $\mathfrak{U}_* \subset \mathfrak{U}$ be the set:

$$\mathfrak{U}_* = \{(J_1, J_2) \in \mathfrak{U} : \Upsilon(J_1, J_2) > 0\},$$

then $\{\mathcal{T}_{J_1,J_2}\}_{(J_1,J_2)\in\mathfrak{U}_*}$ constitutes a two parameter family of real two-dimensional invariant tori. The intrinsic frequencies of these tori will be non-degenerated (in Kolmogorov's sense) if, in addition:

$$b - \frac{d^2}{a} \neq 0, \tag{1.9.6}$$

where the coefficients b and d involved in this last formula are again those of the polynomial $\mathcal{Z}_r$.

Proof. It is enough to look for the equilibrium points of the the system (1.9.3) restricted to the normal direction q and its conjugate momentum p. There, we fix $p = 0$, so $\dot{q} = 0$ and introduce the parameters J_1 and J_2 by $I_1 = J_1, I_2 = 2J_2 q$. Then, from the last equilibrium point condition, $\dot{p} = 0$, we derive the equation,

$$J_2^2 - \partial_1 \mathcal{Z}_r(q, J_1, 2J_2 q) = 0, \tag{1.9.7}$$

which has a trivial solution $(q, J_1, J_2) = (0, 0, 0)$. Therefore, as

$$\frac{\partial}{\partial q}\left(J_2^2 - \partial_1 \mathcal{Z}_r(q, J_1, 2J_2 q)\right)\Big|_{(q=0,J_1=0,J_2=0)} = -\partial_{1,1}^2 \mathcal{Z}_r(0,0,0) = -a,$$

and $a \neq 0$, the local implicit function theorem can be used to establish the existence of,

(i) An open set $\mathfrak{U} \subset \mathbb{R}^2$, with $(0,0) \in \mathfrak{U}$, and

(ii) an analytic function $\Upsilon : \mathfrak{U} \to \mathbb{R}$,

such that

(i') $\Upsilon(0,0) = 0$,

(ii') $J_2^2 - \partial_1 \mathcal{Z}_r\left(\Upsilon(J_1, J_2), J_1, 2J_2\Upsilon(J_1, J_2)\right) = 0$, for all $(J_1, J_2) \in \mathfrak{U}$.

Now, it is immediately checked out that the flow of (1.9.3) on the torus $\mathcal{T}_{J_1,J_2}$ is given by,

$$t \mapsto \begin{pmatrix} \Omega_1(J_1, J_2)\, t + \theta_1^0 \\ \Omega_2(J_1, J_2)\, t + \theta_2^0 \end{pmatrix} \quad \bmod 2\pi, \qquad (\theta_1^0, \theta_2^0) \in \mathbb{T}^2,\ (J_1, J_2) \in \mathfrak{U},$$

with $\Omega_1(J_1, J_2)$, $\Omega_2(J_1, J_2)$ given by (1.9.5a) and (1.9.5b). Here, J_1, J_2 determine the invariant torus of the family and θ_1^0, θ_2^0 are initial conditions for an orbit on it. This ends the proof of the first statement of the theorem. Let us point that, if we want real tori, we have to ask $q > 0$ in (1.9.1), so the "reality condition" is $\Upsilon(J_1, J_2) > 0$.

Next, expanding Υ in power series up to first order, one obtains

$$\begin{aligned} \Upsilon(J_1, J_2) = {} & -\frac{d}{a}J_1 + \left(-\frac{3d^2}{a^3}f_{3,0,0} + \frac{2d}{a^2}f_{2,1,0} - \frac{1}{a}f_{1,2,0}\right)J_1^2 \\ & + \frac{1}{a}J_2^2 + \frac{2ed}{a^2}J_1 J_2 + \mathrm{O}_3(J_1, J_2). \end{aligned} \tag{1.9.8}$$

Similarly, a low order expansion of the frequencies Ω_1, Ω_2 yields,

$$\begin{aligned} \Omega_1(J_1, J_2) = \omega_1 & + \left(b - \frac{d^2}{a}\right)J_1 + \left(-\frac{3d^3}{a^3}f_{3,0,0} + \frac{3d^2}{a^2}f_{2,1,0} - \frac{3d}{a}f_{1,2,0} + 3f_{0,3,0}\right)J_1^2 \\ & + \frac{d}{a}J_2^2 + \left(\frac{2ed^2}{a^2} - \frac{2fd}{a}\right)J_1 J_2 + \mathrm{O}_3(J_1, J_2), \end{aligned} \tag{1.9.9a}$$

$$\begin{aligned} \Omega_2(J_1, J_2) = \omega_2 & + \left(f - \frac{ed}{a}\right)J_1 + J_2 \\ & + \left(-\frac{3ed^2}{a^3}f_{3,0,0} + \frac{2ed}{a^2}f_{2,1,0} - \frac{e}{a}f_{1,2,0} + \frac{d^2}{a^2}f_{2,0,1} - \frac{d}{a}f_{1,1,1} + f_{0,2,1}\right)J_1^2 \\ & + \frac{e}{a}J_2^2 + \left(\frac{2e^2 d}{a^2} - \frac{2cd}{a}\right)J_1 J_2 + \mathrm{O}_3(J_1, J_2) \end{aligned} \tag{1.9.9b}$$

and the determinant of the Jacobian matrix, $\frac{\partial(\Omega_1,\Omega_2)}{\partial(J_1,J_2)}$, is up to zero order in J_1, J_2,

$$\det\begin{pmatrix} b-\frac{d^2}{a}+\mathrm{O}_1(J_1,J_2) & \mathrm{O}_1(J_1,J_2) \\ f-\frac{ed}{a}+\mathrm{O}_1(J_1,J_2) & 1+\mathrm{O}_1(J_1,J_2) \end{pmatrix} = b-\frac{d^2}{a}+\mathrm{O}_1(J_1,J_2).$$

Then, for $|J_1|, |J_2|$ sufficiently small and if the quantity $b-\frac{d^2}{a}$ is different from zero, the frequencies of the invariant tori will change along the family. □

Remark 1.28. From (1.9.8) it can be seen that, in order to check the reality condition $\Upsilon(J_1,J_2)>0$ it suffices, for $|J_1|, |J_2|$ small enough, to ask

$$-\frac{d}{a}J_1>0 \quad \text{and} \quad |J_2|\le|J_1|^{\alpha}, \tag{1.9.10}$$

with $\alpha>1/2$. Later on, in chapter 3, we devise another parametrization for the family of invariant tori, more suitable to control their real character (see theorem 3.1). ♠

To investigate the stability type (that is, the elliptic or hyperbolic character) of the tori in theorem 1.27, we set up the normal first variational equations around an invariant torus $\mathcal{T}_{J_1,J_2}$ of the family: i. e., we consider the system, $\dot{\boldsymbol{z}}=J_3\partial^2\mathcal{Z}^{(r)}(\Upsilon(J_1,J_2),J_1,2J_2\Upsilon(J_1,J_2))\boldsymbol{z}$, restricted to those normal directions to the invariant torus. In our case, they correspond only to the coordinate q, and its conjugate momentum p. So, writing these equations explicitly,

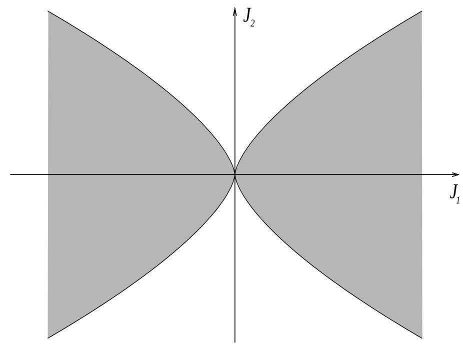


Figure 1.8: $|J_2|\le|J_1|^{2/3}$.

$$\begin{aligned} \dot{X} &= 2\Upsilon(J_1,J_2)\,Y, \\ \dot{Y} &= -\left(\frac{2J_2^2}{\Upsilon(J_1,J_2)}+\partial^2_{1,1}\,\mathcal{Z}_r|_{\mathcal{T}_{J_1,J_2}}\right)X, \end{aligned} \tag{1.9.11}$$

it is easy to check that $\partial^2_{1,1}\,\mathcal{Z}_r|_{\mathcal{T}_{J_1,J_2}}=a+\dots$ and, with the corresponding expansion for $\Upsilon(J_1,J_2)=-\frac{d}{a}J_1+\dots$ in (1.9.8) we can compute the characteristic exponents, $\varrho^{\pm}_{J_1,J_2}$, of the torus

$$\begin{aligned} \varrho^{\pm}_{J_1,J_2} &= \pm\sqrt{-4J_2^2-2\Upsilon(J_1,J_2)\partial^2_{1,1}\mathcal{Z}_r(\Upsilon(J_1,J_2),J_1,2J_2\Upsilon(J_1,J_2))} \\ &= \pm\sqrt{2dJ_1-6J_2^2+\left(2f_{1,2,0}-\frac{6d^2}{a^2}f_{3,0,0}\right)J_1^2-\frac{2ed}{a}J_1J_2+\mathrm{O}_3(J_1,J_2)}. \end{aligned} \tag{1.9.12}$$

This last relation, together with the formulas (1.8.9a) and (1.8.9b) for the characteristic exponents of the periodic orbits, leads to the following proposition.

Proposition 1.29. *Under the assumptions of theorem 1.27 –including the reality condition $\Upsilon>0$–, the type of the bifurcation is determined by the sign of the coefficient a in* (1.8.2). *More precisely,* *Case 1.* *$a > 0$ then, we say that the bifurcation is "direct": there appear elliptic tori around complex-unstable periodic orbits; and if*

Case 2. *$a < 0$ the bifurcation is "inverse": hyperbolic invariant tori unfold around stable periodic orbits. In this case, the family contains also parabolic and elliptic tori.*

Proof. Indeed, when $a > 0$, it must be –in view of the (local) reality conditions–: $dJ_1 < 0$. Therefore the expression (1.9.12) for the characteristic exponents $\varrho^{\pm}_{J_1,J_2}$ shows that they are purely imaginary (as always, provided J_1, J_2 are small enough in absolute value). Hence, the invariant tori are elliptic. On the other hand, d and $J_1(= I_1)$ should have opposite signs for their product to be negative; but then, the corresponding periodic orbit is unstable (see figure 1.7). Similar arguments apply when $a < 0$. □

Remark 1.30. (A note on the parabolic tori). As stated in the last proposition, when $a < 0$, also real parabolic tori appear. Let us explain this point in more detail. If we define $\breve{\varrho}(J_1, J_2)$ as the stuff inside the square root in (1.9.12), i. e.:

$$\begin{aligned}\breve{\varrho}(J_1, J_2) :=& -4J_2^2 - 2\Upsilon(J_1, J_2)\partial^2_{1,1}(\Upsilon(J_1, J_2), J_1, 2J_2\Upsilon(J_1, J_2)) \\ =& \ 2dJ_1 - 6J_2^2 + \left(2f_{1,2,0} - \frac{6d^2}{a^2}f_{3,0,0}\right)J_1^2 - \frac{2ed}{a}J_1J_2 + \mathrm{O}_3(J_1, J_2)\end{aligned}$$

then, provided $d \neq 0$, application of the implicit function theorem at $(J_1, J_2) = (0, 0)$ shows the existence, in the space of parameters (J_1, J_2) of a curve $J_1(J_2) = \frac{3}{d}J_2^2 + \mathrm{O}_3(J_2)$, giving rise to a one-parameter family of real parabolic tori. The same can be done also for $a > 0$; but then (by substitution in (1.9.8)) it can be seen that: $\Upsilon(J_1(J_2), J_2) = -\frac{2}{a}J_2^2 + \mathrm{O}_3(J_2)$ which will take (at least for J_2 sufficiently small) negative values leading thus to complex tori. See similar comment in section 3.2.2 of chapter 3. ♠

The solutions winding the invariant torus (1.9.4), can be put back in rectangular coordinates using the change (1.9.1), and in this way, the corresponding family of solutions for the system (1.8.3),

$$\begin{aligned} \theta_1 &= \Omega_1 t + \theta_1^0, & I_1 &= J_1, \\ x_1 &= \sqrt{2\Upsilon}\cos(\Omega_2 t + \theta_2^0), & y_1 &= -\sqrt{2\Upsilon}\,J_2\sin(\Omega_2 t + \theta_2^0), \\ x_2 &= -\sqrt{2\Upsilon}\sin(\Omega_2 t + \theta_2^0), & y_2 &= -\sqrt{2\Upsilon}\,J_2\cos(\Omega_2 t + \theta_2^0), \end{aligned} \tag{1.9.13}$$

where, to abbreviate, it is noted $\Upsilon = \Upsilon(J_1, J_2)$ and $\Omega_i = \Omega_i(J_1, J_2)$, for $i = 1, 2$.

Solutions (1.9.13) suggest a bifurcation pattern of the type plotted at figure 1.9, where the typical Hopf-like unfolding as described in proposition 1.29 can be appreciated. In particular, as shown there, if we cut the reduced phase space, (x_1, x_2, I_1), by a plane $I_1 = J_1 =$ *constant*, we obtain an equilibrium point (corresponding to the periodic orbit), "surrounded" by a one-parameter family of invariant circles, which are stable curves when the equilibrium point is unstable and vice-versa.

1.9.1 A note on the low order ($r = 2$) normal form

Now we consider the normal form (1.9.2) with $r = 2$ (fourth order Hamiltonian). It is of interest to investigate the dynamics of such Hamiltonian by itself (see Heggie, 1985). We

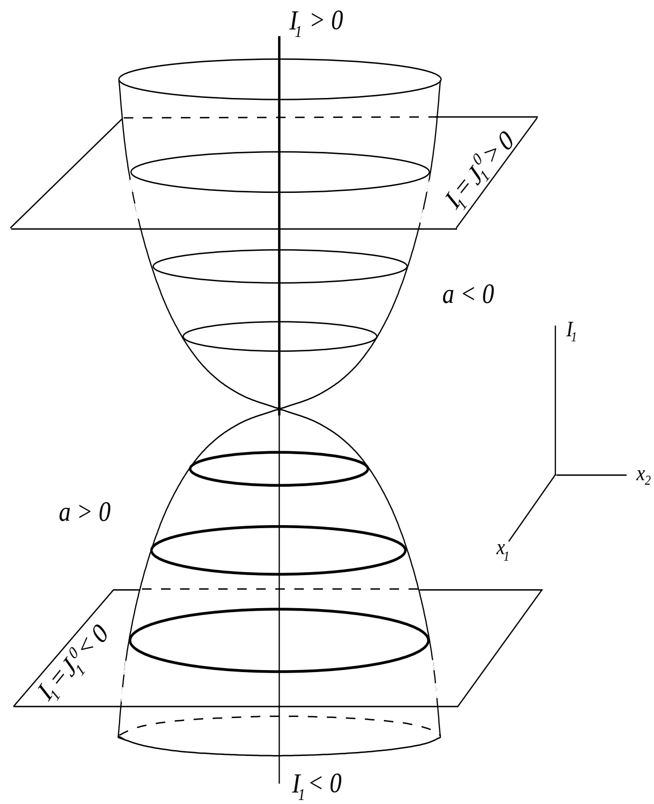


Figure 1.9: Bifurcation of quasiperiodic solutions linked to the transition stable-complex unstable (J_2 fixed). In the figure ticker and narrower traces represent stable and unstable objects respectively (periodic orbits and invariant tori). Remark: though they have been glued together, only the upper or the lower part of the diagram may take place, since whether the direct or the inverse bifurcation actually occurs, depends upon the coefficient a, and it has a fixed value for a given Hamiltonian.

begin introducing a more suitable coordinates by means of the canonical change,

$$Q = \sqrt{2q}, \qquad P = p\sqrt{2q}, \tag{1.9.14}$$

In these new coordinates, the Hamiltonian $Z^{(2)}$ will take the form,

$$Z^{(2)}(Q, P, I_1, I_2) = H_0(I_1, I_2) + H_1(Q, P, I_1, I_2), \tag{1.9.15}$$

with,

$$\begin{aligned} H_0(I_1, I_2) &= \omega_1 I_1 + \omega_2 I_2 + \frac{1}{2} b I_1^2 + \frac{1}{2} c I_2^2 + f I_1 I_2, \\ H_1(Q, P, I_1, I_2) &= \frac{1}{2} P^2 + \frac{1}{2}\left(\frac{I_2^2}{Q^2} - AQ^2\right) + BQ^4, \end{aligned} \tag{1.9.16}$$

with $A = -dI_1 - eI_2$, and $B = a/8$. The actions I_1, I_2 and also the function H_1 are constants of the movement, so $H_1(Q, P, I_1, I_2)$ is actually a one-degree of freedom Hamiltonian (and hence integrable) of type $H(Q, P) = \frac{1}{2}P^2 + V(Q)$, where the potential energy $V(Q)$ is, by identification with the second equation in (1.9.16),

$$V(Q) = \frac{1}{2}\left(\frac{I_2^2}{Q^2} - AQ^2\right) + BQ^4.$$

The flow of this Hamiltonian gives the movement in the phase plane (Q, P) of the normal directions. In particular, equilibrium points and periodic orbits of this reduced Hamiltonian will correspond to two and three-dimensional invariant tori of the complete (in the present context) Hamiltonian (1.9.16) (see next section).

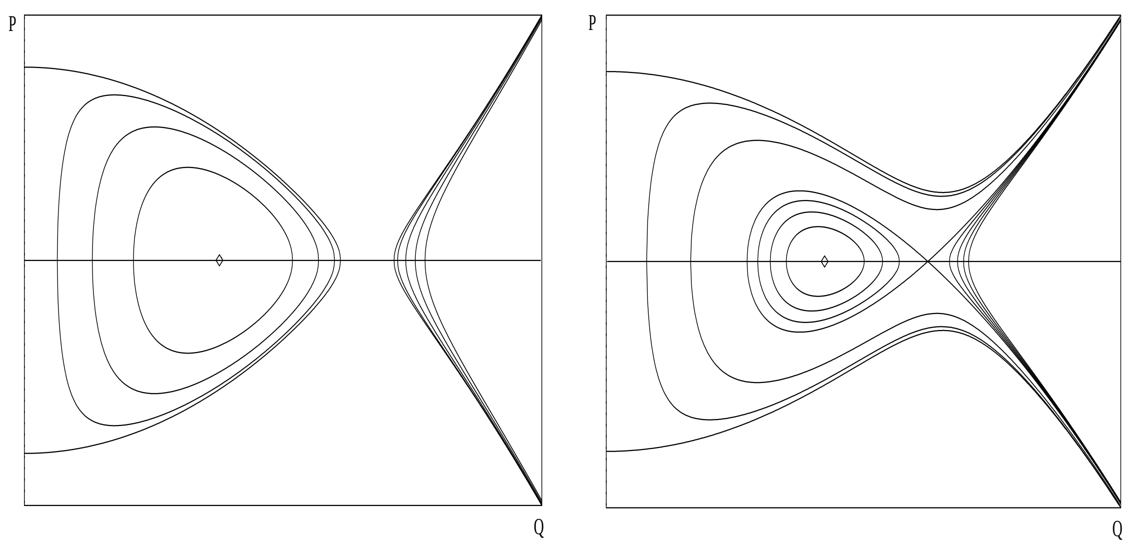
Several different behaviors are appreciated depending upon the coefficients A and B and on the energy $h = H(Q,P)$. If we suppose that the sign of the coefficient $A = -dI_1 - eI_2$, appearing in the potential $V(Q)$, is given mainly by the sign of the product $-dI_1$ (assume, for example $|I_2| < \left|\frac{d}{e}\right| |I_1|$), then the results –concerning two-dimensional invariant tori–, presented here are easily linked with those obtained in our previous study and also with the ones to appear in the forthcoming section.

More precisely, and as before, we must distinguish two cases which correspond to the inverse and direct bifurcation respectively.

Case 1. For $A < 0$ and $B < 0$. Then for a given $0 < h < -\frac{A^2}{16B}$ a 2D elliptic torus surrounded by 3D tori appear. See figure 1.10(a). When the energy is increased and $-\frac{A^2}{16B} < h < -\frac{A^2}{12B}$, then a hyperbolic torus unfolds and the 2D torus with its accompanying family of 3D tori are contained inside the loop formed by the connecting invariant manifolds of the hyperbolic 2D torus as can be appreciated in figure 1.10(b).

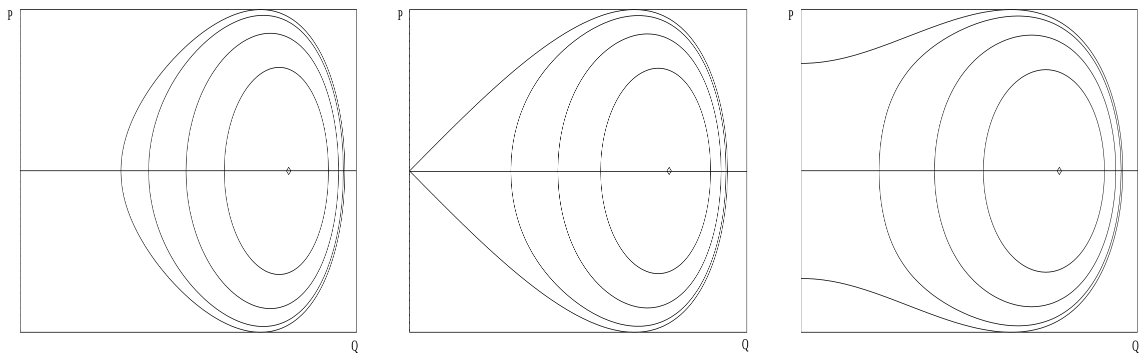
Case 2. When $A > 0$ and $B > 0$ and for a given h, an elliptic 2D invariant torus surrounded by 3D tori appears (figure 1.11). It is worth noting –figure 1.11(b)– that, for $h = 0$ the separatrix curve corresponding to the action $I_2 = 0$ merges from the origin of the (Q,P) phase plane. It corresponds to the (matching) stable and unstable invariant manifolds of the hyperbolic periodic orbit with the same value of I_1.

As the reduced Hamiltonian we are dealing with is a one-degree of freedom Hamiltonian, and the sum of the kinetic term plus a potential $V(Q)$, its phase portrait is straightforward constructed from the shape of the potential function. This is just shown, for the first case above and for both $0 < h < -\frac{A^2}{16B}$ and $-\frac{A^2}{16B} < h < -\frac{A^2}{12B}$ at figures 1.12 and 1.13 respectively. The reader can verify the second case and, in the same way, realize



(a) $0 < h < -\frac{A^2}{16B}$.   (b) $-\frac{A^2}{16B} < h < -\frac{A^2}{12B}$.

Figure 1.10: $A < 0$ and $B < 0$. Different invariant curves for several values of the action I_2 are plotted given a fixed value of the energy, h, in the specified intervals. In particular, closed trajectories correspond to 3D invariant tori around the elliptic 2D torus (an equilibrium point marked with a $\diamond$). Also, escape phase trajectories have been drawn. In (b) an hyperbolic equilibrium point (hyperbolic 2D invariant tori in the complete phase space) together with its invariant manifolds appear. In both figures, the outermost curve corresponds to $I_2 = 0$.



(a) $-\frac{A^2}{16B} < h < 0$. (b) $h = 0$. (c) $h > 0$.

Figure 1.11: $A > 0$ and $B > 0$. As in the previous plot the outermost curve corresponds to $I_2 = 0$. As pointed in the text, for $h = 0$, –figure (b)–, this curve identifies with the stable and unstable invariant manifolds of the elliptic periodic orbit for the same value of I_1. Here, no hyperbolic invariant tori appear. This situation corresponds to the direct bifurcation of the previous section.

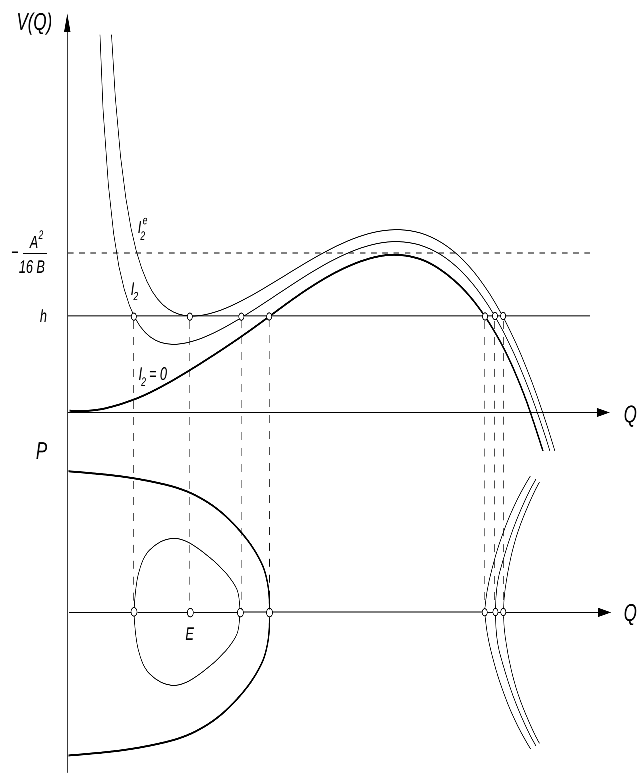


Figure 1.12: ($A < 0$, $B < 0$ and $0 < h < \frac{-A^2}{16B}$). Construction of the phase portrait of figure 1.10(a) from the potential curves $V(Q)$ corresponding to three values of the actions $I_2 = 0$ (thicker line) and $0 < I_2 < I_2^e$, where I_2^e is the action of the elliptic torus –represented by the point marked with E in the phase portrait below–. It can be appreciated that, for the value of the energy, $h < -\frac{A^2}{16B}$, and the action $0 < I_2 < I_2^e$ pointed at the upper part of the figure, the flow is whether confined in a 3D-torus (the closed curve surrounding the elliptic point E), or escape to infinity following one of the escape trajectories drawn in the figure.

that, when $A < 0$ and $B > 0$, we have a two-dimensional elliptic torus surrounded by a family of 3D-invariant tori. This agrees with the appearance of maximal and lower dimensional elliptic invariant tori around stable periodic orbits. On the other hand, when $A > 0$ and $B < 0$, there are no quasi-periodic solutions at all (for $Q > 0$, the potential $V(Q)$ decreases monotonically and goes to the infinity as one approaches rightwards to zero). For a more complete account of the analysis outlined here, see the reference of Heggie quoted at the beginning of the section.

Remark 1.31. Of course, the dynamics described in the last two sections corresponds – as it has been already pointed–, to the dynamics of the truncated normal form, that is with $\mathcal{Z}_r(q, I_1, I_2)$ as a polynomial of finite (arbitrarily high) degree r. Actually, after the normal form reduction process, the transformed Hamiltonian consists of this truncated normal form plus the remainder –see (1.7.68)–. Bounds on this remainder, $\mathfrak{R}^{(r)}$, and the persistence of the bifurcated 2D-tori (in Cantor sets) for the complete (non-integrable) Hamiltonian will be the subject of the next two chapters. ♠

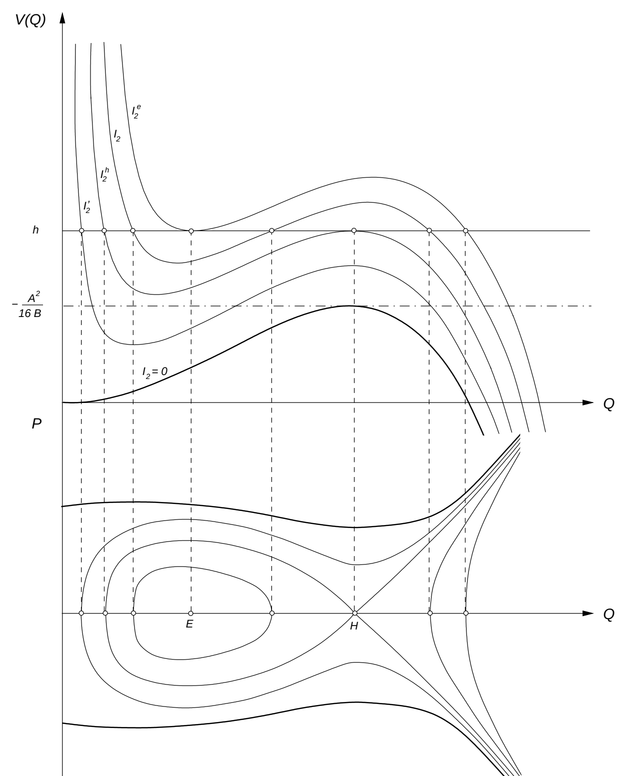


Figure 1.13: ($A < 0$, $B < 0$ and $-A^2/(16B) < h < \frac{-A^2}{12B}$). Detailed phase portrait of figure 1.10(b) from the potential curves $V(Q)$. As in the preceding figure, the potential is represented for several values of the actions: $I_2 = 0$, $0 < I_2' < I_2^h < I_2 < I_2^e$, an a fixed energy level, h, with $-\frac{A^2}{16B} < h < -\frac{A^2}{12B}$. Again, an elliptic two-dimensional invariant torus (the elliptic point marked with E in the phase plane), and 3D-tori (closed invariant curves around) appear. However, additionally, an hyperbolic 2D-torus in the whole phase space rises. It is represented in the (Q, P)-plane by the hyperbolic point H, together with its stable and unstable manifolds. Escape curves are drawn on the right of the invariant manifolds.

Chapter 2

Quantitative estimates on the normal form

2.1 Overview of the chapter

Suppose that we transform a function $f = \sum_{l \geq 1} f_l \in \mathfrak{E}$ through the map T_G in the definition 1.7 on page 21 to obtain $T_G f = \sum_{k \geq 1} F_k$ and we are interested in bounding –using some suitable norm to be defined later–, the sum $\sum_{k \geq r} F_k$. Since F_s, $s = 1, 2, \ldots,$ are obtained as a combination of Poisson brackets of $f_1, \ldots, f_k; G_3, \ldots, G_{s+1}$ of type,

$$\begin{aligned} F_1 &= f_1, \\ F_2 &= f_2 + \{f_1, G_3\}, \\ F_3 &= f_3 + \{f_2, G_3\} + \{f_1, G_4\} + \frac{1}{2}\{\{f_1, G_3\}, G_3\}, \\ &\vdots \end{aligned}$$

Then, *assuming* certain bounds for the components $G_j, j \geq 3$ of the generating function $G = G_3 + G_4 + \cdots + G_r$ (see (2.6.30) in proposition 2.19) one may attempt –with the help of the more specialized recursive formulas of the Giorgilli-Galgani algorithm–, to obtain estimates for the remainder as the sum $\mathfrak{R}^{(r)} = F_{r+1} + F_{r+2} + \ldots$. This is done in section 2.6, using proposition 2.16 stated in section 2.5. This proposition is based on lemma 2.13, completed with lemma A.12 of appendix A. Particularly, we mention that is in lemma 2.13 where the assumptions on the size of G_j, $3 \leq j \leq r$ are (through the bounds (2.5.11)) introduced. Next, in the section 2.6.1, we optimize the derived estimates with respect to the degree r up to which the normal form is carried out. In this way, an optimal r_{opt} order is obtained as a function of the distance to the critical periodic orbit. The results concerning optimization of the order are formalized in proposition 2.19. Finally, in the last section of the chapter, section 2.7, we let $f = H$, where H will be the complexified Hamiltonian (1.6.2). Then it is checked that the early assumed bounds for the terms of the generated function $G = G_3 + \cdots + G_r$ are fulfilled when these are the solutions of the homological equations (1.7.60). Hence, the optimal normalization order apply, in particular, to the complex Hamiltonian (1.6.2) and the remainder of the reduced Hamiltonian satisfy the same bounds of the proposition. Theorem 2.29 summarizes these results and also, gives an order-independent estimate for the size of the normal form, which will be useful in the forthcoming chapter 3.

Though, in the paragraph above, we have, essentially, outlined the plot of the present chapter. However, we have not mentioned that, besides a preliminary section, with some remarks and notation (see below), there are two sections (2.3, 2.4) devoted to derive bounds for the solutions of the homological equations, more precisely, to bound the solutions G_j, Z_j of the homological equations (1.7.60) in terms of the norm of their corresponding r. h. s., F_j. The main result (used later in section 2.7) is lemma 2.7, given at the end of section 2.4.

2.2 Preliminaries

We want to point here that, when analogous normal form computations are applied to a semisimple elliptic equilibrium point of a Hamiltonian (see example B.23), homological equations lead to a linear algebraic diagonal system in the coefficients $g_{\boldsymbol{l},\boldsymbol{m}}$ (of the polynomial G_s), whose solutions are $g_{\boldsymbol{l},\boldsymbol{m}} = \frac{F_{\boldsymbol{l},\boldsymbol{m}}}{\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{l}-\boldsymbol{m}\rangle}$, with $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = s$, $\boldsymbol{l} \neq \boldsymbol{m}$, and $F_{\boldsymbol{l},\boldsymbol{m}}$ the coefficients of F_s, (the right hand side of the homological equations). It is usual to introduce some conditions on the frequencies $\boldsymbol{\omega}^* = (\omega_1, \ldots, \omega_n)$. For example, as in Giorgilli et al. (1989), assume that $|\langle \boldsymbol{\omega}, \boldsymbol{\nu}\rangle| \geq \alpha_r$ for those $\boldsymbol{\nu} \notin \mathfrak{R}$ (being $\mathfrak{R}$ the resonance module considered) and such that, $|\boldsymbol{\nu}|_1 \leq r$. In $\mathfrak{A}_s$ –the space of homogeneous polynomials of degree s as defined in example B.23–, one introduces the norm,

$$\|f\| = \sum_{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = s} |f_{\boldsymbol{l},\boldsymbol{m}}|,$$

and then the term G_s is easily bounded by $\|G_s\| \leq \frac{1}{\alpha_r}\|F_s\|$.

When the homological equations are not diagonal, as actually happens, the computations are more involved. Moreover, we shall work with functions which are no longer polynomials (see below). Then, it is worth introducing before the appropriate norms and some notation.

Notation

Let $\mathfrak{E}$ denote now the space of the analytic functions $f(\theta_1, I_1, \boldsymbol{q}, \boldsymbol{p})$, and 2π-periodic with respect θ_1, defined on $\mathcal{D}(\rho, R)$, with some $\rho < \rho^*$ and $R < R^*$ –see (1.6.3)–. Later on though, this domain will be more precisely defined (section 2.5.1). Therefore a function $f \in \mathfrak{E}$ admits an expansion in Taylor series of type (1.7.1), again,

$$f = \sum_{(l,\boldsymbol{m},\boldsymbol{n}) \in \mathbb{Z}_+ \times \mathbb{Z}_+^2 \times \mathbb{Z}_+^2} f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}}, \tag{2.2.1}$$

(with $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$). In their turn, the coefficients $f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1)$, expand in Fourier series,

$$f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1) = \sum_{k \in Z} f_{k,l,\boldsymbol{m},\boldsymbol{n}} \exp(\mathrm{i}k\theta_1). \tag{2.2.2}$$

We use the developments (2.2.1) and (2.2.2) to introduce in $\mathfrak{E}$ the following norms,

$$|f_{l,\boldsymbol{m},\boldsymbol{n}}|_\rho = \sum_{k\in\mathbb{Z}} |f_{k,l,\boldsymbol{m},\boldsymbol{n}}| \exp(|k|\rho), \tag{2.2.3}$$

$$|f|_{\rho,R} = \sum_{l\in\mathbb{Z}_+} \sum_{\boldsymbol{m}\in\mathbb{Z}_+^2} \sum_{\boldsymbol{n}\in\mathbb{Z}_+^2} |f_{l,\boldsymbol{m},\boldsymbol{n}}|_\rho \, R^{2l+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1}, \tag{2.2.4}$$

(with the definition of degree given by (1.7.3), section 1.7 of the previous chapter, i. e. counting twice the contribution of the degree in I). Some basic properties of these norms are given in the appendix A. If the sums defining these norms are convergent, then

$$\sup_{|\mathrm{Im}\,\theta_1|\le\rho} |f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1)| \le |f_{l,\boldsymbol{m},\boldsymbol{n}}|_\rho, \qquad \sup_{\mathcal{D}(\rho,R)} |f| \le |f|_{\rho,R},$$

i. e., they are bounds for the supremum norms of $f_{l,\boldsymbol{m},\boldsymbol{n}}(\theta_1)$, on the complex strip of width $\rho > 0$, and for f on $\mathcal{D}(\rho, R)$ (see (1.6.3)). The use of these norms will simplify many of the bounds, specially those of the small divisors (see lemma A.1 of appendix A). Alternatively one could consider the supremum norm and use the bounds in Rüssmann (1975).

We have already used the *absolute norm* of a vector $\boldsymbol{x} \in \mathbb{R}^n$ $(\mathbb{C}^n)$ $|\boldsymbol{x}|_1 := \sum_{i=1}^n |x_i|$; $\boldsymbol{x} \in \mathbb{R}^n$ (C^n). Given a $n \times n$ real (complex) matrix $A = (a_{i,j})_{1\le i,j\le n}$, the sum

$$|A|_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{i,j}| \tag{2.2.5}$$

defines a *compatible* matrix norm, since it can be shown that $|A\boldsymbol{x}|_1 \le |A|_1 |\boldsymbol{x}|_1$ (see Fröberg, 1973, chapter 3). We shall use these norms to bound the solutions of the linear system (1.7.19) as the first step to get estimates of the generating function.

Moreover, and for shortness, we shall denote by $\mathfrak{E}^s_{\mathcal{M}}$ the space of homogeneous polynomials of even degree s, constituted only by $\mathcal{M}$-type monomials, i. e.,

$$\mathfrak{E}^s_{\mathcal{M}} := \oplus_{l+M=s/2} \mathfrak{E}_{0,l,M,M}, \tag{2.2.6}$$

(see section 1.7.1).

2.3 Bounds for the solutions of the homological equations

Let us return to the matrix formulation of the homological equations given in section 1.7.1 of chapter 1 by the $(M+1)\times(N+1)$ linear system (1.7.21), where E_ν, $\nu = 1, 2, \ldots, N+1$ were the $(N+1)\times(N+1)$ diagonal matrices diag$[\nu, \ldots, \nu]$, $D_N = \Omega I_{N+1} - P_N$, with P_N the $(N+1)\times(N+1)$ nilpotent matrix (1.7.22), I_{N+1} the identity matrix of the same order and $\Omega = \Omega_{k,M,N}$, defined as $\Omega_{k,M,N} = i\omega_1 k + i\omega_2(M-N)$.

We can split each term of degree s of the generating function G, G_s, as the sum $G_s = G_s^{(1)} + G_s^{(2)}$, with $G_s^{(2)} \in \mathfrak{E}^s_{\mathcal{M}}$ –i. e., $G_s^{(2)}$ contains only $\mathcal{M}$-type monomials (see definition 1.11)–, and $G_s^{(1)} \in \mathfrak{E}_s \setminus \mathfrak{E}^s_{\mathcal{M}}$ (note that $G_s^{(2)} = 0$ for s even). This last one gives rise to linear systems with $\Omega \neq 0$.

Our aim in the present section, is to derive estimates on the norm of $G_s^{(1)}$ –so $\Omega \neq 0$ will be assumed throughout–, whilst the search for bounds on $G_s^{(2)}$ will be relegated to section 2.4.

The first step is, then, to get bounds for the solutions of (1.7.21). This will provide bounds on the coefficients of the expansion of $G_s^{(1)}$. We are going to see that those solutions can be formally expressed as (if necessary, see section 1.7.1, page 25, to review the notation)

$$\begin{aligned} \boldsymbol{g}_M &= D_N^{-1} \boldsymbol{f}_M, \\ \boldsymbol{g}_{M-\nu} &= \sum_{j=0}^{\nu-1} (-1)^{\nu+j} (M-\nu+1)_{\nu-j} D_N^{-\nu+j-1} \boldsymbol{f}_{M-j} + D_N^{-1} \boldsymbol{f}_{M-\nu}, \end{aligned} \tag{2.3.1}$$

$\nu = 1, \dots, M$ and where the Pochhammer symbols

$$(\alpha)_\nu := \alpha(\alpha+1)\cdots(\alpha+\nu-1), \quad (\alpha)_0 := 1,$$

are used. Indeed, since the solutions of (1.7.21) can be first obtained in the following recursive form,

$$\begin{aligned} \boldsymbol{g}_M &= D_N^{-1} \boldsymbol{f}_M, \\ \boldsymbol{g}_{M-\nu} &= -(M-\nu+1) D_N^{-1} \boldsymbol{g}_{M-\nu+1} + D_N^{-1} \boldsymbol{f}_{M-\nu}, \end{aligned} \tag{2.3.2}$$

so taking $\nu = 1$, $\boldsymbol{g}_{M-1}$ turns out to be $\boldsymbol{g}_{M-1} = -M D_N^{-1} \boldsymbol{g}_M + D_N^{-1} \boldsymbol{f}_{M-1}$ and substituting the first of (2.3.1),

$$\boldsymbol{g}_{M-1} = -M D_N^{-2} \boldsymbol{f}_M + D_N^{-1} \boldsymbol{f}_{M-1},$$

which coincides with the second of (2.3.1) for $\nu = 1$. Now, suppose this same relation works also for ν, with $1 < \nu \leq M$. Then for $\nu + 1$:

$$\begin{aligned} \boldsymbol{g}_{M-\nu-1} &= D_N^{-1} \boldsymbol{f}_{M-\nu-1} - (M-\nu) D_N^{-1} \boldsymbol{g}_{M-\nu} \\ &= D_N^{-1} \boldsymbol{f}_{M-\nu-1} - (M-\nu) D_N^{-2} \boldsymbol{f}_{M-\nu} \\ &+ \sum_{j=0}^{\nu-1} (-1)^{\nu+j+1} (M-\nu)(M-\nu+1)_{\nu-j} D_N^{-\nu+j-2} \boldsymbol{f}_{M-j}. \end{aligned}$$

Using that,

$$\begin{aligned} (M-\nu)(M-\nu+1)_{\nu-j} &= (M-\nu)(M-\nu+1)\cdots(M-j) \\ &= (M-\nu)_{\nu-j+1}, \end{aligned}$$

for $j = 1, 2, \dots, \nu+1$, the last formula for $\boldsymbol{g}_{M-\nu-1}$ can be arranged to,

$$\boldsymbol{g}_{M-\nu-1} = D_N^{-1} \boldsymbol{f}_{M-\nu-1} + \sum_{j=0}^{\nu} (-1)^{\nu+1+j} (M-\nu)_{\nu+1-j}\, D_N^{-\nu-1+j-1} \boldsymbol{f}_{M-j}.$$

This ends the induction and hence (2.3.1) is fully justified. Moreover, if as in section 1.7.1, we denote the matrix of (1.7.21) by Λ, it is immediately seen in view of the just obtained solutions for $\boldsymbol{g}$, that its inverse, Λ^{-1}, can be written blockwise as

$$\Lambda^{-1}=\begin{pmatrix} \binom{M}{0}\mathcal{D}_1 & & & & & & \\ \binom{M}{1}\mathcal{D}_2 & \binom{M-1}{0}\mathcal{D}_1 & & & & & \\ \binom{M}{2}\mathcal{D}_3 & \binom{M-1}{1}\mathcal{D}_2 & \binom{M-2}{0}\mathcal{D}_1 & & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ \binom{M}{\nu}\mathcal{D}_{\nu+1} & \binom{M-1}{\nu-1}\mathcal{D}_{\nu} & \binom{M-2}{\nu-2}\mathcal{D}_{\nu-1} & \cdots & \binom{M-\nu}{0}\mathcal{D}_1 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \\ \binom{M}{M}\mathcal{D}_{M+1} & \binom{M-1}{M-1}\mathcal{D}_{M} & \binom{M-2}{M-2}\mathcal{D}_{M-1} & \cdots & \binom{M-\nu}{M-\nu}\mathcal{D}_{M-\nu+1} & \cdots & \binom{0}{0}\mathcal{D}_1 \end{pmatrix}, \tag{2.3.3}$$

where, for convenience, we have introduced $\mathcal{D}_\nu = (-1)^{\nu+1}(\nu-1)!\,D_N^{-\nu}$, and used that $(M)_\nu = \nu!\binom{M+\nu-1}{\nu}$. To determine the powers $D_N^{-\nu}$ ($\nu = 1, 2, \ldots, M+1$), of the inverse matrix D_N^{-1}, we proceed from the definition of D_N, so

$$D_N^{-\nu} = (\Omega \cdot I_{N+1} - P_N)^{-\nu},$$

or, equivalently, $D_N^{-\nu} = \frac{1}{\Omega^\nu}\left(I_{N+1} - \frac{1}{\Omega}P_N\right)^{-\nu}$ and after binomial expansion

$$\begin{aligned} D_N^{-\nu} &= \sum_{j=0}^{N} \frac{(-1)^j}{\Omega^{\nu+j}}\binom{-\nu}{j}P_N^j \\ &= \tfrac{1}{\Omega^\nu}I_{N+1} - \tfrac{1}{\Omega^{\nu+1}}\binom{-\nu}{1}P_N + \tfrac{1}{\Omega^{\nu+2}}\binom{-\nu}{2}P_N^2 + \cdots + \tfrac{(-1)^N}{\Omega^{\nu+N}}\binom{-\nu}{N}P_N^N, \end{aligned}$$

but $\binom{-\nu}{1} = -\frac{\nu}{1!}$, $\binom{-\nu}{2} = \frac{\nu(\nu+1)}{2!}$,... and, in general, for $\nu = 1, \ldots, M+1$,

$$\binom{-\nu}{N} = (-1)^N \frac{\nu(\nu+1)\cdots(\nu+N-1)}{N!} = \frac{(-1)^N}{N!}(\nu)_N. \tag{2.3.4}$$

Direct computation of the powers P_N^j yields

$$P_N^j = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ j!\binom{N}{j} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & j!\binom{N-1}{j} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & j!\binom{N-2}{j} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & j! & 0 & \cdots & 0 \end{pmatrix}.$$

Here, the coefficient $j!\binom{N}{j}$ which produces the first row different from zero is placed at the $(j+1)$–th row, with j ranging from $j = 0$ (given the $(N+1)\times(N+1)$ identity matrix) to $j = N$. Thus, defining

$$a_j(\nu, \Omega) := \tfrac{(\nu)_j}{\Omega^{\nu+j}}, \qquad j = 0, 1, \ldots, N;$$

and by substitution in our previous expansion of $D_N^{-\nu}$, one obtains an explicit expression for these matrices

$$D_N^{-\nu} = \begin{pmatrix} \binom{N}{0}a_0 & & & & & & \\ \binom{N}{1}a_1 & \binom{N-1}{0}a_0 & & & & & \\ \binom{N}{2}a_2 & \binom{N-1}{1}a_1 & \binom{N-2}{0}a_0 & & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ \binom{N}{j}a_j & \binom{N-1}{j-1}a_{j-1} & \binom{N-2}{j-2}a_{j-2} & \cdots & \binom{N-j}{0}a_0 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \\ \binom{N}{N}a_N & \binom{N-1}{N-1}a_{N-1} & \binom{N-2}{N-2}a_{N-2} & \cdots & \binom{N-j}{N-j}a_{N-j} & \cdots & \binom{0}{0}a_0 \end{pmatrix}. \tag{2.3.5}$$

Remark 2.1. Before continuing, we stress here that, actually, all the entities involved in the linear system (1.7.19) do depend, through Ω, on $k \in \mathbb{Z}$, and M, N non-negative integers with $M + N + 2l = s$ with $l = 0, 1, 2, \ldots, \lfloor s/2 \rfloor$ fixed. From the expression of the blocks D_N^{-1} just found, where powers of Ω are present, it follows, in view of (2.3.3), that the same applies to the inverse matrix Λ^{-1}. To avoid, as far as possible, the use of an excessive number of indices, we shall not write them out explicitly, but ask the reader to keep these dependences in mind. ♠

The next lemma furnishes an estimate on the norm of Λ^{-1}.

Lemma 2.2. *For the corresponding k, M and N (see last remark),*

$$\left|\Lambda^{-1}\right|_1 \leq \left(1 + \frac{1}{|\Omega|}\right)^{M+N} \frac{(M+N)!}{|\Omega|}. \tag{2.3.6}$$

Proof. Directly from the structure of the matrix (2.3.3), we have that

$$\left|\Lambda^{-1}\right|_1 = \sum_{\nu=1}^{M+1} \binom{M}{\nu-1} |\mathcal{D}_\nu|_1,$$

and with the definition of $\mathcal{D}_\nu$ given before,

$$\left|\Lambda^{-1}\right|_1 = \sum_{\nu=1}^{M+1} (\nu-1)! \binom{M}{\nu-1} \left|D_N^{-\nu}\right|_1.$$

Then from (2.3.5), it is clear that the norm of $D_N^{-\nu}$ equals to,

$$\left|D_N^{-\nu}\right|_1 = \sum_{j=0}^{N} \binom{N}{j} \frac{(\nu)_j}{|\Omega|^{\nu+j}}.$$

This allows to bound $|\Lambda^{-1}|_1$ in accordance with

$$\begin{aligned}
\left|\Lambda^{-1}\right|_1 &= \sum_{\nu=1}^{M+1}(\nu-1)!\binom{M}{\nu-1}\sum_{j=0}^{N}\binom{N}{j}\frac{(\nu)_j}{|\Omega|^{\nu+j}} \\
&\underset{((\nu)_j\le(\nu)_N)}{\le} \sum_{\nu=1}^{M+1}\binom{M}{\nu-1}\frac{(\nu-1)!\,(\nu)_N}{|\Omega|^{\nu}}\sum_{j=0}^{N}\binom{N}{j}\frac{1}{|\Omega|^{j}} \\
&= \left(1+\frac{1}{|\Omega|}\right)^N\sum_{\nu=0}^{M}\binom{M}{\nu}\frac{\nu!\;(\nu+1)_N}{|\Omega|^{\nu+1}} \\
&\overset{(1)}{=} \left(1+\frac{1}{|\Omega|}\right)^N\frac{(M+N)!}{|\Omega|}\sum_{\nu=0}^{M}\binom{M}{\nu}\frac{1}{|\Omega|^{\nu}}.
\end{aligned}$$

Finally, using again the binomial formula $\sum_{\nu=0}^{M}\binom{M}{\nu}\frac{1}{|\Omega|^{\nu}}=(1+\frac{1}{|\Omega|})^M$, we get the estimate (2.3.6). □

Now, we look at the denominators $|\Omega|$, with $\Omega_{k,M,N}=\mathrm{i}k\omega_1+\mathrm{i}\omega_2(M-N)$ and realize that, even though ω_1 and ω_2 are not commensurable, values of the integers k and the difference $M-N$ can be chosen which make $|\Omega_{k,M,N}|$ smaller than any previously fixed quantity.

Thus, to control the size of these small divisors, the frequencies ω_1, ω_2 are asked to satisfy the following Diophantine conditions,

$$|\langle \boldsymbol{k},\boldsymbol{\omega}\rangle| \ge \frac{\gamma}{|\boldsymbol{k}|_1^{\tau}} \tag{2.3.7}$$

for $\tau>1$ and for a certain $\gamma>0$. Here, $\boldsymbol{k}\in\mathbb{Z}^2\setminus\{\boldsymbol{0}\}$ and $\boldsymbol{\omega}^*=(\omega_1,\omega_2)$.

With the definitions,

$$\Xi(\tau,\gamma) := \left\{\boldsymbol{\omega}\in\mathbb{R}^n : |\langle \boldsymbol{k},\boldsymbol{\omega}\rangle| \ge \gamma|\boldsymbol{k}|_1^{-\tau}\ \text{ for all } \boldsymbol{k}\in\mathbb{Z}^n\setminus\{\boldsymbol{0}\}\right\} \tag{2.3.8}$$

and,

$$\Xi(\tau) := \bigcup_{\gamma>0}\Xi(\tau,\gamma), \tag{2.3.9}$$

the following theorem (see Lochack and Meunier, 1988, Appendix 4) is well known.

Theorem 2.3. *According to the values of τ, there are three different cases,*

(i) $0<\tau<n-1$; *then* $\Xi(\tau)=\emptyset$.

(ii) $\tau=n-1$; $\Xi(\tau)$ *has Lebesgue measure zero and Hausdorff measure* n.

(iii) $\tau>n-1$; $\mathbb{R}^n\setminus\Xi(\tau)$ *has zero Lebesgue measure. More precisely, if* B_R *is a ball in* $\mathbb{R}^n$ *of radius* R, *then,*

$$\mathit{meas}\,\{\mathbb{R}^n\setminus\Xi(\tau,\gamma)\cap B_R\} \le C_\tau\gamma R^{n-1},$$

(1) $\nu!\;(\nu+1)_N=\nu!\,(\nu+1)(\nu+2)\cdots(\nu+N)=(\nu+N)!\le(M+N)!$, since $\nu=0,1,2,\ldots,M$.

where C_τ is a constant for a fixed τ.

In our case $n = 2$, so taking $\tau > 1$, the corresponding set $\Xi(\tau)$ of "Diophantine" frequencies has full Lebesgue measure.

From the estimates on $|\Lambda^{-1}|_1$ of lemma 2.2 and the Diophantine conditions just imposed, it is now possible to give more explicit bounds of $G_s^{(1)}$ in a smaller domain. Given $0 < \delta < \rho$ and $\chi = e^{-\delta}$, we have

$$\left|G_s^{(1)}\right|_{\rho-\delta,R\chi} = \sum_{k\in\mathbb{Z}} \sum_{l=0}^{\lfloor s/2\rfloor} \sum_{\substack{M+N\\=s-2l}} \sum_{m=0}^{M} \sum_{n=0}^{N} |g_{k,l,m,M-m,N-n,n}|\, R^s e^{-\delta(|k|+s)} e^{\rho|k|} \tag{2.3.10}$$

(where the sum does not contain terms with $k = 0$ and $M = N$ at the same time). But,

$$\left|\boldsymbol{g}_{k,l,M,N}\right|_1 = \sum_{m=0}^{M} \sum_{n=0}^{N} |g_{k,l,m,M-m,N-n,n}| \tag{2.3.11}$$

is the solution of (1.7.19), with the subscripts k, l, M, N written explicitly, so $\left|\boldsymbol{g}_{k,l,M,N}\right|_1 \le \left|\Lambda^{-1}_{k,l,M,N}\right|_1 \left|\boldsymbol{f}_{k,l,M,N}\right|_1$ (by the consistency property of the matrix norm). Hence, substitution in (2.3.10), gives the bounds,

$$\left|G_s^{(1)}\right|_{\rho-\delta,R\chi} \le \sum_{k\in\mathbb{Z}} \sum_{l=0}^{\lfloor s/2\rfloor} \sum_{M=0}^{s-2l} \left|\Lambda^{-1}_{k,l,M,s-2l-M}\right|_1 \left|\boldsymbol{f}_{k,l,M,s-2l-M}\right|_1 R^s e^{-\delta(|k|+s)} e^{\rho|k|}. \tag{2.3.12}$$

Introducing now the quantity,

$$\widetilde{\alpha}_s := \sup_{\substack{k\in\mathbb{Z},\, 0\le l\le[s/2]\\ 0\le M\le s-2l}} \left|\Lambda^{-1}_{k,l,M,s-2l-M}\right|_1 e^{-\delta(|k|+s)}, \tag{2.3.13}$$

(2.3.12) reads,

$$\begin{aligned}\left|G_s^{(1)}\right|_{\rho-\delta,R\chi} &\le \widetilde{\alpha}_s \sum_{k\in\mathbb{Z}} \sum_{l=0}^{\lfloor s/2\rfloor} \sum_{M=0}^{s-2l} \left|\boldsymbol{f}_{k,l,M,s-2l-M}\right|_1 R^s e^{\rho|k|} \\ &= \widetilde{\alpha}_s \left|F_s^{(1)}\right|_{\rho,R},\end{aligned} \tag{2.3.14}$$

where Poisson series, $F_s^{(1)} \in \mathfrak{E}_s$, stands for the right hand side of the homological equations (1.7.11), without $\mathcal{M}$-type monomials (see definition 1.11) to conform with our early assumption $\Omega_{k,M,N} \neq 0$ for non-negative whole numbers, M, N, with $2l + M + N = s$.

Thus, we need reasonable bounds of the factor $\widetilde{\alpha}_s$. Using the Diophantine conditions (2.3.7),

$$|\Omega| = |k\omega_1 + \omega_2(M-N)| \ge \frac{\gamma}{(|k| + |M-N|)^\tau}$$

in the bounds for $|\Lambda^{-1}|_1$ on lemma 2.2

$$\left|\Lambda^{-1}_{k,l,M,N}\right|_1 \le \left(1 + \frac{(|k| + |M-N|)^\tau}{\gamma}\right)^{M+N} \frac{(|k| + |M-N|)^\tau}{\gamma}(M+N)!. \tag{2.3.15}$$

From here we can derive estimates for the $\widetilde{\alpha}_s$ constants defined in (2.3.13), explicitly,

$$\widetilde{\alpha}_s \leq \left(1+\frac{(|k|+|M-N|)^\tau}{\gamma}\right)^{M+N} \frac{(|k|+|M-N|)^\tau}{\gamma}(M+N)!\,\mathrm{e}^{-\delta(|k|+M+N)}$$
$$\overset{(2)}{\leq} \left(1+\frac{(|k|+|M-N|)^\tau}{\gamma}\right)^{s+1} s!\,\mathrm{e}^{-\delta(|k|+|M-N|)}. \quad (2.3.16)$$

Consider now the function,

$$h(x) := (1+\tfrac{x^\tau}{\gamma})^{s+1}\mathrm{e}^{-\delta x},$$

defined for $x \geq 0$[(3)]. If $0<\delta<1$ is small enough, its derivative,

$$h'(x) = \frac{\mathrm{e}^{-\delta x}}{\gamma}\left(1+\frac{x^\tau}{\gamma}\right)^s g(x),$$

with

$$g(x) := (s+1)\tau x^{\tau-1} - \delta x^\tau - \delta\gamma,$$

has a zero, $\widehat{x}$, in the interval $\mathfrak{J} = (x_1, x_2)$, where

$$x_1 = \max\left\{1, \tfrac{(\tau-1)(s+1)}{\delta}\right\} \qquad \text{and} \qquad x_2 = \frac{\tau(s+1)}{\delta},$$

provided $g(1)>0$. Then $\widehat{x}$ corresponds to a point of (absolute) maximum of $h(x)$[(4)]. To show this, we first check that $g(x_1)>0$, $g(x)$ decreases monotonically for $x>x_1$ and $g(x_2) = -\delta\gamma < 0.$, $\widehat{x} \in \mathfrak{J}$. It should be clear from the expression of the derivative of $h(x)$ above, that $h'(x)>0$, for $x \in (1,\widehat{x})$ and $h'(x)<0$, for $x>\widehat{x}$. Thus $h(x) \leq h(\widehat{x})$, whenever $x \geq 1$. Finally, note that the condition $g(1)>0$ is equivalent to the requirement

$$\frac{(s+1)\tau}{\delta} - \gamma > 1, \quad (2.3.17)$$

for $s \geq 3$ and $\tau > 1$, $\gamma > 0$ fixed. Indeed, this can be achieved taking $0<\delta<1$ sufficiently small. In fact, as $s \geq 3$, it will be enough to take, $\delta < \min\{\frac{4\tau}{1+\gamma}, 1\}$. This considerations will be contemplated again in section 2.5.1, to conveniently adjust the size of the domain.

Moreover, $g(\widehat{x}) = 0$ implies $1+\frac{\widehat{x}^\tau}{\gamma} = \frac{(s+1)\tau}{\delta\gamma}\widehat{x}^{\tau-1}$, and $\widehat{x} < \frac{\tau(s+1)}{\delta}$ with $\tau-1>0$, so we have

$$1+\frac{\widehat{x}^\tau}{\gamma} = \frac{\tau(s+1)}{\delta\gamma}\widehat{x}^{\tau-1} < \frac{\tau(s+1)}{\delta\gamma}\left(\frac{\tau(s+1)}{\delta}\right)^{\tau-1} = \frac{1}{\gamma}\left(\frac{\tau(s+1)}{\delta}\right)^{\tau},$$

(2) Taking into account that: $|M-N| \leq M+N$, so $\mathrm{e}^{-\delta(|k|+M+N)} < \mathrm{e}^{-\delta(|k|+|M-N|)}$; moreover $M+N = s-2l \leq s \Rightarrow (M+N)! \leq s!$, and $(|k|+|M-N|)^\tau/\gamma \leq 1+(|k|+|M-N|)^\tau/\gamma$.

(3) We take $x = |k|+|M-N|$ and further consider x as a continuous variable, but $|k|+|M-N| \geq 1$, since M, N, k are integers and both k and the difference $M-N$ can not be zero simultaneously. Hence, it suffices to define the function $h(x)$ for $x \geq 1$. Nevertheless, it is easier to discuss the position of its maximum if both, $h(x)$ and the auxiliary function $g(x)$, are considered in the whole nonnegative semiaxis.

(4) One checks out immediately that, under the assumption $g(1)>0$, there is another zero of $g(x)$ in the interval $(0, \min\{1, x_1\})$, and no more ones could exist though, because $g(x)$ decreases for $x > x_1$.

and together with the obvious inequality $\mathrm{e}^{-\delta\widehat{x}} < \mathrm{e}^{-(\tau-1)(s+1)}$, they produce the following bound for $h(x)$

$$h(x) \leq h(\widehat{x}) \leq \left(\frac{\mathrm{e}}{\gamma}\right)^{s+1} \left(\frac{\tau(s+1)}{\delta}\right)^{\tau(s+1)} \mathrm{e}^{-\tau(s+1)}, \tag{2.3.18}$$

valid for all $x \geq 1$. The lemma below summarizes the arguments given in this section and shows the estimates on the generating function, $G_s^{(1)}$, to be used along the rest of the chapter.

Lemma 2.4. *With the notation above, and with $0 < \delta < \rho$, $\chi = \mathrm{e}^{-\delta}$, and whenever the frequencies ω_1, ω_2 fulfill the Diophantine condition (2.3.7) for some $\gamma > 0$, $\tau > 1$, the piece of the s-th degree term of the function free of $\mathcal{M}$-type monomials, $G_s^{(1)}$, is bounded by,*

$$\left|G_s^{(1)}\right|_{\rho-\delta,R\chi} \leq \alpha_s \left|F_s^{(1)}\right|_{\rho,R} \tag{2.3.19}$$

with the coefficients α_s defined as,

$$\alpha_s := \sqrt{2\pi}\mathrm{e}^2 \left(\frac{s+1}{\gamma}\right)^{s+1} \left(\frac{\tau(s+1)}{\delta \mathrm{e}}\right)^{\tau(s+1)}, \qquad \textit{if } s \geq 3. \tag{2.3.20}$$

Proof. (2.3.16), with the bounds (2.3.18) on h above, give the following estimates for $\widetilde{\alpha}_s$

$$\widetilde{\alpha}_s \leq s! \left(\frac{\mathrm{e}}{\gamma}\right)^{s+1} \left(\frac{\tau(s+1)}{\delta \mathrm{e}}\right)^{\tau(s+1)},$$

and with the use of the Stirling's formula (see Puig Adam, 1939), $s! = \sqrt{2\pi} s^{s+\frac{1}{2}} \mathrm{e}^{-s+\frac{\xi}{12s}}$, for $s > 0$ and $0 < \xi < 1$. In our case $s \geq 3$, so $\frac{\xi}{12s} < 1$ and then $\mathrm{e}^{\frac{\xi}{12s}} < \mathrm{e}$. Furthermore $s^{s+\frac{1}{2}} < (s+1)^{s+1}$, so $s! < \sqrt{2\pi}\mathrm{e}^2(s+1)^{s+1}\mathrm{e}^{-(s+1)}$. Finally, substitution of this bound for $s!$ in the last inequality leads to $\widetilde{\alpha}_s \leq \alpha_s$ with α_s as defined in (2.3.20). □

2.4 Study of the resonant terms

As it has been repeatedly pointed, up to now, we have only dealt with the solutions of the "nonresonant" (in the sense already explained) part of the homological equations. Time has come, then, to revisit section 1.7.2 and face up to the problem of finding bounds for the solutions of the reduced homological equations (1.7.25).

Additional notation and definitions

To proceed beyond, it is worth revisiting the different variables we use in the representation of polynomials. Originally, we considered polynomials $P \in \mathfrak{E}^s_{\mathcal{M}}$, with s even –see definition (2.2.6)–, in the action I_1, the positions q_1, q_2 and their conjugate momenta p_1, p_2. Later, in section 1.7.2, the coordinates were grouped in the special variables $\boldsymbol{\eta}^* = (\eta_1, \eta_2, \eta_3, \eta_4)$, related with the former ones through,

$$\boldsymbol{\eta}^*(\boldsymbol{q}, \boldsymbol{p}) = (q_1p_1, q_1p_2, q_2p_2, q_2p_1). \tag{2.4.1}$$

Furthermore, even another set of variables, $\boldsymbol{\xi}^* = (\xi_1, \xi_2, \xi_3, \xi_4)$, was introduced by means of a linear change $\boldsymbol{\eta}(\boldsymbol{\xi})$,

$$\boldsymbol{\eta}(\boldsymbol{\xi}) = \begin{pmatrix} -i & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -i & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix}, \tag{2.4.2}$$

and its inverse $\boldsymbol{\xi}(\boldsymbol{\eta})$,

$$\boldsymbol{\xi}(\boldsymbol{\eta}) = \begin{pmatrix} i/2 & 0 & i/2 & 0 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{pmatrix}. \tag{2.4.3}$$

The homological equations were solved in these variables ξ_1, ξ_2, ξ_3 and ξ_4, since when expressed with respect to them, the independent term was easily separable –by projection on the corresponding subspaces–, into resonant and nonresonant components. In the forthcoming, it will be convenient to distinguish explicitly between the expression of polynomials with respect to those three different set of variables, –i. e., $(I_1, \boldsymbol{q}, \boldsymbol{p})$; $(I_1, \boldsymbol{\eta})$ and $(I_1, \boldsymbol{\xi})$. So, given a polynomial P, we shall denote

$$P(I_1, \boldsymbol{q}, \boldsymbol{p}) = \widetilde{P}(I_1, \boldsymbol{\eta}(\boldsymbol{q}, \boldsymbol{p})), \tag{2.4.4a}$$
$$\widetilde{P}(I_1, \boldsymbol{\eta}) = \check{P}(I_1, \boldsymbol{\xi}(\boldsymbol{\eta})), \tag{2.4.4b}$$
$$\check{P}(I_1, \boldsymbol{\xi}) = \widetilde{P}(I_1, \boldsymbol{\eta}(\boldsymbol{\xi})). \tag{2.4.4c}$$

In other words, P will stand for the polynomial in $(I_1, \boldsymbol{q}, \boldsymbol{p})$ and $\widetilde{P}$, $\check{P}$ for its expression in the sets of variables $(I_1, \boldsymbol{\eta})$ and $(I_1, \boldsymbol{\xi})$ respectively. So, if $G_s^{(2)}$ is the piece of the s-degree term, G_s, of the generating function formed by $\mathcal{M}$-type monomials, as pointed in section 1.7.2, and in accordance with our just introduced notation, $G_s^{(2)}$ may be expressed as,

$$\check{G}_s^{(2)} = \sum_{\substack{l+\nu+m+n=s/2 \\ (\nu \geq 1)}} f_{l,\nu,m,n} I_1^l \xi_2^\nu \xi_1^m \xi_3^n + \sum_{\substack{l+\nu+m+n=s/2 \\ (\nu \geq 0)}} g_{l,\nu,m,n} I_1^l \xi_4^\nu \xi_1^m \xi_3^n, \tag{2.4.5}$$

see equation (1.7.46b). Here, an additional subscript, l, has been added to account for the degree of the action I_1.

Definition 2.5. *Given a polynomial $P = \sum_{\boldsymbol{m}} a_{\boldsymbol{m}} \boldsymbol{z}^{\boldsymbol{m}}$, defined on the complex domain,*

$$\mathcal{D}(\widehat{R}) = \{\boldsymbol{z} \in \mathbb{C}^5 : |z_j| \leq \widehat{R}, j = 1, 2, 3, 4, 5\},$$

we introduce the auxiliary norm $\|\cdot\|_{\widehat{R}}$,

$$\|P\|_{\widehat{R}} := \sum_{\boldsymbol{m}} |a_{\boldsymbol{m}}| \widehat{R}^{|\boldsymbol{m}|_1}. \tag{2.4.6}$$

Lemma 2.6. *The norm defined by* (2.4.6) *satisfies,* *(i) If* $P \in \mathfrak{E}^s_{\mathcal{M}}$*, then* $\|\widetilde{P}\|_{R^2} = |P|_{\rho,R}$ *(where* $|\cdot|_{\rho,R}$ *is the norm defined by* (2.2.4)*).*

(ii) $\|PQ\|_{\widehat{R}} \le \|P\|_{\widehat{R}}\|Q\|_{\widehat{R}}$.

(iii) $\|\widetilde{P}(\boldsymbol{\eta})\|_{\widetilde{R}} \le \|\check{P}(\boldsymbol{\xi})\|_{\widetilde{R}}$, $\|\check{P}(\boldsymbol{\xi})\|_{\check{R}} \le \|\widetilde{P}(\boldsymbol{\eta})\|_{2\check{R}}$.

Proof. The first item follows immediately from the definition, while the second and the third ones can be thought of as special cases for polynomials of item (ii), lemma A.2, and of lemma A.4 (applied to particular transformations $\boldsymbol{\xi}(\boldsymbol{\eta})$, $\boldsymbol{\eta}(\boldsymbol{\xi})$) respectively, taking into account that,

$$\max_{i=1,2,3,4}\{\|\xi_i(\boldsymbol{\eta})\|_{\widetilde{R}}\} \le \widetilde{R}, \qquad \text{and} \qquad \max_{i=1,2,3,4}\{\|\eta_i(\boldsymbol{\xi})\|_{\check{R}}\} \le 2\check{R}.$$

□

With the definition 2.5, the norm $\|\cdot\|_{R^2\chi^2}$, with $\chi = \exp(-\delta)$, of $\check{G}^{(2)}_s$ in (2.4.5) is

$$\|\check{G}^{(2)}_s\|_{R^2\chi^2} = \left(\sum_{\substack{l+\nu+m+n=s/2 \\ (\nu\ge 1)}} |f_{l,\nu,m,n}| + \sum_{\substack{l+\nu+m+n=s/2 \\ (\nu\ge 0)}} |g_{l,\nu,m,n}| \right) R^s \exp(-s\delta). \tag{2.4.7}$$

But in view of (1.7.50), the second of the sums in the r. h. s. of (2.4.7) is bounded without effort. Indeed, if we express the independent term of the homological equations restricted to the space $\mathfrak{E}^s_{\mathcal{M}}$, in the variables ξ_1, ξ_2, ξ_3 and ξ_4,

$$\check{F}^{(2)}_s = \sum_{\substack{l+\nu+m+n=s/2 \\ (\nu\ge 0)}} \widehat{f}_{l,\nu,m,n}\, I_1^l \xi_2^\nu \xi_1^m \xi_3^n + \sum_{\substack{l+\nu+m+n=s/2 \\ (\nu\ge 1)}} \widehat{g}_{l,\nu,m,n}\, I_1^l \xi_4^\nu \xi_1^m \xi_3^n \tag{2.4.8}$$

–where all the coefficients $\widehat{f}_{l,\nu,m,n}$, $\widehat{g}_{l,\nu,m,n}$ are real, as in (1.7.46a)–, and we recall that the coefficients $\widehat{f}_{l,\nu,m,0}$ correspond to resonant monomials, so one can define:

$$\check{Z}_s = \sum_{\substack{l+\nu+m=s/2 \\ (\nu\ge 0)}} \widehat{f}_{l,\nu,m,0}\, I_1^l \xi_2^\nu \xi_1^m, \tag{2.4.9}$$

and hence, according to the norm in definition 2.5 and the relations (2.4.4a)-(2.4.4c), the functions $Z_s(\boldsymbol{q}, I_1, \boldsymbol{p})$, $s = 3, \ldots, r$, holding the resonant terms, can be bounded, applying lemma 2.6, as:

$$\begin{aligned} |Z_s|_{\rho,R} &= \|\widetilde{Z}_s\|_{R^2} = \|\check{Z}_s\|_{R^2} \\ &\le \|\check{F}_s\|_{R^2} = 2^{s/2}\|\check{F}_s\|_{R^2/2} \le 2^{s/2}\|\widetilde{F}_s\|_{R^2} = 2^{s/2}|F_s|_{\rho,R}. \end{aligned} \tag{2.4.10}$$

Then, we get the estimates,

$$\begin{aligned} \sum_{\substack{l+\nu+m+n=s/2 \\ (\nu\ge 0)}} |g_{l,\nu,m,n}| &= \sum_{\nu=0}^{s/2-l-1} \sum_{\substack{\nu+m+n=s/2-l \\ (n\ne 0)}} \frac{|\widehat{g}_{l,\nu+1,m,n-1}|}{n} \\ &\le \sum_{\nu=0}^{s/2-l-1} \sum_{\substack{\nu+m+n=s/2-l \\ (n\ne 0)}} |\widehat{g}_{l,\nu+1,m,n-1}| \end{aligned} \tag{2.4.11}$$

(we recall that the arbitrary coefficients $g_{l,\nu,m,0}$ were taken equal to 0). However, it is a more involved task to obtain suitable bounds for the first sum in (2.4.7). We have split such a task into two basic steps.

Step 1. With s and l fixed, we compute explicit solutions of linear algebraic systems of type (1.7.53), now, $\Lambda'_{l,\nu}\boldsymbol{f}_{l,\nu} = \widehat{\boldsymbol{f}}_{l,\nu}$, with

$$\boldsymbol{f}_{l,\nu} = \begin{pmatrix} f_{l,\nu+1,0,s/2-l-\nu-1} \\ f_{l,\nu+1,1,s/2-l-\nu-2} \\ f_{l,\nu+1,2,s/2-l-\nu-3} \\ \vdots \\ f_{l,\nu+1,s/2-l-\nu-2,1} \\ f_{l,\nu+1,s/2-l-\nu-1,0} \end{pmatrix}, \quad \widehat{\boldsymbol{f}}_{l,\nu} = \begin{pmatrix} \widehat{f}_{l,\nu,0,s/2-l-\nu} \\ \widehat{f}_{l,\nu,1,s/2-l-\nu-1} \\ \widehat{f}_{l,\nu,2,s/2-l-\nu-2} \\ \vdots \\ \widehat{f}_{l,\nu,s/2-l-\nu-2,2} \\ \widehat{f}_{l,\nu,s/2-l-\nu-1,1} \end{pmatrix}, \tag{2.4.12}$$

–i. e., they are vectors with the same structure as in (1.7.52), with a new subscript "l" and $s/2-l$ in substitution of M–. In the same way, the matrix $\Lambda'_{l,\nu}$ is given by (1.7.54), but with $s/2-l$ instead of M everywhere.

Step 2. The components of the vector $\boldsymbol{f}_{l,\nu}$ are directly bounded in terms of those of $\widehat{\boldsymbol{f}}_{l,\nu}$. If in addition, one introduces the vectors $\boldsymbol{f}_l, \widehat{\boldsymbol{f}}_l \in \mathbb{R}^d$, $d = (s/2-l)(s/2-l+1)/2$,

$$\boldsymbol{f}^*{}_l = (\boldsymbol{f}^*{}_{l,1}, \boldsymbol{f}^*{}_{l,2}, \boldsymbol{f}^*{}_{l,3}, \ldots, \boldsymbol{f}^*{}_{l,s/2-l}), \tag{2.4.13a}$$

$$\widehat{\boldsymbol{f}}^*{}_l = (\widehat{\boldsymbol{f}}^*{}_{l,0}, \widehat{\boldsymbol{f}}^*{}_{l,1}, \widehat{\boldsymbol{f}}^*{}_{l,2}, \ldots, \widehat{\boldsymbol{f}}^*{}_{l,s/2-l-1}), \tag{2.4.13b}$$

then, the sum $|\boldsymbol{f}_l|_1 = \sum |f_{l,\nu,m,n}|$ in (2.4.7) can be bounded as $|\boldsymbol{f}_l|_1 \le \frac{1}{2}|\widehat{\boldsymbol{f}}_l|_1$.

The first step is accomplished after some direct algebra on the equations (1.7.51a)–(1.7.51c); from there, the solution of the system

$$\Lambda'_{l,\nu}\boldsymbol{f}_{l,\nu} = \widehat{\boldsymbol{f}}_{l,\nu} \tag{2.4.14}$$

is derived straightforward from the general recurrence relation,

$$\begin{aligned} f_{l,\nu+1,m,s/2-l-\nu-m-1} = &-\frac{1}{s/2-l+\nu-m+1}\widehat{f}_{l,\nu,m,s/2-l-\nu-m} \\ &+\frac{s/2-l-\nu-m+1}{s/2-l+\nu-m+1} f_{l,\nu+1,m-2,s/2-l-\nu-m+1}, \end{aligned} \tag{2.4.15}$$

with $2 \le m \le s/2 - l - \nu - 1$. The first terms of the solution are,

$$
\begin{aligned}
f_{l,\nu+1,0,s/2-l-\nu-1} &= -\frac{1}{s/2-l+\nu+1}\widehat{f}_{l,\nu,0,s/2-l-\nu},\\
f_{l,\nu+1,1,s/2-l-\nu-2} &= -\frac{1}{s/2-l+\nu}\widehat{f}_{l,\nu,1,s/2-l-\nu-1},\\
f_{l,\nu+1,2,s/2-l-\nu-3} &= -\tfrac{1}{s/2-l+\nu-1}\widehat{f}_{l,\nu,2,s/2-l-\nu-2}\\
&\quad+\tfrac{s/2-l-\nu-1}{(s/2-l+\nu-1)(s/2-l+\nu+1)}\widehat{f}_{l,\nu,0,s/2-l-\nu},\\
f_{l,\nu+1,3,s/2-\nu-l-4} &= -\tfrac{1}{s/2-l+\nu-2}\widehat{f}_{l,\nu,3,s/2-l-\nu-3}\\
&\quad+\tfrac{s/2-l-\nu-2}{(s/2-l+\nu-2)(s/2-l+\nu)}\widehat{f}_{l,\nu,1,s/2-l-\nu-1},\\
f_{l,\nu+1,4,s/2-l-\nu-5} &= -\tfrac{1}{s/2-l+\nu-3}\widehat{f}_{l,\nu,4,s/2-l-\nu-4}\\
&\quad+\tfrac{s/2-l-\nu-3}{(s/2-l+\nu-3)(s/2-l+\nu-1)}\widehat{f}_{l,\nu,2,s/2-l-\nu-2}\\
&\quad-\tfrac{(s/2-l-\nu-3)(s/2-l-\nu-1)}{(s/2-l+\nu-3)(s/2-l+\nu-1)(s/2-l+\nu+1)}\widehat{f}_{l,\nu,0,s/2-l-\nu},\\
&\ \ \vdots
\end{aligned}
$$

and so on. By induction it can be shown that the general term is,

$$
\begin{aligned}
f_{l,\nu+1,m,s/2-l-\nu-m-1} &= -\tfrac{1}{s/2-l+\nu-m+1}\,\widehat{f}_{l,\nu,m,s/2-l-\nu-m}\\
&+\sum_{\alpha=1}^{\lfloor m/2\rfloor}(-1)^{\alpha+1}\tfrac{(s/2-l+\nu-m-1)!!\,(s/2-l-\nu-m+2\alpha-1)!!}{(s/2-l-\nu-m-1)!!\,(s/2-l+\nu-m+2\alpha+1)!!}\widehat{f}_{l,\nu,m-2\alpha,s/2-l-\nu-m+2\alpha}.
\end{aligned}
\tag{2.4.16}
$$

Indeed, for the formula above has just been checked for $m = 2, 3, 4$. Suppose, thus, that it works also for $2 < m < s/2 - l - \nu - 1$. Therefore, substitution in the recurrence relation (2.4.15), written for $m + 1$, yields

$$
\begin{aligned}
f_{l,\nu+1,m+1,s/2-l-\nu-m-2} &= \tfrac{-1}{s/2-l+\nu-m}\widehat{f}_{l,\nu,m+1,s/2-l-\nu-m-1}\\
&\quad+\tfrac{s/2-l-\nu-m}{(s/2-l+\nu-m)(s/2-l+\nu-m+2)}\widehat{f}_{l,\nu,m+1,s/2-l-\nu-m-1}\\
&-\sum_{\alpha=1}^{\lfloor (m-1)/2\rfloor}(-1)^{\alpha+1}\tfrac{(s/2-l+\nu-m-2)!!(s/2-l-\nu-m+2\alpha)!!}{(s/2-l-\nu-m-2)!!(s/2-l+\nu-m+2\alpha+2)!!}\widehat{f}_{l,\nu,m-1-2\alpha,s/2-l-\nu-m+2\alpha+1}
\end{aligned}
\tag{2.4.17}
$$

and after the inclusion of the second term in the r. h. s. into the sum, followed by a displacement of the index α to $\alpha + 1$, the last expression arranges to,

$$
\begin{aligned}
f_{l,\nu+1,m+1,s/2-l-\nu-m-2} &= \tfrac{-1}{s/2-l+\nu-m}\widehat{f}_{l,\nu,m+1,s/2-l-\nu-m-1}\\
&+\sum_{\alpha=1}^{\lfloor (m+1)/2\rfloor}(-1)^{\alpha+1}\tfrac{(s/2-l+\nu-m-2)!!(s/2-l-\nu-m-1+2\alpha-1)!!}{(s/2-l-\nu-m-2)!!(s/2-l+\nu-m-1+2\alpha+1)!!}\widehat{f}_{l,\nu,m+1-2\alpha,s/2-l-\nu-m-1+2\alpha}
\end{aligned}
\tag{2.4.18}
$$

which matches (2.4.16) with $m + 1$ in the place of m. This completes the induction and hence the first step.

The second step begins with the quest for appropriate bounds on the different components of $\boldsymbol{f}_{l,\nu}$. For $m=0,1$ we have directly

$$\begin{aligned}|f_{l,\nu+1,0,s/2-l-\nu-1}| &= \tfrac{1}{s/2-l+\nu+1}|\widehat{f}_{l,\nu,0,s/2-l-\nu}|,\\ |f_{l,\nu+1,1,s/2-l-\nu-2}| &= \tfrac{1}{s/2-l+\nu}|\widehat{f}_{l,\nu,1,s/2-l-\nu-1}|\end{aligned}$$

while a glance on the general term (2.4.16) of the solution for $m\geq 2$ leads to the estimate

$$\begin{aligned}|f_{l,\nu+1,m,s/2-l-\nu-m-1}| \leq{}& \tfrac{1}{s/2-l+\nu-m+1}|\widehat{f}_{l,\nu,m,s/2-l-\nu-m}|\\ &+\sum_{\alpha=1}^{[m/2]}\tfrac{1}{s/2-l+\nu-m+2\alpha+1}|\widehat{f}_{l,\nu,m-2\alpha,s/2-l-\nu-m+2\alpha}|.\end{aligned} \tag{2.4.19}$$

Let us put $M=s/2-l$ and $j=m-2\alpha$. Then, note that the coefficient of

$$|\widehat{f}_{l,\nu,m-2\alpha,s/2-l-\nu-m+2\alpha}| = |\widehat{f}_{l,\nu,j,M-j-\nu}|,$$

is $\frac{1}{M-j+1+\nu}$, which is independent of the index α in the sum (2.4.19) above. Therefore:

$$\sum_{m=0}^{M-1-\nu}|f_{l,\nu+1,m,M-\nu-m-1}| \leq \sum_{j=0}^{M-\nu-1}\frac{N_{\#}}{M-j+1+\nu}|\widehat{f}_{l,\nu,j,M-j-\nu}|,$$

being

$$N_{\#} = \#\left\{m \text{ such that the term } |\widehat{f}_{l,\nu,j,M-j-\nu}| \text{ appear in the sum (2.4.19)}\right\},$$

but the first m giving rise to $|\widehat{f}_{l,\nu,j,M-j-\nu}|$ is $m=j$, (and $\alpha=0$), the second is $m=j+2$ (and $\alpha=1$), the third is $m=j+4$ ($\alpha=2$) and so on. So it must be $N_{\#}=\lfloor\frac{M-1-\nu-j}{2}\rfloor+1$. Hence:

$$\frac{N_{\#}}{M-j+1+\nu} \leq \frac{\frac{1}{2}(M-j+\nu+1)}{M-j+\nu+1} \leq \frac{1}{2},$$

and therefore:

$$\sum_{\nu=0}^{s/2-l-1}\sum_{m=0}^{s/2-l-\nu-1}|f_{l,\nu+1,m,s/2-l-\nu-m-1}| \leq \frac{1}{2}\sum_{\nu=0}^{s/2-l-1}\sum_{j=0}^{s/2-l-\nu-1}|\widehat{f}_{l,\nu,j,s/2-l-j-\nu}|.$$

Using this in (2.4.7) and taking (2.4.11) into account,

$$\begin{aligned}\|\check{G}_s^{(2)}\|_{R^2\chi^2} \leq{}& \left(\sum_{\substack{l+\nu+m+n=s/2\\(\nu\geq 0)}}|\widehat{f}_{l,\nu,m,n}| + \sum_{\substack{l+\nu+m+n=s/2\\(\nu\geq 1)}}|\widehat{g}_{l,\nu,m,n}|\right)\frac{1}{2}R^s\exp(-s\delta)\\ \leq{}& \frac{1}{2}\exp(-s\delta)\|\check{F}_s^{(2)}\|_{R^2}.\end{aligned} \tag{2.4.20}$$

From the expressions (2.4.2) and (2.4.3) for the linear transformation $\boldsymbol{\eta}(\boldsymbol{\xi})$ and its corresponding inverse, $\boldsymbol{\eta}(\boldsymbol{\xi})$, it follows that $\|\xi_i(\boldsymbol{\eta})\|_{\chi^2 R^2} = \chi^2 R^2$ and $\|\eta_i(\boldsymbol{\xi})\|_{R^2/2} \leq R^2$, for $i = 1, 2, 3, 4$. Hence, using items (i) and (iii), of lemma 2.6,

$$\begin{aligned} |G_s^{(2)}|_{\rho-\delta,R\chi} = \|\widetilde{G}_s^{(2)}\|_{R^2\chi^2} \leq \|\check{G}_s^{(2)}\|_{R^2\chi^2} &\leq \frac{1}{2}\exp(-s\delta)\|\check{F}_s^{(2)}\|_{R^2} \\ &\leq \frac{1}{2}\exp(-s\delta)\|\widetilde{F}_s^{(2)}\|_{2R^2} = \frac{1}{2}2^{s/2}\exp(-s\delta)\|\widetilde{F}_s^{(2)}\|_{R^2} \\ &\leq \frac{1}{2}2^{s/2}\exp(-s\delta)|F_s^{(2)}|_{\rho,R}, \end{aligned} \tag{2.4.21}$$

which gives an inequality of type,

$$|G_s^{(2)}|_{\rho-\delta,R\chi} \leq \breve{\beta}_s |F_s^{(2)}|_{\rho,R},$$

with $\breve{\beta}_s := \frac{1}{2}2^{s/2}\exp(-s\delta)$; but $\breve{\beta}_s \leq \alpha_s$ for $s \geq 3$ –with α_s given by (2.3.20)–, provided that the coefficient γ in the Diophantine conditions (2.3.7) not to be too large (we can suppose by the moment that $\gamma < 1$, see remark 2.10). Moreover, it is:

$$\begin{aligned} |G_s|_{\rho-\delta,R\chi} &= |G_s^{(1)}|_{\rho-\delta,R\chi} + |G_s^{(2)}|_{\rho-\delta,R\chi}, \\ |F_s|_{\rho,R} &= |F_s^{(1)}|_{\rho,R} + |F_s^{(2)}|_{\rho,R}, \end{aligned}$$

so then,

$$\begin{aligned} |G_s|_{\rho-\delta,R\chi} &= |G_s^{(1)}|_{\rho-\delta,R\chi} + |G_s^{(2)}|_{\rho-\delta,R\chi} \\ &\leq \alpha_s |F_s^{(1)}|_{\rho,R} + \breve{\beta}_s |F_s^{(2)}|_{\rho,R} \\ &\leq \alpha_s \left(|F_s^{(1)}|_{\rho,R} + |F_s^{(2)}|_{\rho,R}\right) \\ &\leq \alpha_s |F_s|_{\rho,R}. \end{aligned}$$

We state this conclusion in a new lemma which extends the preceding lemma 2.4 and gives the effective estimates for the solutions of the s^{th}-degree homological equations, G_s, to be used in the forthcoming.

Lemma 2.7. *Consider the homological equations* (1.7.11) *for the terms of degree s, with $s = 3, 4, \ldots$, that is:*

$$L_{H_2} G_s + Z_s = F_s.$$

With the same notation of lemma 2.4, with $0 < \delta < \rho$, $\chi = \mathrm{e}^{-\delta}$, the frequencies ω_1, ω_2 satisfying the Diophantine condition (2.3.7) *for some $0 < \gamma < 1$ and some $\tau > 1$, the terms G_s of the generating function can be estimated –in a slightly reduced domain–, according to*

$$|G_s|_{\rho-\delta,R\chi} \leq \alpha_s |F_s|_{\rho,R}, \tag{2.4.22}$$

where α_s is the quantity defined by (2.3.20). *Moreover, for the resonant terms Z_s, we have:*

$$|Z_s|_{\rho,R} \leq 2^{s/2} |F_s|_{\rho,R}.$$

From here, the idea is –once the generating function of the canonical change has been determined and bounded–, to use such bounds to estimate the quantities $f_{l,k}$ in the recursive formula (1.7.6). As their sum gives the components of the transformed function $T_G f$, say F_s, with $T_G f = \sum_{s\geq 1} F_s$ (see definition 1.7 in the previous chapter), this process will allow us to bound those components for each degree s. So, if the nonlinear reduction has been carried out up to some finite degree, r (thus, taking $G_{r+1} = G_{r+2} = \cdots = 0$), to compute the size of the remainder we still need to determine the size of the sum $\sum_{s>r} F_s$. These questions are investigated in the sections to come.

2.5 Bounds on the transformed function

In order to obtain estimates of the s-degree terms, F_s, in the transformed function $F = T_G f$, since $F_s = \sum_{l=1}^{s} f_{l,s-l}$, $f_{l,0} = f_l$, some knowledge about the size of $f_{l,m}$ is first required. From now on and up to the end of this chapter, we shall use the following notation and definitions

2.5.1 New notation and definitions

We suppose the function $f = \sum_{l\geq 1} f_l$ is defined and analytic in the complex domain $\mathcal{D}(\rho^*, R^*)$, given by (1.6.3), with R^*, ρ^* small enough to make $|f|_{\rho^*,R^*} < +\infty$. Let c be precisely the norm of f, i. e.,

$$c := |f|_{\rho^*,R^*}. \tag{2.5.1}$$

Next we reduce our initial domain taking R_0 and ρ_0 two positive quantities satisfying the inequalities,

$$R_0 < \min\{1, c, R^*\}, \quad \rho_0 < \min\left\{1, \rho^*, \tfrac{32}{\mathrm{e}\gamma^{1/\tau}}, \tfrac{128\tau}{\gamma+1}\right\}, \tag{2.5.2}$$

and, still, consider ρ, R with $0 < \rho < \rho_0$, $0 < R < R_0$. Thus, the complex domain is reduced to $\mathcal{D}(\rho, R)$. From the definition of the norm $|\cdot|_{\rho,R}$ and from (2.5.1) one gets the following *analytic bounds* on each f_s:

$$|f_s|_{\rho,R} \leq c\left(\frac{R}{R_0}\right)^s, \tag{2.5.3}$$

for $s = 1, 2, \ldots$; they are deduced straightforward since

$$|f_s|_{\rho,R} = |f_s|_{\rho_0,R_0}\left(\frac{R}{R_0}\right)^s \leq |f_s|_{\rho^*,R^*}\left(\frac{R}{R_0}\right)^s \leq c\left(\frac{R}{R_0}\right)^s.$$

Clearly, $|f_s|_{\rho^*,R^*} \leq |f|_{\rho^*,R^*}$, and therefore $|f_s|_{\rho^*,R^*} \leq |f|_{\rho^*,R^*} \leq c$ (this trick has been used to state the last inequality in the chain above). Nevertheless, we advance that this last domain, $\mathcal{D}(\rho_0, R_0)$, will be later successively shrunken down up to a smaller one, in which all the bounds for the functions $f_{l,\nu}$ should be valid. We define (see corollary A.7 of appendix A):

$$\rho_2 = \rho, \qquad \rho_\nu := \rho_2 - 2\sum_{\sigma=3}^{\nu}\delta_\sigma, \tag{2.5.4a}$$

$$R_2 = R, \qquad R_\nu := R\exp\left(-2\sum_{\sigma=3}^{\nu}\delta_\sigma\right). \tag{2.5.4b}$$

In particular, for the sequence $\{\delta_\nu\}_{3\le\nu\le s+2}$ the following determination is chosen: define for some $r\in\mathbb{N}$, $r\ge 3$, and any $s\ge 1$ fixed,

$$\delta := \frac{\rho/16}{r-1}, \tag{2.5.5}$$

and then, for the terms δ_ν, we take $\delta_\nu=\delta_\nu^{(s)}$, with:

$$\delta_\nu^{(s)} := \begin{cases} \delta, & \text{if } 3\le\nu\le r,\\ \dfrac{\delta}{s}, & \text{if } r<\nu\le s+2, \end{cases} \tag{2.5.6}$$

(so $\delta_3^{(s)}=\delta_4^{(s)}=\cdots=\delta_{s+2}^{(s)}$ if $s\le r-2$). Hence, for the sum of the terms,

$$\sum_{\nu=3}^{s+2}\delta_\nu^{(s)}=\sum_{\nu=3}^{r}\delta_\nu^{(s)}+\sum_{\nu=r+1}^{s+2}\delta_\nu^{(s)}=(r-2)\delta+(s-r+2)\frac{\delta}{s}\le(r-1)\frac{\rho/16}{r-1}=\frac{\rho}{16},$$

if $s>r-2$. Conversely, if $s\le r-2$ then

$$\sum_{\nu=3}^{s+2}\delta_\nu^{(s)}=s\delta=s\frac{\rho/16}{r-1}<\frac{\rho}{16}.$$

Therefore, for all $3\le\nu\le s+2$,

$$\sum_{\sigma=3}^{\nu}\delta_\sigma^{(s)}\le\frac{\rho}{16}. \tag{2.5.7}$$

Remark 2.8. The superscript (s) in definition (2.5.6) of δ_ν was added only to emphasize their specific dependence on the degree s. We want to stress that for any given $s_1\ge 1$, the sequence $\{\delta_\nu^{(s_1)}\}_{3\le\nu\le s_1+2}$ will determine the successive reduction of the domains only when we seek bounds for F_{s_1}. For a term of different degree, say F_{s_2} with $s_2\ne s_1$, one must construct another *different* sequence, $\{\delta_\nu^{(s_1)}\}_{3\le\nu\le s_1+2}$, letting $s=s_2$ in (2.5.6). Once this precision is understood, the superscript (s) can be thrown away, as we shall do in what follows to avoid an overload of notation. ♠

Remark 2.9. With (2.5.7) and from the definitions (2.5.4a), (2.5.4b), it is clear that,

$$7\rho/8\le\rho_\nu,\qquad R\exp(-\rho/8)\le R_\nu, \tag{2.5.8}$$

for all $\nu=3,4,\dots,s+2$. We can take $\mathcal{D}(7\rho/8,R\exp(-\rho/8))$ as the common domain where all the estimates in the norm of $f_{l,\nu}$ will work for all $\nu=3,4,\dots,s+2$, and all $s\ge 1$. ♠

Remark 2.10. The condition on ρ_0 in (2.5.2), together with the definition (2.5.5) for δ guarantees,

(C1) $1\le\alpha_j\le\alpha_k$, for $3\le j\le k\le r$, because it makes $\gamma(\delta e)^\tau<1$ in (2.3.20). Indeed,

$$\gamma(\delta\exp(1))^\tau=\gamma\left(\tfrac{\rho\exp(1)}{16(r-1)}\right)^\tau\underset{(r\ge3)}{\le}\gamma\left(\tfrac{\rho\exp(1)}{32}\right)^\tau\underset{(\rho<\rho_0)}{<}\gamma\left(\tfrac{1}{32}\tfrac{32}{\exp(1)}\tfrac{\exp(1)}{\gamma^{1/\tau}}\right)^\tau=\gamma\frac{1}{\gamma}=1.$$

Note also that with this choice of ρ_0 the condition $\gamma<1$ in lemma 2.7 can be dropped.

(C2) With $\rho < \rho_0$, the condition (2.3.17) is fulfilled. Really,

$$\tfrac{(s+1)\tau}{\delta} - \gamma = 16\tfrac{(s+1)(r-1)\tau}{\rho} - \gamma \underset{(3\le s\le r)}{\ge} \tfrac{128\tau}{\rho} - \gamma \underset{(\rho<\rho_0)}{>} \tfrac{128\tau}{\rho_0} - \gamma > \tfrac{128\tau(\gamma+1)}{128\tau} - \gamma = 1.$$

Item (C1) states that the $r-2$ quantities $\alpha_3, \alpha_4, \ldots, \alpha_r$ form a nondecreasing sequence with all its terms greater than 1. This will play –as we shall see later on in section 2.7–, an important role in the process of bounding the terms of the generating function. ♠

To state lemma 2.13, which provides bounds for the terms $f_{l,\nu}$ (those given in definition (1.7.6)), we need to introduce a pair of definitions –included also in appendix A–, (lemma A.11).

Definition 2.11. *Let m, n be two nonnegative integers, with $n > 0$; then*

$$[m, n] := m - n \left\lfloor \frac{m}{n} \right\rfloor, \tag{2.5.9}$$

i. e., the square bracket of the integers $m \ge 0$ and $n > 0$, $[m, n]$, is the remainder of the integer division of m by n.

Definition 2.12. *Consider the sequence of $r-1$ numbers, $\{\beta_s\}_{2\le s\le r}$, such that,*

$$1 = \beta_2 \le \beta_3 \le \beta_4 \le \cdots \le \beta_{r-1} \le \beta_r.$$

Given two positive integers r, s, the quantities $\mathcal{W}(r, s)$ are defined by the products,

$$\mathcal{W}(r, s) := \left(\prod_{\nu=1}^{r-2} \beta_{\nu+2}\right)^{\left\lfloor \frac{s}{r-2} \right\rfloor} \left(\prod_{\nu=0}^{[s,r-2]} \beta_{\nu+2}\right). \tag{2.5.10}$$

With these two previous definitions in mind, we go on with the enunciation of the just mentioned lemma.

Lemma 2.13. *For some given $r \ge 3$, let $s \ge 1$ be a fixed integer, consider the sequence $\{\delta_\nu\}_{3\le\nu\le s+2}$ defined by* (2.5.6)[5]*, the corresponding sequences $\{\rho_\nu\}_{3\le\nu\le s+2}$, $\{R_\nu\}_{3\le\nu\le s+2}$ introduced in* (2.5.4a) *and* (2.5.4b) *respectively and $\{\beta_\nu\}_{2\le\nu\le r}$ a non decreasing sequence with $\beta_2 = 1$. Assume that,*

$$\begin{aligned} |G_3|_{\rho_3+\delta_3, R_3\exp(\delta_3)} &\le \beta_3 b \left(\frac{R}{R_0}\right)^3, \\ |G_\nu|_{\rho_\nu+\delta_3, R_\nu\exp(\delta_3)} &\le \beta_3\beta_4\ldots\beta_\nu \frac{a^{\nu-3}b}{\delta_3^{2(\nu-3)}} \left(\frac{R}{R_0}\right)^\nu, \quad 3 < \nu \le r. \end{aligned} \tag{2.5.11}$$

(with a, b positive constants). Then, the following bounds on $f_{l,\nu}$ apply,

$$|f_{l,\nu}|_{\rho_{\nu+2}, R_{\nu+2}} \le \mathcal{W}(r, \nu) \frac{\vartheta_\nu c}{\delta_3^{2\nu}} \left(\frac{R}{R_0}\right)^{l+\nu} \tag{2.5.12}$$

[5] With the superindex (s) dropped, (see remark 2.8).

for all $l, \nu \in \mathbb{N}$ with $l \geq 1$, $0 \leq \nu \leq s$. Here, the terms ϑ_ν are defined recursively through,

$$
\begin{aligned}
\vartheta_0 &= 1, \\
\vartheta_\nu &= \frac{\delta_3}{\delta_{\nu+2}} \sum_{j=1}^{\min\{\nu, r-2\}} \frac{j}{\nu} a^{j-1} d \vartheta_{\nu-j}, \quad 1 \leq \nu \leq s
\end{aligned}
\tag{2.5.13}
$$

with the constants $d = \frac{17b\exp(1)}{2R_0^2}$ and c given by (2.5.1)*; whilst the symbol $\mathcal{W}(r,s)$ stands for the products* (2.5.10) *introduced in definition 2.12.*

Remark 2.14. We note that to obtain the bounds for $f_{l,\nu}$ we have stated *a priori* estimations for the terms G_ν of the generating function, where the constants a, b will be conveniently exacted later in proposition 2.28. ♠

Remark 2.15. Before proceeding with the proof of lemma 2.13, we point out that bounds for sums of Poisson brackets of type $\sum_{j=1}^{\nu} \frac{j}{\nu}\{f_{l,\nu-j}, G_{2+j}\}$ will be required, but every term therein can be bounded directly applying corollary A.7, so

$$
\begin{aligned}
&\sum_{j=1}^{\nu} \tfrac{j}{\nu} |\{f_{l,\nu-j}, G_{2+j}\}|_{\rho_{\nu+2}, R_{\nu+2}} \leq \\
&\qquad \leq \sum_{j=1}^{\nu} \tfrac{j}{\nu} \tfrac{17\exp(1)}{2R^2\delta_{\nu+2}\delta_{j+2}} |f_{l,\nu-j}|_{\rho_{\nu-j+2}, R_{\nu-j+2}} |G_{2+j}|_{\rho_{2+j}+\delta_{2+j}, R_{2+j}\exp(\delta_{2+j})}.
\end{aligned}
\tag{2.5.14}
$$

This provides recursive bounds on these sums. Without explicit mention, we shall apply the above formula throughout. ♠

Proof of lemma 2.13. The estimate (2.5.12) works for $\nu = 1$ and for all $l \geq 1$,

$$
\begin{aligned}
|f_{l,1}|_{\rho_3, R_3} &= |\{f_{l,0}, G_3\}|_{\rho_3, R_3} \\
&\leq \frac{17\exp(1)}{2\delta_3\delta_3 R^2} |f_{l,0}|_{\rho_2, R_2} |G_3|_{\rho_3+\delta_3, R_3\exp(\delta_3)} \\
&\leq \beta_3 \frac{17bc\exp(1)}{2\delta_3\delta_3 R^2} \left(\frac{R}{R_0}\right)^{l+3},
\end{aligned}
$$

but $\mathcal{W}(r,1) = \beta_3$; this, together with the arrangement $\frac{1}{R^2}(\frac{R}{R_0})^{l+3} = \frac{1}{R_0^2}(\frac{R}{R_0})^{l+1}$ and further identification of the constant d, will lead to

$$
|f_{l,1}|_{\rho_3, R_3} \leq \mathcal{W}(r,1) \frac{dc}{\delta_3^2} \left(\frac{R}{R_0}\right)^{l+1},
$$

which agrees with (2.5.12) for $\nu = 1$, for according with (2.5.13), is $\vartheta_1 = d$. Assume thus, the same inequality is verified also for $1 \leq \nu \leq k-1$; therefore, taking the norm $|\cdot|_{\rho_{k+2}, R_{k+2}}$, at both sides of $f_{l,k} = \sum_{j=1}^{k} \frac{j}{k}\{f_{l,k-j}, G_{2+j}\}$ –see remark 2.15–, and realizing that $G_{r+1} = G_{r+2} = \cdots = 0$, one deduces readily,

$$
\begin{aligned}
|f_{l,k}|_{\rho_{k+2},R_{k+2}} &\leq \tfrac{17\exp(1)}{2R^2\delta_3\delta_{k+2}} \sum_{j=1}^{\min\{k,r-2\}} \tfrac{j}{k}\, |f_{l,k-j}|_{\rho_{k-j+2},R_{k-j+2}} |G_{2+j}|_{\rho_{j+2}+\delta_{j+2},R_{j+2}\exp(\delta_{j+2})} \\
&\leq \tfrac{17b\exp(1)}{2R^2\delta_3\delta_{k+2}} \sum_{j=1}^{\min\{k,r-2\}} \tfrac{j}{k}\, a^{j-1}\vartheta_{k-j}\beta_3\cdots\beta_{j+2}\mathcal{W}(r,k-j)\tfrac{c}{\delta_3^{2(k-j)}\delta_3^{2(j-1)}} \left(\tfrac{R}{R_0}\right)^{l+k+2} \\
&\leq \mathcal{W}(r,k)\frac{c}{\delta_3^{2k}} \left(\frac{\delta_3}{\delta_{k+2}} \sum_{j=1}^{\min\{k,r-2\}} \tfrac{j}{k}\tfrac{17b\exp(1)}{2R_0^2} a^{j-1}\vartheta_{k-j} \right) \left(\frac{R}{R_0}\right)^{l+k} \\
&= \mathcal{W}(r,k)\frac{\vartheta_k\, c}{\delta_3^{2k}} \left(\frac{R}{R_0}\right)^{l+k},
\end{aligned}
$$

where to obtain the first inequality we apply: $\delta_{j+2} = \delta_3$ for $j \leq r-2$, and for the third we use that, according with lemma A.11, it must be

$$
\beta_3\beta_4\cdots\beta_{j+2}\mathcal{W}(r,k-j) \leq \mathcal{W}(r,k),
$$

for all $1 \leq j \leq \min\{k, r-2\}$. Finally, identification of the constant d in the sum between parenthesis and the whole sum (times the quotient δ_3/δ_{k+2}) with the term ϑ_k close the proof. $\square$

The lemma above may be completed with lemma A.12 of the appendix A, which provides bounds for the terms ϑ_s and, on the other hand, the bounds (2.5.11) are proved to be valid –for the solutions $G_3, \ldots, G_r$ of the homological equations (1.7.60)–, in proposition 2.28, at the end of this chapter. This yields estimates for the norm of $f_{l,s}$, $l \geq 1$, $s \geq 1$ in the somewhat reduced domain, $\mathcal{D}(7\rho/8, R\exp(-\rho/8))$. Such result constitutes the matter of the next proposition.

Proposition 2.16. *Under the same hypothesis of lemma 2.13, the following estimates on $f_{l,s}$ apply,*

$$
|f_{l,s}|_{7\rho/8,R\exp(-\rho/8)} \leq \mathcal{W}(r,s)\frac{cdg^{s-1}}{\delta_r^{2s}} \left(\frac{R}{R_0}\right)^{l+s}, \tag{2.5.15}
$$

where,

$$
g := \max\{1, \mathrm{e}d + 2a\mathrm{e}\}, \tag{2.5.16}
$$

and a subscript r has been added to remark the dependence of δ on r,

$$
\delta_r = \frac{\rho/16}{r-1}.
$$

These estimates will be the keystone to deal with the bounding of the remainder.

2.6 Bounds for the remainder of the transformation

Let us precise what we mean when talking about the "bounding of the remainder". Up to now, we have been considering the formal transformation $T_G f = F_1 + F_2 + F_3 + \cdots + F_r + \ldots$.

This is an infinite process, which in practical computations is carried out only up to some finite order (degree) r –i. e., the last term computed is F_r–. So it is natural to define, the *remainder of order* r of the transformation (or of the transformed function) as,

$$\mathfrak{R}^{(r)} = F_{r+1} + F_{r+2} + F_{r+3} + \dots \tag{2.6.1}$$

In the previous section it has been justified that one can consider the functions $f_{l,s}$ and the terms $F_1, F_2, \dots, F_r, \dots$ defined in the domain, $\mathcal{D}(7\rho/8, R\exp(-\rho/8))$. To avoid an overload of notation we introduce the following convention for the norm,

$$|||\cdot||| = |\cdot|_{7\rho/8, R\exp(-\rho/8)}. \tag{2.6.2}$$

This will not cause any confusion, because the domain will no longer be reduced along this section. Keeping in mind the definition of the Giorgilli-Galgani algorithm, definition 1.7, one realizes easily that,

$$\begin{aligned}
|||F_{r+1}||| \le{}& |||f_{r+1,0}||| + |||f_{r,1}||| + |||f_{r-1,2}||| + \dots \\
&\dots + |||f_{3,r-2}||| + |||f_{2,r-1}||| + |||f_{1,r}|||, \\
|||F_{r+2}||| \le{}& |||f_{r+2,0}||| + |||f_{r+1,1}||| + |||f_{r,2}||| + \dots \\
&\dots + |||f_{4,r-2}||| + |||f_{3,r-1}||| + |||f_{2,r}||| + |||f_{1,r+1}|||, \\
|||F_{r+3}||| \le{}& |||f_{r+3,0}||| + |||f_{r+2,1}||| + |||f_{r+1,2}||| + \dots \\
&\dots + |||f_{5,r-2}||| + |||f_{4,r-1}||| + |||f_{3,r}||| + |||f_{2,r+1}||| + |||f_{1,r+2}|||, \\
&\vdots \\
|||F_{r+k}||| \le{}& |||f_{r+k,0}||| + |||f_{r+k-1,1}||| + |||f_{r+k-2,2}||| + \dots \\
&\dots + |||f_{k+2,r-2}||| + |||f_{k+1,r-1}||| + |||f_{k,r}||| \\
&+ |||f_{k-1,r+1}||| + |||f_{k-2,r+2}||| + \dots + |||f_{1,r+k-1}|||, \\
&\vdots
\end{aligned}$$

Therefore, the sum of the norms $|||F_{r+1}||| + |||F_{r+2}||| + \dots$, may be grouped in two terms, $\mathcal{S}_1$ and $\mathcal{S}_2$, defined as,

$$\begin{aligned}
\mathcal{S}_1 ={}& |||f_{r+1,0}||| + |||f_{r+2,0}||| + |||f_{r+3,0}||| + \dots + |||f_{r+k,0}||| + \dots \\
&+ |||f_{r,1}||| + |||f_{r+1,1}||| + |||f_{r+2,1}||| + \dots + |||f_{r+k-1,1}||| + \dots \\
&+ |||f_{r-1,2}||| + |||f_{r,2}||| + |||f_{r+1,2}||| + \dots + |||f_{r+k-2,2}||| + \dots \\
&+ |||f_{r-2,3}||| + |||f_{r-1,3}||| + |||f_{r,3}||| + \dots + |||f_{r+k-3,3}||| + \dots \\
&+ \dots \\
&+ |||f_{3,r-2}||| + |||f_{4,r-2}||| + |||f_{5,r-2}||| + \dots + |||f_{k+2,r-2}||| + \dots,
\end{aligned} \tag{2.6.3}$$

and,

$$
\begin{aligned}
\mathcal{S}_2 \;=\;& |||f_{2,r-1}||| + |||f_{3,r-1}||| + |||f_{4,r-1}||| + \cdots + |||f_{k,r-1}||| + \dots \\
+\;& |||f_{1,r}||| + |||f_{2,r}||| + |||f_{3,r}||| + \cdots + |||f_{k,r}||| + \dots \\
+\;& |||f_{1,r+1}||| + |||f_{2,r+1}||| + |||f_{3,r+1}||| + \cdots + |||f_{k,r+1}||| + \dots \\
+\;& \dots \\
+\;& |||f_{1,2r-4}||| + |||f_{2,2r-4}||| + |||f_{3,2r-4}||| + \cdots + |||f_{k,2r-4}||| + \dots \\
+\;& \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
+\;& |||f_{1,2r-3}||| + |||f_{2,2r-3}||| + |||f_{3,2r-3}||| + \cdots + |||f_{k,2r-3}||| + \dots \\
+\;& |||f_{1,2r-2}||| + |||f_{2,2r-2}||| + |||f_{3,2r-2}||| + \cdots + |||f_{k,2r-2}||| + \dots \\
+\;& |||f_{1,2r-1}||| + |||f_{2,2r-1}||| + |||f_{3,2r-1}||| + \cdots + |||f_{k,2r-1}||| + \dots \\
+\;& \dots \\
+\;& |||f_{1,3r-6}||| + |||f_{2,3r-6}||| + |||f_{3,3r-6}||| + \cdots + |||f_{k,3r-6}||| + \dots \\
+\;& \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
+\;& |||f_{1,3r-5}||| + |||f_{2,3r-5}||| + |||f_{3,3r-5}||| + \cdots + |||f_{k,3r-5}||| + \dots \\
+\;& |||f_{1,3r-4}||| + |||f_{2,3r-4}||| + |||f_{3,3r-4}||| + \cdots + |||f_{k,3r-4}||| + \dots \\
+\;& |||f_{1,3r-3}||| + |||f_{2,3r-4}||| + |||f_{3,3r-4}||| + \cdots + |||f_{k,3r-4}||| + \dots \\
+\;& \dots \\
+\;& |||f_{1,4r-8}||| + |||f_{2,4r-8}||| + |||f_{3,4r-8}||| + \cdots + |||f_{k,4r-8}||| + \dots \\
+\;& \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
+\;& |||f_{1,j(r-2)+1}||| + |||f_{2,j(r-2)+1}||| + \cdots + |||f_{k,j(r-2)+1}||| + \dots \\
+\;& |||f_{1,j(r-2)+2}||| + |||f_{2,j(r-2)+2}||| + \cdots + |||f_{k,j(r-2)+2}||| + \dots \\
+\;& |||f_{1,j(r-2)+3}||| + |||f_{2,j(r-2)+3}||| + \cdots + |||f_{k,j(r-2)+3}||| + \dots \\
+\;& \dots \\
+\;& |||f_{1,(j+1)(r-2)}||| + |||f_{2,(j+1)(r-2)}||| + \cdots + |||f_{k,(j+1)(r-2)}||| + \dots \\
+\;& \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}
\tag{2.6.4}
$$

Using the proposition 2.16, we can bound the sums

$$
|||f_{r+1,0}||| + |||f_{r+2,0}||| + \dots, \quad |||f_{r,1}||| + |||f_{r+1,1}||| + \dots, \dots
$$

entering in the r. h. s. of (2.6.3),

$$\sum_{j\geq r+1} |||f_{j,0}||| \leq c_1 \left[\left(\frac{R}{R_0}\right)^{r+1} + \left(\frac{R}{R_0}\right)^{r+2} + \dots \right],$$

$$\sum_{j\geq r} |||f_{j,1}||| \leq c_2\beta_3 \left[\left(\frac{R}{R_0}\right)^{r+1} + \left(\frac{R}{R_0}\right)^{r+2} + \dots \right] \left(\frac{g}{\delta_r^2}\right),$$

$$\sum_{j\geq r-1} |||f_{j,2}||| \leq c_2\beta_3\beta_4 \left[\left(\frac{R}{R_0}\right)^{r+1} + \left(\frac{R}{R_0}\right)^{r+2} + \dots \right] \left(\frac{g}{\delta_r^2}\right)^2,$$

$$\vdots$$

$$\sum_{j\geq 3} |||f_{j,r-2}||| \leq c_2\beta_3\beta_4\cdots\beta_r \left[\left(\frac{R}{R_0}\right)^{r+1} + \left(\frac{R}{R_0}\right)^{r+2} + \dots \right] \left(\frac{g}{\delta_r^2}\right)^{r-2},$$

with the constants c_1 and c_2 given by,

$$c_1 = c, \qquad c_2 = \frac{cd}{g}. \tag{2.6.5}$$

Adding up the terms in all the partial sums above, we get the following bounds for $\mathcal{S}_1$,

$$\begin{aligned}\mathcal{S}_1 \leq{}& c_1 \left(1 - \frac{R}{R_0}\right)^{-1} \left(\frac{R}{R_0}\right)^{r+1} + \\ &+ c_2 \left[\beta_3 \frac{g}{\delta_r^2} + \beta_3\beta_4 \left(\frac{g}{\delta_r^2}\right)^2 + \dots + \beta_3\beta_4\cdots\beta_r \left(\frac{g}{\delta_r^2}\right)^{r-2} \right] \left(1 - \frac{R}{R_0}\right)^{-1} \left(\frac{R}{R_0}\right)^{r+1},\end{aligned}$$

and where the geometric series,

$$\left(\frac{R}{R_0}\right)^{r+1} + \left(\frac{R}{R_0}\right)^{r+2} + \left(\frac{R}{R_0}\right)^{r+3} + \dots = \frac{(R/R_0)^{r+1}}{1 - R/R_0}$$

has been summed ($R/R_0 < 1$). If we introduce ξ and ξ_s as,

$$\begin{aligned}\xi &= \left(\frac{R}{R_0}\right)^{1/2}, \\ \xi_s &= \beta_3\beta_4\cdots\beta_s \left(\frac{g}{\delta_r^2}\xi\right)^{s-2}, \qquad \text{for } 3 \leq s \leq r\end{aligned} \tag{2.6.6}$$

and take into account the nondecreasing character of the sequence $\{\beta_s\}_{3\leq s\leq r}$ (with $\beta_3 \geq 1$), the definition (2.5.5) of δ_r, and that $g \geq 1$ (see proposition 2.16), we can derive an easier estimate for $\mathcal{S}_1$. Explicitly,

$$\mathcal{S}_1 \leq \frac{c_1\xi^{2r+2}}{1-\xi^2} + \frac{c_2(r-2)\xi_r\xi^{r+4}}{1-\xi^2}. \tag{2.6.7}$$

For $\mathcal{S}_2$, one proceeds in the same way: proposition 2.16 allows to bound separately, every one of the sums $|||f_{2,r-1}||| + |||f_{3,r-1}||| + \cdots + |||f_{k,r-1}||| + \ldots$, $|||f_{1,r}||| + |||f_{2,r}||| + \cdots + |||f_{k,r}||| + \ldots$, $|||f_{1,r+1}||| + |||f_{2,r+1}||| + \cdots + |||f_{k,r+1}||| + \ldots$, $\ldots$ in (2.6.4), i. e.,

$$
\begin{aligned}
\textstyle\sum_{k\geq 2} |||f_{k,r-1}||| &\leq c_2\beta_3(\beta_3\cdots\beta_r)\left[\left(\tfrac{R}{R_0}\right)^{r+1} + \left(\tfrac{R}{R_0}\right)^{r+2} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{r-1}, \\
\textstyle\sum_{k\geq 1} |||f_{k,r}||| &\leq c_2\beta_3\beta_4(\beta_3\beta_4\cdots\beta_r)\left[\left(\tfrac{R}{R_0}\right)^{r+1} + \left(\tfrac{R}{R_0}\right)^{r+2} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{r}, \\
\textstyle\sum_{k\geq 1} |||f_{k,r+1}||| &\leq c_2\beta_3\beta_4\beta_5(\beta_3\beta_4\cdots\beta_r)\left[\left(\tfrac{R}{R_0}\right)^{r+2} + \left(\tfrac{R}{R_0}\right)^{r+3} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{r+1}, \ldots \\
\textstyle\sum_{k\geq 1} |||f_{k,2r-4}||| &\leq c_2(\beta_3\beta_4\cdots\beta_r)^2\left[\left(\tfrac{R}{R_0}\right)^{2r-3} + \left(\tfrac{R}{R_0}\right)^{2r-2} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{2r-4},
\end{aligned}
$$

....

$$
\begin{aligned}
\textstyle\sum_{k\geq 1} |||f_{k,2r-3}||| &\leq c_2\beta_3(\beta_3\beta_4\cdots\beta_r)^2\left[\left(\tfrac{R}{R_0}\right)^{2r-2} + \left(\tfrac{R}{R_0}\right)^{2r-1} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{2r-3}, \\
\textstyle\sum_{k\geq 1} |||f_{k,2r-2}||| &\leq c_2\beta_3\beta_4(\beta_3\beta_4\cdots\beta_r)^2\left[\left(\tfrac{R}{R_0}\right)^{2r-1} + \left(\tfrac{R}{R_0}\right)^{2r} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{2r-2}, \\
\textstyle\sum_{k\geq 1} |||f_{k,2r-1}||| &\leq c_2\beta_3\beta_4\beta_5(\beta_3\beta_4\cdots\beta_r)^2\left[\left(\tfrac{R}{R_0}\right)^{2r} + \left(\tfrac{R}{R_0}\right)^{2r+1} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{2r-1}, \ldots \\
\textstyle\sum_{k\geq 1} |||f_{k,3r-6}||| &\leq c_2(\beta_3\beta_4\cdots\beta_r)^3\left[\left(\tfrac{R}{R_0}\right)^{3r-5} + \left(\tfrac{R}{R_0}\right)^{3r-4} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{3r-6},
\end{aligned}
$$

....

$$\vdots$$

$$
\begin{aligned}
\sum_{k\geq 1} |||f_{k,j(r-2)+1}||| &\leq c_2\beta_3(\beta_3\beta_4\cdots\beta_r)^j\times \\
&\quad\times\left[\left(\tfrac{R}{R_0}\right)^{j(r-2)+2} + \left(\tfrac{R}{R_0}\right)^{j(r-2)+3} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{j(r-2)+1}, \\
\textstyle\sum_{k\geq 1} |||f_{k,j(r-2)+2}||| &\leq c_2\beta_3\beta_4(\beta_3\beta_4\cdots\beta_r)^j\times \\
&\quad\times\left[\left(\tfrac{R}{R_0}\right)^{j(r-2)+3} + \left(\tfrac{R}{R_0}\right)^{j(r-2)+4} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{j(r-2)+2}, \\
\sum_{k\geq 1} |||f_{k,j(r-2)+3}||| &\leq c_2\beta_3\beta_4\beta_5(\beta_3\beta_4\cdots\beta_r)^j\times \\
&\quad\times\left[\left(\tfrac{R}{R_0}\right)^{j(r-2)+4} + \left(\tfrac{R}{R_0}\right)^{j(r-2)+5} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{j(r-2)+3}, \ldots \\
\sum_{k\geq 1} |||f_{k,(j+1)(r-2)}||| &\leq c_2(\beta_3\beta_4\cdots\beta_r)^{j+1}\times \\
&\quad\times\left[\left(\tfrac{R}{R_0}\right)^{j(r-2)+4} + \left(\tfrac{R}{R_0}\right)^{j(r-2)+5} + \ldots\right]\left(\tfrac{g}{\delta_r^2}\right)^{(j+1)(r-2)}, \ldots
\end{aligned}
$$

....

$$\vdots$$

and, after addition of the terms on the right and on the left of the $\le$ symbol (using as before, the expression for the sum of the geometric series $(R/R_0)^m + (R/R_0)^{m+1} + \dots$ in the square brackets), we get –in terms of ξ and ξ_s defined by (2.6.6)–,

$$\mathcal{S}_2 \le \frac{c_2\xi^2}{1-\xi^2}\left(\xi_3\xi + \xi_4\xi^2 + \xi_5\xi^3 + \cdots + \xi_r\xi^{r-2}\right) \times \\ \times [\xi_r\xi^{r-2} + (\xi_r\xi^{r-2})^2 + (\xi_r\xi^{r-2})^3 + \cdots + (\xi_r\xi^{r-2})^j + \dots].$$

So, since $|||\mathfrak{R}^{(r)}||| \le \mathcal{S}_1 + \mathcal{S}_2$, we shall take such a sum as a bound of the remainder, i. e.,

$$|||\mathfrak{R}^{(r)}||| \le \frac{c_1\xi^{2r+2}}{1-\xi^2} + c_2\left\{(r-2)\xi^4 + \left(\sum_{j=3}^{r}\xi_j\xi^{j-2}\right)\left[\sum_{j\ge 0}\left(\xi_r\xi^{r-2}\right)^j\right]\right\}\frac{\xi_r\xi^r}{1-\xi^2}. \tag{2.6.8}$$

Therefore, given a fixed order r up to which the normal form is computed, it follows from the last inequality that the transformation $\phi^{G^{(r)}}$ generated by $G^{(r)} = \sum_{s=3}^{r} G_s$ –i. e., such that $f \circ \phi^{G^{(r)}} = T_{G^{(r)}} f$–, will be defined and analytic in some subset of $\mathcal{D}(7\rho/8, R\exp(-\rho/8))$, provided $R < R_0$ is small enough to satisfy,

$$\xi_r\xi^{r-2} < 1. \tag{2.6.9}$$

Conversely, for a sufficiently small $R < R_0$ (hence a *fixed* size of the domain), we may wonder what should be the order of the normal form, r, to minimize $|||\mathfrak{R}^{(r)}|||$, or at least, some appropriate bound of it. Thus, as in Giorgilli et al. (1989), "one can look for the optimal normalization order r_{opt} as a function of R, by minimizing the bound (2.6.8) with respect to r".

2.6.1 The optimal normalization order

Prior to the computations leading to r_{opt}, it is necessary to concrete the analytical expression of the terms β_s, $s = 3, \dots, r$. A suitable choice for them (see proposition 2.28) turns out to be,

$$\beta_s = c_3\left(\frac{c_4}{\rho}r\right)^{\tau(s+1)} \exp\left(12\tau\int_{s-1}^{s} x\ln x\,dx\right), \tag{2.6.10}$$

with the constants,

$$c_3 = \sqrt{2\pi}\mathrm{e}^2\ (>1), \qquad c_4 = \frac{64\tau}{\gamma_1\mathrm{e}}\ (\ge 16). \tag{2.6.11}$$

Another point to remark is that, regardless the more or less accurate selection of the quantities β_s, the r. h. s. of (2.6.8) is –as a function of r and ξ–, too intricate to allow analytic optimization with respect to r. However, one can appreciate that the most significative terms of that bound is the product $\xi_r\xi^r$. Hence, to determine r_{opt} it seems reasonable, for $R > 0$ fixed, to find the minimum of $\xi_r\xi^{r-2}$ as a function of r, and then ask R to be small enough to make $\xi_3\xi^{r-2} < 1$. In view of the r. h. s. of (2.6.8), we see that this optimization assures the analyticity of the transformation, provided the sum $\sum_{j=3}^{r_{opt}}\xi_j\xi^{j-2}$ is bounded.

Actually, to simplify even more the calculations, what is minimized is not $\xi_r\xi^{r-2}$, but the bound purposed by the lemma below.

Lemma 2.17. *The product $\xi_r\xi^{r-1}$ admits the upper bound,*

$$\xi_r\xi^{r-2} \le \exp\left\{(r-2)\ln(c_5\xi) + c_6\int_2^r x\ln x\,dx + c_7\right\}, \tag{2.6.12}$$

valid for $r \ge 3$ and with the constants c_5, c_6, c_7 given by,

$$c_5 := c_3 g, \qquad c_6 := 16\tau + 4\tau\ln\frac{c_4}{\rho}, \qquad c_7 := 4\tau\ln 2 - 2\tau. \tag{2.6.13}$$

Proof. Inequality (2.6.12) is derived straightforward from the definition of ξ_r, the above expression for β_s, $s = 3,\dots,r$, the aid of the (immediate) auxiliary relations,

$$\frac{1}{\delta_r} = \frac{16}{\rho}(r-2) \le \frac{c_4}{\rho}r, \qquad \text{since } c_4 \ge 16; \tag{2.6.14}$$

$$2r - 4 + \tau\sum_{j=4}^{r+1} j \le \tau r^2, \qquad \text{for } r \ge 3; \tag{2.6.15}$$

$$r^2 \le 4\int_2^r x\ln x\,dx, \qquad \text{for } r \ge 3. \tag{2.6.16}$$

and the evaluation of the integral,

$$\int_2^r x\ln x\,dx = \frac{r^2}{2}\ln r - \frac{1}{4}r^2 + 1 - 2\ln 2. \tag{2.6.17}$$

Making the computations explicitly,

$$\begin{aligned}
\xi_r\xi^{r-2} &= \beta_3\beta_4\cdots\beta_r\left(\frac{g}{\delta_r^2}\xi\right)^{r-2}\xi^{r-2}\\
&\le \left(\frac{c_4}{\rho}r\right)^{\tau\sum_{s=3}^r(s+1)}\left(\frac{1}{\delta_r^2}\xi\right)^{r-2}\exp\left(12\tau\sum_{s=3}^r\int_{s-1}^s x\ln x\,dx\right)\times(c_3 g\xi)^{r-2}\\
&\underset{\left\{\frac{1}{\delta_r}\le\frac{c_4}{\rho}r\right\}}{\le} \left(\frac{c_4}{\rho}r\right)^{\tau\sum_{s=3}^r(s+1)+2r-4}\exp\left(12\tau\sum_{s=3}^r\int_{s-1}^s x\ln x\,dx\right)\times(c_5\xi^2)^{r-2}\\
&\underset{(2.6.15)}{\le} \left(\frac{c_4}{\rho}r\right)^{\tau r^2}\exp\left(12\tau\int_2^r x\ln x\,dx\right)\times(c_5\xi^2)^{r-2}\\
&= \exp\left\{(r-2)\ln(c_5\xi^2) + \tau r^2\ln\frac{c_4}{\rho} + \tau r^2\ln r + 12\tau\int_2^r x\ln x\,dx\right\},
\end{aligned} \tag{2.6.18}$$

but, in virtue of the inequality (2.6.16) and the integral (2.6.17), we see that the sum

$$\begin{aligned}
\tau r^2\ln r + \tau r^2\ln\frac{c_4}{\rho} &= 2\tau\int_2^r x\ln x\,dx + \left(\frac{\tau}{2} + \tau\ln\frac{c_4}{\rho}\right)r^2 + 4\tau\ln 2 - 2\tau\\
&\le \left(4\tau + 4\tau\ln\frac{c_4}{\rho}\right)\int_2^r x\ln x\,dx + 4\tau\ln 2 - 2\tau,
\end{aligned}$$

and finally, application of the above inequality to (2.6.18) and further identification of the constants c_5, c_6 and c_7 lead to (2.6.12). $\square$

Let us now consider r a real continuous variable and the function

$$h(r) := \exp\left\{(r-2)\ln(c_5\xi^2) + c_6\int_2^r x\ln x\,dx + c_7\right\}, \qquad (r \geq 3)$$

–i. e., $h(r)$ is defined by the the r. h. s. of (2.6.12)–. We denote by $\check{r}$ the value of r minimizing this function. Then, $\check{r}$ must be a solution of the equation $h'(r) = 0$, which may be expressed as,

$$\mathrm{e}^{\ln r}\ln r = -\frac{1}{c_6}\ln(c_5\xi^2); \tag{2.6.19}$$

and still, letting $w = \ln r$ and $z = -\frac{1}{c_6}\ln(c_5\xi^2)$, it takes the form,

$$w\mathrm{e}^w = z. \tag{2.6.20}$$

Provided $z \geq 0$, this equation has just one solution, as the function $w\mathrm{e}^w$ increases monotonically from 0 to $+\infty$ when w goes from 0 to ∞[6]. The function $W : \mathbb{C} \to \mathbb{C}$ such that $W(z)\mathrm{e}^{W(z)} = z$ is a special function known as the Lambert W function –see Corless et al. (1996), for a concise introduction to the W-logy[7]–. As a complex function $W(z)$ is multivalued, with infinite number of branches denoted by $W_k(z)$, $k \in \mathbb{Z}$. From the division of the complex plane into branches proposed in the paper of Corless et al.:

(i) the branch $W_0(z)$, called the *principal branch*, contains the real axis from $-1/\mathrm{e}$ up to ∞; it has a second-order branch point at $z = -1/\mathrm{e}$ which corresponds to $w = -1$, with branch cut $\{z \in \mathbb{R} : -\infty < z < -1/\mathrm{e}\}$. This branch point is shared with $W_{-1}(z)$, $W_1(z)$.

(ii) $W_{-1}(z)$, $W_1(z)$ each have a double branch cut: $\{z \in \mathbb{R} : -\infty < z < -1/\mathrm{e}\}$ and $\{z \in \mathbb{R} : -\infty < z < 0\}$. By convention, the branch cuts are closed on the top and it turns out that this choice for the closure implies that $W_{-1}(z)$ is real for $z \in [-1/\mathrm{e}, 0)$, so W_0 and $W_{-1}(z)$ are the only branches of $W(z)$ taking real values.

(iii) All other branches $W_k(z)$, $k = \pm 2, \pm 3, \ldots$, have only one branch cut, the one matching the real negative axis. Thus, these branches are similar to those of the logarithm.

In our case z will take real large values, since we consider R small ($R/R_0 \ll 1$); so from the items above, one deduces that the solutions of (2.6.19) we are looking for will be conveniently expressed by the principal branch $W_0(z)$. More precisely,

$$\check{r} = \mathrm{e}^{W_0\left(\ln(c_5\xi^2)^{-1/c_6}\right)}, \tag{2.6.21}$$

and take as the optimal normalizing order, $r_{opt} = \lfloor \check{r} \rfloor$.

[6]It can be seen easily that for $-1/\mathrm{e} < z < 0$, (2.6.20) has two *negative* solutions, both in the interval $(0, -1)$.

[7]W-logy: the science of the Lambert W function.

Remark 2.18. Though it has not been explicitly pointed out, the choice of the principal branch of $W(z)$ carries out an implicit assumption on the smallness of R, as it is necessary to impose that $c_5\xi^2 < 1$, or, equivalently:

$$\frac{R}{R_0} < \frac{1}{c_5},$$

to have $z > 0$ and hence $w > 0$, giving rise to solutions $\check{r} > 1$, as desired. ♠

In what follows, we assume that the nonlinear normalization process has been carried out up to the –in the foregoing explained sense–, optimal order r_{opt}. Then,

$$\xi_{r_{opt}}\xi^{r_{opt}-2} \le \exp\left\{(r_{opt}-2)\ln(c_5\xi^2) + c_6\int_2^{r_{opt}} x\ln x\,dx + c_7\right\}, \tag{2.6.22}$$

and introducing $0 < \chi = \frac{\check{r}-r_{opt}}{\check{r}} < 1$, we have,

$$\begin{aligned} c_6\int_2^{r_{opt}} x\ln x\,dx + c_7 &= \frac{c_6}{2}(1-\chi)r_{opt}\check{r}\ln\check{r} + \frac{c_6}{2}r_{opt}\check{r}(1-\chi)\ln(1-\chi) \\ &\quad - \frac{c_6}{4}r_{opt}^2 + c_6(1-2\ln 2) + c_7 \\ &= -\frac{1}{2}(1-\chi)r_{opt}\ln(c_5\xi^2) + \frac{c_6}{2}r_{opt}\check{r}(1-\chi)\ln(1-\chi) \\ &\quad - \frac{c_6}{4}r_{opt}^2 + c_6(1-2\ln 2) + c_7, \end{aligned}$$

but $\frac{c_6}{2}r_{opt}\check{r}(1-\chi)\ln(1-\chi) < 0$, $-\frac{c_6}{4}r_{opt} < 0$ (we recall that all the constants are positive) and also explicit computation gives[(8)]

$$c_6(1-2\ln 2) + c_7 = (14\tau + 4\tau\ln\frac{c_4}{\rho})(1-2\ln 2) < 0.$$

So according with (2.6.22) $\xi_{r_{opt}}\xi^{r_{opt}-2}$ can be still bounded by,

$$\xi_{r_{opt}}\xi^{r_{opt}-2} \le \exp\left\{\left[\frac{r_{opt}}{2}(1+\chi)-2\right]\ln(c_5\xi^2)\right\} \le \left(c_5\frac{R}{R_0}\right)^{\frac{r_{opt}}{2}-2} < 1. \tag{2.6.23}$$

This last inequality is true, provided $\frac{R}{R_0} \le \frac{1}{c_5}$ (as pointed in remark 2.18), and small enough to make $r_{opt}(R) > 4$. Also, both conditions will from now on assumed.

It is still necessary to check the bounded character[(9)] of $\sum_{j=3}^{r_{opt}}\xi_j\xi^{j-2}$. In fact, its terms decay faster than the corresponding ones in $\sum_{j\ge 0}\frac{1}{j!}$. To show this last point, let us compute, with $3 \le j \le r_{opt}-1$, the quotient

$$\begin{aligned} \frac{\xi_{j+1}\xi^{j-1}}{\xi_j\xi^{j-2}} &= \beta_{j+1}\frac{g}{\delta_{r_{opt}}^2}\xi^2 \\ &\le c_3\left(\frac{c_4}{\rho}r_{opt}\right)^{\tau(j+2)+2}\exp\left(12\tau\int_j^{j+1} x\ln x dx\right)\xi^2. \end{aligned}$$

(8) Here, we suppose $\frac{c_4}{\rho} > 1$ –and hence $\ln\frac{c_4}{\rho} > 0$–, but since $c_4 \ge 16$, this is guaranteed if now (and in the sequel) we take $\rho < 1$.

(9) Note that $\sum_{j=3}^{r_{opt}}\xi_j\xi^{j-2}$ should be bounded independently of r_{opt}, since $r_{opt}(R) \to \infty$ when $R \to 0$.

By explicit cast of the quadrature one derives, after some arrangements:

$$12\tau \int_j^{j+1} x \ln x dx \leq \ln(j+1)^{12\tau j+6\tau} - 3\tau.$$

Now, taking into account:

(i) the value of the constants c_3 and c_6, i. e.,

$$c_3 = \sqrt{2\pi}\mathrm{e}^2 (< \mathrm{e}^3), \qquad c_6 = 16\tau + 4\tau \ln \tfrac{c_4}{\rho},$$

(ii) $j + 1 \leq r_{opt} \leq \lfloor \check{r} \rfloor \leq \check{r}$,

(iii) and that $\check{r}$ is the (unique) solution of the equation (2.6.19), so $\xi^2 = \frac{1}{c_5}\check{r}^{-c_6\check{r}}$,

one has, after some trivial computations:

$$\begin{aligned} \frac{\xi_{j+1}\xi^{j-1}}{\xi_j \xi^{j-2}} &\leq \left(\frac{c_4}{\rho}\right)^{\tau(j+2)+2} \check{r}^{13\tau j + 8\tau + 2}\xi^2 \\ &= \frac{1}{c_5} \exp\left\{ (\tau(j+2) + 2 - 4\tau\check{r}\ln\check{r}) \ln \frac{c_4}{\rho} \right\} \times \exp\left\{ (-16\tau\check{r} + 13\tau j + 8\tau + 2)\ln\check{r} \right\}, \end{aligned} \tag{2.6.24}$$

but again, using item (ii) above, it is straightforward to check the inequalities,

$$\exp\left\{ (\tau(j+2) + 2 - 4\tau\check{r}\ln\check{r}) \ln \frac{c_4}{\rho} \right\} \leq 1,$$

and

$$\exp\left\{ (-16\tau\check{r} + 13\tau j + 8\tau + 2)\ln\check{r} \right\} \leq \exp\{-3\tau j \ln(j+1)\} = (j+1)^{-3\tau j},$$

so, their product times $1/c_5$ in (2.6.24) leads to the bound,

$$\frac{\xi_{j+1}\xi^{j-1}}{\xi_j \xi^{j-2}} = \beta_{j+1} \times \frac{g}{\delta^2_{r_{opt}}} \times \frac{R}{R_0} \leq \frac{1/c_5}{(j+1)^{3\tau j}}, \tag{2.6.25}$$

for all $3 \leq j \leq r_{opt} - 1$; but the r. h. s. of this last inequality cannot be greater that $\frac{1}{(j+1)}$ (we recall that $c_5, \tau > 1$). This proves our assertion on the decay of the terms in $\sum_{j\geq 3}^{r_{opt}} \xi_j \xi^{j-2}$. Using such (i. e., $1/(j+1)$) rough bound for the quotient above, one finds that the sum may be estimated according with

$$\sum_{j=3}^{r_{opt}} \xi_j \xi^{j-2} \leq \left(1 + 3!\, c_5^3 \sum_{j\geq 4} \frac{1}{j!} \left(\frac{1}{c_5}\right)^j \right) \xi_3 \xi \leq 3!\, c_5^3 \mathrm{e}^{\frac{1}{c_5}} \xi_3 \xi, \tag{2.6.26}$$

and it is thus "controlled" (for R small enough), since

$$\begin{aligned} \xi_3 \xi = \beta_3 \frac{g}{\delta^2_{r_{opt}}} \xi^2 &= c_3 \left(\frac{c_4}{\rho} r_{opt}\right)^{4\tau} \frac{g}{\delta^2_{r_{opt}}} \exp\left(12\tau \int_2^3 x \ln x \, dx \right) \frac{R}{R_0} \\ &\leq \check{c}_3 \left(\frac{c_4}{\rho}\right)^{4\tau+2} \check{r}^{4\tau+2} \frac{R}{R_0} = \check{c}_4 \left(\frac{c_4}{\rho}\right)^{4\tau+2} \check{r}^{4\tau+2-c_6\check{r}}, \end{aligned} \tag{2.6.27}$$

where the constants $\breve{c}_3$, $\breve{c}_4$ are independent of R. Here, we have used that $r_{opt} = \lfloor \breve{r} \rfloor \leq \breve{r}$ and that $\breve{r}$ is a solution of the equation $\ln r^r = \ln(c_5 \frac{R}{R_0})^{-1/c_6}$ (see (2.6.19)), so $\frac{R}{R_0} = \frac{1}{c_5} \breve{r}^{-c_6 \breve{r}}$. But $\breve{r}(R)$ tends to infinity when $R \to 0$ and $c_6 > 0$. Then it is clear, from the rightmost term of the expression above that, taking $R < R_0$ small enough one can make $\xi_3 \xi$ smaller than any prefixed constant, for example $\xi_3 \xi < 1$.

With all these elements, we can write down an effective bound for the remainder. Before the statement of the corresponding proposition (which summarizes the foregoing discussion), we introduce for the domains, the abbreviations $\mathcal{D}_0$ and $\mathcal{D}_1$ as,

$$\mathcal{D}_0 := \mathcal{D}(\rho_0, R_0), \tag{2.6.28}$$

$$\mathcal{D}_1 := \mathcal{D}(7\rho/8, R\exp(-\rho/8)), \tag{2.6.29}$$

with ρ_0 and R_0 given by (2.5.2).

Proposition 2.19. *Consider the generating function $G^{(r)} = \sum_{s=3}^{r} G_s$, defined in the domain $\mathcal{D}_1$ and such that*

$$\begin{aligned} |||G_3||| &\leq \beta_3 b \left(\frac{R}{R_0}\right)^3, \\ |||G_s||| &\leq \beta_3 \beta_4 \cdots \beta_s \frac{a^{s-3} b}{\delta_r^{2(s-3)}} \left(\frac{R}{R_0}\right)^s, \quad 3 < s \leq r, \end{aligned} \tag{2.6.30}$$

with the terms,

$$\beta_s = c_3 \left(\frac{c_4}{\rho} r\right)^{\tau(s+1)} \exp\left(12\tau \int_{s-1}^{s} x \ln x \, dx\right), \qquad s = 3, \ldots, r \tag{2.6.31}$$

and the constants,

$$c_3 > 1, \qquad c_4 \geq 16. \tag{2.6.32}$$

Furthermore, let f be a complex function defined and analytic in $\mathcal{D}_0$. Then, for $0 < \rho < \rho_0$ and for $0 < R < R_0$ sufficiently small:

(i) The canonical transformation $\phi^{G^{(r)}}$ given by $f \circ \phi^{G^{(r)}} = T_{G^{(r)}} f$, is defined and analytic in the domain $\mathcal{D}_1$.

(ii) Let

$$T_{G^{(r)}} f = F_1 + F_2 + F_3 + \cdots + F_r + \mathfrak{R}^{(r)},$$

be the expansion of the transformed function of f, where

$$\mathfrak{R}^{(r)} = F_{r+1} + F_{r+2} + F_{r+3} + \ldots$$

Then, there exists an "optimal" normalizing order $r_{opt} = \lfloor \breve{r} \rfloor$, depending on R through,

$$\breve{r}(R) = \mathrm{e}^{W_0\left(\ln\left(c_5 \frac{R}{R_0}\right)^{-1/c_6}\right)}, \tag{2.6.33}$$

(where W_0 denotes the principal branch of the Lambert W function), such that,

$$|||\mathfrak{R}^{(r_{opt})}||| \leq c_8 \left(1 - \frac{R}{R_0}\right)^{-1} \left(c_5 \frac{R}{R_0}\right)^{\frac{r_{opt}(R)}{2} - 1}, \tag{2.6.34}$$

and c_5, c_6, c_8 are positive constants which depend upon ρ, τ, γ and on R_0 but not on R.

(iii) *The remainder* $\mathfrak{R}^{(r_{opt})}$ *goes to zero with* R/R_0 *faster than any analytic order in* R/R_0. *More precisely,*

$$\mathfrak{R}^{(r_{opt})} = \mathrm{o}\left(\left(\frac{R}{R_0}\right)^n\right), \qquad (R/R_0 \to 0) \tag{2.6.35}$$

for any given positive integer n.

Proof. The first item and (2.6.33) in (ii) has been already proved. The bound (2.6.34) is proved straightforward from (2.6.8) and the bound (2.6.26). Finally, the assertion of (iii) is derived from

$$\frac{|||\mathfrak{R}^{(r_{opt})}|||}{(R/R_0)^n} \leq c_8 c_5^n \left(1 - \frac{R}{R_0}\right)^{-1} \left(c_5 \frac{R}{R_0}\right)^{\frac{r_{opt}(R)}{2} - n - 1}.$$

Thus, (2.6.35) follows immediately if one knows that $W_0(z)$ tends to (positive) infinity when $z \to +\infty$ (see remark below), so $\lim_{R/R_0 \to 0} r_{opt} = +\infty$ and the r. h. s. of the last inequality –for all arbitrary but fixed $n \in \mathbb{N}$–, goes to 0 as R/R_0 does. □

Remark 2.20. In the semisimple case (i. e., when the eigenvalues of the monodromy matrix around the periodic orbit are pairwise different), it has been proved (see Jorba and Villanueva, 1997a), that the remainder is exponentially small with R, with bounds of type[(10)].

$$|||\mathfrak{R}^{(r_{opt})}||| \leq \mathit{constant} \cdot \exp\left(-\mathit{constant}\left(\frac{1}{R}\right)^{\frac{1}{\tau+1}}\right).$$

In this sense, the results shown in the proposition above are worse, as can immediately be deduced from the behavior of $W_0(z)$ at infinity[(11)]:

$$W(z) = \ln(z) - \ln\ln(z) + \mathrm{O}\left(\frac{\ln\ln z}{\ln z}\right);$$

(see de Bruijn, 1958), for $z \in \mathbb{R}$ with $z \gg 1$. Then, it is clear that, the r. h. s., of (2.6.34) do decay with R slower than exponentially. ♠

2.7 On the bounds for G_s

It should be clear that the theses of proposition 2.19 are tied to the conditions (2.6.30) –derived from lemma 2.13–, with the given expression for the quantities β_s in (2.6.31). In this section, we consider the finite order (degree) generating function, $G^{(r)} = G_3 + \cdots + G_r$

(10) The paper of Jorba and Villanueva discuss seminormal forms are around elliptic lower dimensional invariant tori, so their results can be applied in the particular case of nonresonant elliptic periodic orbits.

(11) In §2.4 of de Bruijn's book, it is shown that, for z real and large enough, the solutions of equation (2.6.20) admit the following (convergent) development:

$$w = \ln z - \ln\ln z + \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} c_{k,m} (\ln\ln z)^{m+1} (\ln z)^{-k-m-1},$$

Corless et al. (1996)–using two points not noted in de Bruijn's proof–, extended this development to give the asymptotics for all nonprincipal branches of $W(z)$ both at (complex) infinity and at zero.

set up in chapter 1. There, from an analytic linearly reduced Hamiltonian, $H : \mathcal{D}_0 \to \mathbb{C}$, $H = H_2 + H_3 + \cdots$, with

$$H_2(\boldsymbol{q}, I_1, \boldsymbol{p}) = \omega_1 I_1 + \mathrm{i}\omega_2(q_1p_1 + q_2p_2) + q_2p_1, \tag{2.7.1}$$

$$H_s(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) = \sum_{k\in\mathbb{Z}} \sum_{2l+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1=s} h_{k,l,\boldsymbol{m},\boldsymbol{n}} I_1^l \boldsymbol{q}^{\boldsymbol{m}} \boldsymbol{p}^{\boldsymbol{n}} \exp(\mathrm{i}k\theta_1), \qquad (s \geq 3), \tag{2.7.2}$$

$G^{(r)}$ was formally constructed such that the generated canonical change, $\phi^{G^{(r)}}$ transforms H to give,

$$H \circ \phi^{G^{(r)}}(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) = \sum_{s=2}^{r} Z_s(\boldsymbol{q}, I_1, \boldsymbol{p}) + \mathfrak{R}^{(r)}(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}),$$

(theorem 1.24). The main target of the present section, is just to discuss whether proposition 2.19 can be applied to such $G^{(r)}$. If so, this will guarantee the analyticity of the reduced Hamiltonian in some domain $\mathcal{D}_2 \subset \mathcal{D}_1$, for R small enough and justify that, at least locally, the remainder $\mathcal{R}^{(r)}$ can be dealt as a small perturbation of the normal form $Z^{(r)} = \sum_{s=2}^{r} Z_s$. Hence we must prove that the hypotheses on the size of G_s, $s = 3, 4, \ldots$, asked in proposition 2.19 hold.

In view of the equation (2.4.22), for the bounds of the solutions of the homological equations, we realize that for the purpose described above, it is worth finding out bounds of the terms F_s, but these are given in a recursive way by the formulas (1.7.61) in prop 1.21.

Remark 2.21. We shall work assuming for the Hamiltonian H the same conditions than for the function f in section 2.5.1. In particular, $H = \sum_{s\geq 2} H_s$ is defined and analytic in $\mathcal{D}(\rho^*, R^*)$; $|H|_{\rho^*,R^*} = c$ and hence,

$$|H_s|_{\rho,R} \leq c \left(\frac{R}{R_0}\right)^s,$$

where $0 < \rho < \rho_0$, $0 < R < R_0$, with ρ_0, R_0 satisfying (2.5.2). ♠

Moreover by (2.4.10), the terms $Z_s(\boldsymbol{q}, I_1, \boldsymbol{p})$, $3 \leq s \leq r$ can be bounded in $\mathcal{D}(\rho_{s-1}, R_{s-1})$, as:

$$|Z_s|_{\rho_{s-1},R_{s-1}} \leq 2^{s/2} |F_s|_{\rho_{s-1},R_{s-1}}. \tag{2.7.3}$$

Proposition 2.22. *Let δ be given by (2.5.5) and define, for $2 \leq k \leq r$,*

$$\rho_2 := \rho, \qquad \rho_k := \rho_2 - 2(k-2)\delta, \tag{2.7.4}$$

$$R_2 := R, \qquad R_k := R_2 \exp\left(-2(k-2)\delta\right). \tag{2.7.5}$$

Furthermore, consider $\{\alpha_j\}_{2\leq j\leq r}$, the non decreasing sequence of positive numbers defined by (2.3.20), if $s \geq 3$ and $\alpha_2 = 1$. The following bounds apply,

$$|H_{2+l,k}|_{\rho_{k+2},R_{k+2}} \leq \theta_{l,k} \frac{\alpha_3 \cdots \alpha_{k+2} c^{k+1}}{\delta^{2k} R_0^{2k}} \left(\frac{R}{R_0}\right)^{2+l+k} \qquad (1 \leq k \leq r-2), \tag{2.7.6}$$

$$|F_{k+2}|_{\rho_{k+1},R_{k+1}} \leq \eta_k \frac{\alpha_3 \cdots \alpha_{k+1} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} \left(\frac{R}{R_0}\right)^{k+2} \qquad (2 \leq k \leq r-2); \tag{2.7.7}$$

with $F_3 = H_3$ and the terms $\eta_1, \ldots, \eta_{r-2}$; $\theta_{l,0}, \theta_{l,1}, \ldots, \theta_{l,r-l-2}$ $(0 \le l \le r-2)$ are defined recursively by the relations

$$\theta_{l,k} = \frac{\gamma_2}{k} \sum_{j=1}^{k} j \eta_j \theta_{l,k-j} \qquad (k \ge 1), \tag{2.7.8}$$

$$\eta_k = \frac{\gamma_2}{k} \sum_{j=1}^{k-1} j 2^{(k-j+2)/2} \eta_j \eta_{k-j} + \frac{1}{k} \sum_{j=1}^{k} j \theta_{j,k-j} \qquad (k \ge 2), \tag{2.7.9}$$

with

$$\gamma_2 = \frac{17\mathrm{e}}{2}, \qquad \eta_1 = 1, \qquad \theta_{l,0} = 1 \quad \textit{(for all } l \ge 0\textit{)}\,, \tag{2.7.10}$$

From the inequality (2.4.22) and the bounds for F_k, $k = 4, \ldots, r$ in the proposition above, the next corollary follows straightforward.

Corollary 2.23. *The terms G_k of order k, $k = 3, \ldots, r$ of the generating functions are bounded according to,*

$$|G_{k+2}|_{\rho_{k+2}+\delta, R_{k+2}\exp(\delta)} \le \eta_k \frac{\alpha_3 \cdots \alpha_{k+2} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} \left(\frac{R}{R_0}\right)^{k+2}, \tag{2.7.11}$$

for $1 \le k \le r-2$.

Proof of proposition 2.22. We use the reduction algorithm described in section 1.7.3. For $k = 1$, and the definition of $H_{l,k}$ we have,

$$|H_{2+l,1}|_{\rho_3,R_3} = |L_{G_3} H_{2+l,0}|_{\rho_3,R_3} = |\{H_{2+l}, G_3\}|_{\rho_3,R_3}.$$

Again, using corollary A.7, with $\delta_\nu = \delta$ for all ν to bound the Poisson bracket throughout.

$$\begin{aligned}
|H_{2+l,1}|_{\rho_3,R_3} &\le \frac{17\mathrm{e}}{2\delta^2 R^2} |H_{2+l}|_{\rho,R} |G_3|_{\rho_3+\delta,R_3\exp(\delta)} \\
&\le \frac{17\mathrm{e}}{2} \frac{\alpha_3 c^2}{\delta^2 R_0^2} \left(\frac{R}{R_0}\right)^{2+l+1} \\
&= \theta_{l,1} \frac{\alpha_3 c^2}{\delta^2 R_0^2} \left(\frac{R}{R_0}\right)^{2+l+1}
\end{aligned}$$

since, $|G_3|_{\rho_3+\delta,R_3\exp(\delta)} = |G_3|_{\rho-\delta,R\exp(-\delta)} \le \alpha_3 |F_3|_{\rho,R} = \alpha_3 |H_3|_{\rho,R} \le \alpha_3 c \left(\frac{R}{R_0}\right)^3$, and by (2.7.8) $\theta_{l,1} = \gamma_2$. In the same way, for $k = 2$,

$$\begin{aligned}
|F_4|_{\rho_3,R_3} &\le \frac{1}{2} |\{Z_3, G_3\}|_{\rho_3,R_3} + \frac{1}{2} |H_{3,1}|_{\rho_3,R_3} + |H_4|_{\rho_3,R_3} \\
&\le \frac{17\mathrm{e}/2}{2\delta^2 R^2} |Z_3|_{\rho_3+2\delta,R_3\exp(2\delta)} |G_3|_{\rho_3+\delta,R_3\exp(\delta)} + \frac{1}{2} |H_{3,1}|_{\rho_3,R_3} + |H_4|_{\rho,R} \\
&\le \frac{17\mathrm{e}/2}{2\delta^2 R^2} 2^{3/2} \alpha_3 |F_3|^2_{\rho,R} + \frac{1}{2} \theta_{1,1} \frac{\alpha_3 c^2}{\delta^2 R_0^2} \left(\frac{R}{R_0}\right)^4 + c \left(\frac{R}{R_0}\right)^4 \\
&\overset{(12)}{\le} \left(\frac{\gamma_2}{2} 2^{3/2} \eta_1 \eta_1 + \frac{1}{2} \theta_{1,1} + 1\right) \frac{\alpha_3 c^2}{\delta^2 R_0^2} \left(\frac{R}{R_0}\right)^4,
\end{aligned}$$

but from (2.7.9), $\eta_2 = \frac{\gamma_2}{2} 2^{3/2}\eta_1\eta_1 + \frac{1}{2}\theta_{1,1} + \theta_{2,0}$, and hence,

$$|F_4|_{\rho_3,R_3} \leq \eta_2 \frac{\alpha_3 c^2}{\delta^2 R_0^2}\left(\frac{R}{R_0}\right)^4.$$

Remark 2.24. We recall here that $Z_s = 0$ for s odd. However, we shall ignore this fact and use the bounds:

$$|Z_s|_{\rho_{s-1},R_{s-1}} \leq 2^{s/2}|F_s|_{\rho_{s-1},R_{s-1}}, \qquad s = 3,\dots,r,$$

throughout (see (2.7.3)). ♠

Similarly, $H_{2+l,2} = \frac{1}{2}\{H_{2+l,1}, G_3\} + \{H_{2+l,0}, G_4\}$ so,

$$\begin{aligned}
|H_{2+l,2}|_{\rho_4,R_4} &\leq \frac{17\mathrm{e}/2}{2R^2\delta^2}|H_{2+l,1}|_{\rho_3,R_3}|G_3|_{\rho_3+\delta,R_3\exp(\delta)} + \frac{17\mathrm{e}}{2R^2\delta^2}|H_{l+2,0}|_{\rho,R}|G_4|_{\rho_4+\delta,R_4\exp(\delta)}\\
&\leq \frac{\gamma_2}{2}\theta_{l,1}\frac{\alpha_3^2 c^3}{\delta^4 R^2 R_0^2}\left(\frac{R}{R_0}\right)^{l+6} + \gamma_2\eta_2\frac{\alpha_3\alpha_4 c^3}{\delta^4 R^2 R_0^2}\left(\frac{R}{R_0}\right)^{l+6}\\
&\leq \frac{\gamma_2}{2}\left(\theta_{l,1}\eta_1 + 2\theta_{l,0}\eta_2\right)\frac{\alpha_3\alpha_4 c^3}{\delta^4 R_0^4}\left(\frac{R}{R_0}\right)^{l+4}\\
&= \theta_{l,2}\frac{\alpha_3\alpha_4 c^3}{\delta^4 R_0^4}\left(\frac{R}{R_0}\right)^{l+4}.
\end{aligned}$$

Assume now that (2.7.6) and (2.7.7) work for all ν, $2 < \nu < k$ (and $k \leq r-2$). Therefore (2.7.11) should be valid also for $2 < \nu < k$. Then, for $\nu = k$ and $1 \leq j \leq k$,

$$\begin{aligned}
|\{Z_{k-j+2}, G_{2+j}\}|_{\rho_{k+1},R_{k+1}} &\leq \frac{17\mathrm{e}}{2R^2\delta^2}|Z_{k+2-j}|_{\rho_{k-j+1},R_{k-j+1}}|G_{2+j}|_{\rho_{2+j}+\delta,R_{2+j}\exp(\delta)}\\
&\leq \frac{17\mathrm{e}}{2R^2\delta^2}2^{(k+2-j)/2}|F_{k-j+2}|_{\rho_{k-j+1},R_{k-j+1}}|G_{2+j}|_{\rho_{2+j}+\delta,R_{2+j}\exp(\delta)}\\
&\leq \frac{17\mathrm{e}}{2R^2\delta^2}2^{(k+2-j)/2}\eta_{k-j}\frac{\alpha_3\cdots\alpha_{k-j+1}c^{k-j}}{\delta^{2(k-j-1)}R_0^{2(k-j-1)}}\left(\frac{R}{R_0}\right)^{k-j+2}\\
&\quad\times \eta_j\frac{\alpha_3\cdots\alpha_{j+2}c^j}{\delta^{2(j-1)}R_0^{2(j-1)}}\left(\frac{R}{R_0}\right)^{j+2}.
\end{aligned}$$

Using the non decreasing character of the sequence $\{\alpha_j\}_{3\leq j\leq r}$, so $\alpha_{i+2} \leq \alpha_{k-j+i+1}$, for all $i = 1,\dots,j$, we have

$$\begin{aligned}
\alpha_3\cdots\alpha_{k-j+1}\alpha_3\cdots\alpha_{j+2} &= \left(\prod_{i=3}^{k-j+1}\alpha_i\right)\left(\prod_{i=1}^{j}\alpha_{i+2}\right)\\
&\leq \left(\prod_{i=3}^{k-j+1}\alpha_i\right)\left(\prod_{i=1}^{j}\alpha_{k-j+i+1}\right)\\
&= \alpha_3\cdots\alpha_{k+1}.
\end{aligned}$$

(12) For the third term we use that, by (2.5.2), $R_0 < c$, $R_0 < 1$ then, $\frac{c^2}{R_0^2} \geq c\frac{c}{R_0} \geq c$; together with the obvious inequality $\frac{\alpha_3}{\delta^2} > 1$.

Thus, the Poisson bracket above can be bounded as,

$$|\{Z_{k-j+2}, G_{2+j}\}|_{\rho_{k+1},R_{k+1}} \leq \frac{17\mathrm{e}}{2} 2^{(k+2-j)/2} \eta_{k-j}\eta_j \frac{\alpha_3 \cdots \alpha_{k+1} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} \left(\frac{R}{R_0}\right)^{k+2}. \tag{2.7.12}$$

Next, using the induction hypotheses, we can estimate the sum

$$\begin{aligned}\sum_{j=1}^{k} \frac{j}{k} |H_{2+j,k-j}|_{\rho_{k+1},R_{k+1}} &\leq \frac{1}{k} \sum_{j=1}^{k} j |H_{2+j,k-j}|_{\rho_{k-j+2},R_{k-j+2}} \\ &\leq \frac{1}{k} \sum_{j=1}^{k-1} j \theta_{j,k-j} \frac{\alpha_3 \cdots \alpha_{k-j+2} c^{k-j+1}}{\delta^{2(k-j)} R_0^{2(k-j)}} \left(\frac{R}{R_0}\right)^{k+2} + |H_{2+k,0}|_{\rho,R},\end{aligned} \tag{2.7.13}$$

and for $|H_{2+k,0}|_{\rho,R}$, it is immediate that

$$\begin{aligned}|H_{2+k,0}|_{\rho,R} &\leq c \left(\frac{R}{R_0}\right)^{k+2} \\ &\leq \frac{\alpha_3 \cdots \alpha_{k+1} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} \left(\frac{R}{R_0}\right)^{k+2},\end{aligned} \tag{2.7.14}$$

where it has been used that,

(1) as $c/R_0 > 1$, with $R_0 < 1$ it is,

$$\frac{c^{k-j+1}}{R_0^{2(k-j)}} = \left(\frac{c}{R_0}\right)^{k-j+1} \frac{1}{R_0^{k-j-1}} \leq \left(\frac{c}{R_0}\right)^{k} \frac{1}{R_0^{k-j-1}} \leq \left(\frac{c}{R_0}\right)^{k} \frac{1}{R_0^{k-2}} = \frac{c^k}{R_0^{2(k-1)}},$$

(2) $\alpha_3 \cdots \alpha_{k-j+2} \leq \alpha_3 \cdots \alpha_{k+1}$,

for all $1 \leq j \leq k$. Putting together (2.7.12), (2.7.13), (2.7.14),

$$\begin{aligned}|F_{k+2}|_{\rho_{k+1},R_{k+1}} &\leq \sum_{j=1}^{k-1} \frac{j}{k} |\{Z_{k+2-j}, G_{2+j}\}|_{\rho_{k+1},R_{k+1}} + \sum_{j=1}^{k} \frac{j}{k} |H_{2+j,k-j}|_{\rho_{k+1},R_{k+1}} \\ &\leq \left(\frac{\gamma_2}{k} \sum_{j=1}^{k-1} j 2^{(k-j+2)/2} \eta_j \eta_{k-j} + \frac{1}{k} \sum_{j=1}^{k} j \theta_{j,k-j}\right) \frac{\alpha_3 \cdots \alpha_{k+1} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} \left(\frac{R}{R_0}\right)^{k+2} \\ &= \eta_k \frac{\alpha_3 \cdots \alpha_{k+1} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} \left(\frac{R}{R_0}\right)^{k+2}.\end{aligned}$$

Analogously, making use of the induction hypothesis and the properties (1), (2) above,

$$
\begin{aligned}
|H_{2+l,k}|_{\rho_{k+2},R_{k+2}} &\leq \sum_{j=1}^{k} \frac{j}{k} |\{G_{2+j}, H_{2+l,k-j}\}|_{\rho_{k+2},R_{k+2}} \\
&\leq \sum_{j=1}^{k} \frac{j}{k} \frac{17\mathrm{e}}{2R^2\delta^2} |G_{2+j}|_{\rho_{j+2}+\delta, R_{j+2}\exp(\delta)} |H_{2+l,k-j}|_{\rho_{k-j+2},R_{k-j+2}} \\
&\leq \sum_{j=1}^{k-1} \frac{j}{k} \frac{17\mathrm{e}}{2R^2\delta^2} \eta_j \frac{\alpha_3 \cdots \alpha_{j+2} c^j}{\delta^{2(j-1)} R_0^{2(j-1)}} \left(\frac{R}{R_0}\right)^{j+2} \times \\
&\quad \times \theta_{l,k-j} \frac{\alpha_3 \cdots \alpha_{k-j+2} c^{k-j+1}}{\delta^{2(k-j)} R_0^{2(k-j)}} \left(\frac{R}{R_0}\right)^{2+l+k-j} \\
&\quad + \frac{17\mathrm{e}}{2R^2\delta^2} \eta_k \frac{\alpha_3 \cdots \alpha_{k+2} c^k}{\delta^{2(k-1)} R_0^{2(k-1)}} c \left(\frac{R}{R_0}\right)^{4+l+k} \\
&\overset{(13)}{\leq} \left(\frac{\gamma_2}{k} \sum_{j=1}^{k} j\theta_{l,k-j}\eta_j\right) \frac{\alpha_3 \cdots \alpha_{k+2} c^{k+1}}{\delta^{2k} R_0^{2k}} \left(\frac{R}{R_0}\right)^{2+l+k} \\
&= \theta_{l,k} \frac{\alpha_3 \cdots \alpha_{k+2} c^{k+1}}{\delta^{2k} R_0^{2k}} \left(\frac{R}{R_0}\right)^{2+l+k} .
\end{aligned}
$$

This ends the induction and closes the proof of the proposition. □

Our purpose is to obtain bounds for the terms of the generating function, G_s, $s = 3, \ldots, r$. Corollary 2.23, just provides estimates of type (2.5.11), (with $\delta_3 = \cdots = \delta_r = \delta$). But therein, a factor η_k defined recursively in (2.7.9) appears. Hence, it is worth obtaining estimates for these coefficients. To this end, one defines the quantities,

$$a_{l,k} = \frac{4}{k} \sum_{j=1}^{k} j b_j a_{l,k-j}, \qquad \text{for } l, k \geq 0 \ , \tag{2.7.15}$$

$$b_k = \frac{4}{k} \sum_{j=1}^{k-1} j b_j b_{k-j} + \frac{1}{k} a_{1,k-1}, \qquad \text{for } k \geq 2 \ , \tag{2.7.16}$$

with $b_1 = 1 = a_{l,0}$, for $l \geq 0$. Some relevant properties of these coefficients are described by the next lemma.

Lemma 2.25. *The coefficients $a_{l,k}$, b_k defined by (2.7.15) and (2.7.16) respectively, satisfy the following properties:*

(i) $b_s = a_{1,s-1}$, *for* $s \geq 1$.

(ii) $a_{1,k} \geq 4\frac{k+1}{k} a_{1,k-1}$, *for all* $k \geq 1$.

(iii) $a_{l,s} = a_{l-1,s}$, *for all* $l \geq 1$ *and* $s \geq 0$.

(13) Using again that $\alpha_3 \cdots \alpha_{k-j+2}\alpha_3 \cdots \alpha_{j+2} \leq \alpha_3 \cdots \alpha_{k-j+2}\alpha_{k-j+3} \cdots \alpha_{k+2}$, since $\alpha_{i+2} \leq \alpha_{k-j+i+2}$, for all $i = 1, \ldots, j$ and for all $2 \leq k \leq r-2$.

(iv) $a_{1,k-1} \geq \sum_{j=2}^{k} j a_{j,k-j}$, *for* $k \geq 1$.

(v) $b_k \leq 4^{k-1} k!$, *for all* $k \geq 2$.

Proof. (i) $b_1 = a_{1,0} = 1$ by definition. Assume it is verified from 2 up to $s-1$. Therefore,

$$a_{1,s-1} = \frac{4}{s-1} \sum_{j=1}^{s-1} j a_{1,j-1} a_{1,s-j-1},$$

and

$$\begin{aligned} b_s &= \frac{4}{s} \sum_{j=1}^{s-1} j a_{1,j-1} a_{1,s-j-1} + \frac{1}{s} a_{1,s-1} \\ &= \frac{s-1}{s} a_{1,s-1} + \frac{1}{s} a_{1,s-1} \\ &= a_{1,s-1}; \end{aligned}$$

since $\frac{4}{s-1} \sum_{j=1}^{s-1} j a_{1,j-1} a_{1,s-j-1} = a_{1,s-1}$, in accordance with (2.7.15) and the induction hypothesis. This proves the first item.

(ii) follows straightforward from the definition of $a_{l,k}$ and from (i). Setting $l = 1$ in (2.7.15),

$$a_{l,k} = \frac{4}{k} \sum_{j=1}^{k} j b_j a_{1,k-j} \geq \frac{4}{k} \left(b_1 a_{1,k-1} + k b_k a_{1,0} \right) = 4 \frac{k+1}{k} a_{1,k-1}$$

(all the quantities b_j and $a_{1,k-j}$ are positive, $b_1 = a_{1,0} = 1$ by definition and $b_k = a_{1,k-1}$ by (i)).

(iii) For $s = 0$ it is satisfied for all $l \geq 1$: $a_{l,0} = a_{l-1,0} = 1$ by definition. If it should work also for all $1 \leq k \leq s-1$ and $l \geq 1$; then, working with (2.7.15),

$$a_{l,s} = \frac{4}{s} \sum_{j=1}^{s} j b_j a_{l,s-j} = \frac{4}{s} \sum_{j=1}^{s} j b_j a_{l-1,s-j} = a_{l-1,s};$$

because, by the induction hypothesis $a_{l,s-1} = a_{l-1,s-1}, a_{l,s-2} = a_{l-1,s-2}, \ldots, a_{l,1} = a_{l-1,1}$ and, trivially, $a_{l,0} = a_{l-1,0}$.

(iv) The inequality works for $k = 2$, since, setting $k = 2$ in (ii),

$$a_{1,1} \geq 4 \frac{1+1}{1} a_{1,0} = 8 \geq \sum_{j=2}^{2} j a_{j,2-j} = 2 a_{2,0} = 2.$$

Assume the inequality is also valid from 2 up to k. Then,

$$a_{1,k-1} + a_{1,k-1} \geq \sum_{j=2}^{k} j a_{j,k-j} + a_{1,k-1},$$

which implies,

$$2 a_{1,k-1} \geq \sum_{j=1}^{k} j a_{j,k-j};$$

and applying now the results of the second and third items:

$$a_{1,k} \geq 4\frac{k+1}{k}a_{1,k-1} \geq 2\frac{k+1}{k}\sum_{j=1}^{k} j a_{j,k-j} \underset{(by\ (\mathrm{iii}))}{=} 2\frac{k+1}{k}\sum_{j=1}^{k} j a_{j+1,k-j}$$
$$= 2\frac{k+1}{k}\sum_{s=2}^{k+1}(s-1)a_{s,k+1-s},$$

(where a shift $s = j+1$ in the summation index has been introduced in the last step). But it is easy to check that $2\frac{k+1}{k}(s-1) \geq s$, when $s \geq 2$. This ends the induction and proves item (iv).

(v) is satisfied trivially for $k=1$, as $b_1 = 1$. Let us suppose it works for k and inquire what happens for $k+1$. By (i), $b_{k+1} = a_{1,k}$ and then, by definition (2.7.15):

$$b_{k+1} = \frac{4}{k}\sum_{j=1}^{k} j b_j a_{1,k-j} = \frac{4}{k}\sum_{j=1}^{k} j b_j b_{k-j+1}$$
$$\leq \frac{4}{k}\sum_{j=1}^{k} j 4^{j-1}4^{k-j} j!(k-j+1)! = 4^k \sum_{j=1}^{k}\frac{j}{k} j!(k-j+1)!,$$

but it happens that the sum in the las term is not greater than $(k+1)!$. Again, we proceed by induction: it is checked for $k=1$:

$$\sum_{j=1}^{1}\frac{jj!}{k}(k-j+1)! = \frac{1\cdot 1!}{1}(1-1+1)! = 1 \leq (1+1)! = 2.$$

Let us assume this works for k and check its validity for $k+1$:

$$\sum_{j=1}^{k+1}\frac{jj!}{k+1}(k-j+2)! = \frac{k}{k+1}\sum_{j=1}^{k}\frac{jj!}{k}(k-j+2)! + (k+1)!$$
$$\leq \sum_{j=1}^{k}\frac{jj!}{k}(k-j+1)!(k-j+2) + (k+1)!$$
$$\leq (k+1)\sum_{j=1}^{k}\frac{jj!}{k}(k-j+1)! + (k+1)!$$
$$\leq (k+1)(k+1)! + (k+1)!$$
$$= (k+2)!,$$

and where, besides the induction hypothesis at the third inequality; we have used the fact that $k-j+2 \leq k+1$ for all $j \geq 1$ (and hence for all summation indices at the second term of the r. h. s. of the expression above). This last argument completes the proof. $\square$

Next, from $a_{l,k}$ and b_k we introduce $\widetilde{\theta}_{l,k}$ and $\widetilde{\eta}_k$ by the definitions:

$$\widetilde{\theta}_{l,k} = \gamma_2^k 2^{k^2} a_{l,k}, \tag{2.7.17}$$
$$\widetilde{\eta}_k = \gamma_2^{k-1} 2^{k^2} b_k, \tag{2.7.18}$$

with $\gamma_2 \geq 1$ (in our case $\gamma_2 = 17\mathrm{e}/2$), and the formulas (2.7.15), (2.7.16) for $a_{l,k}$ and b_k respectively, give rise to recursive inequalities for $\widetilde{\theta}_{l,k}$ and $\widetilde{\eta}_k$. Immediately,

$$\begin{aligned}\widetilde{\theta}_{l,k} &= \gamma_2^k 2^{k^2} a_{l,k} \\ &= \frac{4\gamma_2}{k}\sum_{j=1}^{k} j\gamma_2^{j-1}2^{j^2}b_j\gamma_2^{k-j}2^{(k-j)^2}a_{l,k-j}2^{2j(k-j)} \geq \frac{4\gamma_2}{k}\sum_{j=1}^{k} j\,\widetilde{\eta}_j\widetilde{\theta}_{l,k-j} \geq \frac{\gamma_2}{k}\sum_{j=1}^{k} j\,\widetilde{\eta}_j\widetilde{\theta}_{l,k-j}.\end{aligned}$$

Here, apart from the definitions of $\widetilde{\theta}_{l,k-j}$ and $\widetilde{\eta}_j$, it has been used also that $2^{2j(k-j)} \geq 1$, for all $1 \leq j \leq k$. On the other hand,

$$\begin{aligned}\widetilde{\eta}_k &= \gamma_2^{k-1}2^{k^2}b_k \\ &= \frac{4\gamma_2}{k}\sum_{j=1}^{k-1} j\gamma_2^{j-1}2^{j^2}b_j\gamma_2^{k-j-1}2^{(k-j)^2}b_{k-j}2^{2j(k-j)} + \frac{1}{k}\gamma_2^{k-1}2^{k^2-1}2a_{1,k-1} \\ &\geq \frac{4\gamma_2}{k}\sum_{j=1}^{k-1} j2^{(k-j+2)/2}\widetilde{\eta}_j\widetilde{\eta}_{k-j} + \frac{1}{k}\sum_{j=1}^{k} j\widetilde{\theta}_{j,k-j}.\end{aligned}$$

Remark 2.26. $2a_{1,k-1} \geq \sum_{j=1}^{k} ja_{j,k-j}$ (see the proof of lemma 2.25), and $4j(k-j) \geq k-j+2 \Leftrightarrow (4j-1)(k-j) \geq 2$, for all $1 \leq j \leq k-1$ and for all $k \geq 2$, so $2^{2j(k-j)} \geq 2^{(k-j+2)/2}$, for the same values of j and k. From here:

$$\gamma_2^{k-1}2^{k^2-1}2a_{1,k-1} \geq \sum_{j=1}^{k} j\gamma_2^{k-j}2^{(k-j)^2}a_{j,k-j} = \sum_{j=1}^{j} j\widetilde{\theta}_{j,k-j},$$

since $k^2-1 = (k-j+j)^2-1 = (k-j)^2+j^2+2j(k-j)-1 \geq (k-j)^2$, for all $1 \leq j \leq k$ and for all $k \geq 1$ and $\gamma_2^{k-1} \geq \gamma_2^{k-j}$ (because $\gamma_2 = 17\mathrm{e}/2 \geq 1$ and $k-1 \geq k-j$ for all $1 \leq j \leq k$). ♠

In particular, it is true that

$$\widetilde{\theta}_{l,k} \geq \frac{\gamma_2}{k}\sum_{j=1}^{k} j\widetilde{\eta}_j\widetilde{\theta}_{l,k-j}, \tag{2.7.19}$$

$$\widetilde{\eta}_k \geq \frac{\gamma_2}{k}\sum_{j=1}^{k-1} j2^{(k-j+2)/2}\widetilde{\eta}_j\widetilde{\eta}_{k-j} + \frac{1}{k}\sum_{j=1}^{k} j\widetilde{\theta}_{j,k-j},. \tag{2.7.20}$$

As $\widetilde{\theta}_{l,0} = 1 = \theta_{l,0}$ for all $l \geq 0$, and $\widetilde{\eta}_1 = 2 \geq \eta_1 = 1$, induction shows that $\widetilde{\theta}_{l,k} \geq \theta_{l,k}$, for all $l, k \geq 0$ and $\widetilde{\eta}_k \geq \eta_k$, for all $k \geq 1$. Indeed, let us suppose it is true for $1 \leq \nu \leq k-1$ and for all $l \geq 0$ and show that therefore it is accomplished also for k. Directly form (2.7.20):

$$\begin{aligned}\widetilde{\eta}_k &\geq \frac{\gamma_2}{k}\sum_{j=1}^{k-1} j2^{(k-j+2)/2}\eta_j\eta_{k-j} + \frac{1}{k}\sum_{j=1}^{k} j\theta_{j,k-j} \\ &= \eta_k,\end{aligned}$$

(applying the induction hypothesis term by term), so with the same arguments,

$$\widetilde{\theta}_{l,k} \geq \frac{\gamma_2}{k} \sum_{j=1}^{k} j\eta_j \theta_{l,k-j} = \theta_{l,k}.$$

Hence, $\eta_k \leq \widetilde{\eta}_k = \gamma_2^{k-1} 2^{k^2} b_k$ and using item (v) of lemma 2.25, we state

Lemma 2.27. *For all $k \geq 1$, the terms η_k defined by the recursive formulas of proposition 2.22, are bounded by*

$$\eta_k \leq \gamma_2^{k-1} 2^{k^2} 4^{k-1} k!, \tag{2.7.21}$$

for all $k \geq 1$.

With the above estimates on the coefficients η_k, if we define

$$\begin{aligned} \widetilde{\beta}_2 &= 1, \\ \widetilde{\beta}_k &= 4^{k-2}(k-2)\alpha_k, \qquad k = 3, \ldots, r; \end{aligned} \tag{2.7.22}$$

with α_k given by (2.3.20), then

$$\begin{aligned} \widetilde{\beta}_s &= c_3 4^{s-2}(s-2) \left(\frac{s+1}{\gamma_1}\right)^{s+1} \left(\frac{\tau(s+1)}{\delta \mathrm{e}}\right)^{\tau(s+1)} \\ &\leq c_3 \left(\frac{c_4}{\rho} r\right)^{\tau(s+1)} (s+1)^{(\tau+2)(s+1)}, \end{aligned}$$

and using $(s+1)^{(\tau+2)(s+1)} \leq \exp\left(12\tau \int_{s-1}^{s} x \ln x \, dx\right)$ for $s \geq 3$, we realize that

$$\widetilde{\beta}_s \leq \beta_s, \qquad s = 3, \ldots, r, \tag{2.7.23}$$

with β_s defined by (2.6.31), and the constants

$$c_3 = \sqrt{2\pi}\mathrm{e}^2, \qquad c_4 = \frac{64\tau}{\gamma_1 \mathrm{e}}, \qquad \gamma_1 = \min\{1, \gamma\},$$

(so c_3, c_4 satisfy conditions (2.6.11)). Finally, the above estimates on the coefficients η_k lead us naturally to:

Proposition 2.28. *The terms $G_3, \ldots, G_r$ of the generating function $G = \sum_{s=3}^{r} G_s$ satisfy the bounds of lemma 2.13, i. e.,*

$$\begin{aligned} |G_3|_{\rho_3+\delta_3, R_3 \exp(\delta_3)} &\leq \beta_3 b \left(\frac{R}{R_0}\right)^3, \\ |G_\nu|_{\rho_\nu+\delta_3, R_\nu \exp(\delta_3)} &\leq \beta_3 \beta_4 \cdots \beta_\nu \frac{a^{\nu-3} b}{\delta_3^{2(\nu-3)}} \left(\frac{R}{R_0}\right)^\nu, \quad 3 < \nu \leq r, \end{aligned} \tag{2.7.24}$$

where the quantities $\{\beta_s\}_{2 \leq s \leq r}$ are given by

$$\begin{aligned} \beta_2 &= 1, \\ \beta_s &= c_3 \left(\frac{c_4}{\rho} r\right)^{\tau(s+1)} \exp\left(12\tau \int_{s-1}^{s} x \ln x \, dx\right), \qquad s = 3, \ldots, r, \end{aligned}$$

with

$$c_3 = \sqrt{2\pi}\mathrm{e}^2, \qquad c_4 = \frac{64\tau}{\gamma_1 \mathrm{e}}, \qquad \gamma_1 = \min\{1, \gamma\}, \tag{2.7.25}$$

$$b = \frac{c}{2}, \qquad a = \frac{2c\gamma_2}{R_0^2}. \tag{2.7.26}$$

Proof. For G_3:

$$\begin{aligned}
|G_3|_{\rho_3+\delta_3, R_3\exp(\delta_3)} &\le \alpha_3 c \left(\frac{R}{R_0}\right)^3 \\
&= 2\alpha_3 \frac{c}{2}\left(\frac{R}{R_0}\right)^3 \le 4\alpha_3 \frac{c}{2}\left(\frac{R}{R_0}\right)^3 = \widetilde{\beta}_3 b \left(\frac{R}{R_0}\right)^3 \\
&\le \beta_3 b \left(\frac{R}{R_0}\right)^3 .
\end{aligned}$$

And, for G_ν, $3 < \nu \le r$, using (2.7.11) and (2.7.21),

$$\begin{aligned}
|G_\nu|_{\rho_\nu+\delta_3, R_\nu\exp(\delta_3)} &\le \frac{\alpha_3 \cdots \alpha_\nu c^{\nu-2}}{\delta_3^{2(\nu-3)} R_0^{2(\nu-3)}} \gamma_2^{\nu-3} 4^{\nu-3} (\nu-2)! 2^{(\nu-2)^2} \left(\frac{R}{R_0}\right)^\nu \\
&= \frac{(1\cdot 4^1\alpha_3)(2\cdot 4^2\alpha_4)\cdots((\nu-2)4^{\nu-2}\alpha_\nu)}{\delta_3^{2(\nu-3)}} \frac{c}{2}\left(\frac{2\gamma_2 c}{R_0^2}\right)^{\nu-3}\left(\frac{R}{R_0}\right)^\nu \\
&= \frac{\widetilde{\beta}_3\cdots\widetilde{\beta}_\nu}{\delta_3^{2(\nu-3)}} b a^{\nu-3}\left(\frac{R}{R_0}\right)^\nu \\
&\le \frac{\beta_3\cdots\beta_\nu}{\delta_3^{2(\nu-3)}} b a^{\nu-3}\left(\frac{R}{R_0}\right)^\nu .
\end{aligned}$$

where we have made use of the relations:

$$2^{2\nu-6}2^{(\nu-2)^2} = \tfrac{1}{2}2^{\nu-3}2^{\nu-2}2^{(\nu-2)^2} = \tfrac{1}{2}2^{\nu-3}2^{(\nu-2)(\nu-1)}, \tag{2.7.27}$$

$$4^1 4^2 \cdots 4^{\nu-2} = 4^{1+2+\cdots+\nu-2} = 4^{(\nu-2)(\nu-1)/2} = 2^{(\nu-1)(\nu-2)}, \tag{2.7.28}$$

to pass from the first to the second term on the r. h. s. in the inequality above. □

Therefore, the results of this section justify the use of such bounds for the components G_s of the generating function. The theorem below summarizes the quantitative study of the reduced (or normalized, up to some optimal order r_{opt}) Hamiltonian, i. e.:

$$H(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) = Z^{(r_{opt})}(\boldsymbol{q}, I_1, \boldsymbol{p}) + \mathfrak{R}^{(r_{opt})}(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}), \tag{2.7.29}$$

(see theorem 1.24 in the previous chapter). Where $Z^{(r_{opt})} = \sum_{s=2}^{r_{opt}} Z_s$ was the normal form itself while $\mathfrak{R}^{(r_{opt})}(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p})$ is now the remainder of the reduced Hamiltonian. In other words, now we fix the "generic" function f to be the linear normalized and complexified Hamiltonian (see (1.6.2) in section 1.6 of chapter 1).

Theorem 2.29. *With the hypothesis and assumptions in theorem 1.24, if the normalization of the linear reduced (and complexified) Hamiltonian defined in $\mathcal{D}(\rho_0, R_0)$ is carried out up to the order $r_{opt} = \lfloor \check{r}(R) \rfloor$, with $0 < R < R_0$ small enough and $\check{r}(R)$ given by formula* (2.6.21) *(proposition 2.19). Then, the normalized Hamiltonian* (2.7.29) *is defined in $\mathcal{D}(7\rho/8, R\exp(-\rho/8))$ (with $0 < \rho < \rho_0$) and the following bounds for the Remainder*

$$|||\mathfrak{R}^{(r_{opt})}||| \leq c_8 \left(1 - \frac{R}{R_0}\right)^{-1} \left(c_5 \frac{R}{R_0}\right)^{\frac{r_{opt}(R)}{2} - 1}, \tag{2.7.30}$$

and for the sum $Z_{\geq 6}^{(r_{opt})} \equiv \sum_{s=6}^{r_{opt}} Z_s$,

$$|||Z_{\geq 6}^{(r_{opt})}||| \leq c_9 \frac{R}{R_0}^6, \tag{2.7.31}$$

hold for $0 < R < R_0$ small enough. Here c_5, c_8, c_9 are constants depending on ρ, τ, γ, R_0 but not on R and $|||\cdot|||$ is the norm stated by (2.6.2)*, i. e.: $|||\cdot||| \equiv |\cdot|_{7\rho/8, R\exp(-\rho/8)}$.*

Proof. All this theorem, but the last bound (2.7.31) for the sum $Z_6 + \cdots + Z_{r_{opt}}$, can be derived from the proposition 2.19, letting $f = H$, the linear normalized and complexified Hamiltonian (1.6.2). To derive (2.7.31), we shall consider the estimate:

$$|||Z_6||| \leq \check{c}_6 \left(\frac{R}{R_0}\right)^6, \tag{2.7.32}$$

with $\check{c}_6$ independent on R. Here, we recall that Z_6 is an homogeneous polynomial of degree six (we mean, adapted degree, as defined by (1.7.3) in section 1.7 of chapter 1) and the previous terms in the normal form do not change when this is carried out up to some higher order). Moreover, using (2.7.3) and (2.7.7) together with the inequality (2.7.32) given above, it is clear that the sum $Z_{\geq 6}^{(r_{opt})}$ can be bounded as,

$$|||Z_{\geq 6}^{(r_{opt})}||| \leq \left\{ \check{c}_6 + \sum_{j=7}^{r_{opt}-1} \check{Z}_j \right\} \left(\frac{R}{R_0}\right)^6 \tag{2.7.33}$$

being

$$\check{Z}_j = 2^{\frac{j+1}{2}} c^{j-1} \widetilde{\eta}_{j-1} \frac{\alpha_3 \alpha_4 \ldots \alpha_j}{\delta_{r_{opt}}^{2(j-2)} R_0^{2(j-2)}} \left(\frac{R}{R_0}\right)^{j-5}, \qquad j = 7, \ldots, r_{opt-1},$$

with $\widetilde{\eta}_j = \gamma_2^{j-1} 2^{j^2} 4^{j-1} j! \geq \eta_j$ (see lemma 2.27). But the ratio between any term in and its previous one in the sum on the right hand side are computed to be:

$$\frac{\check{Z}_{j+1}}{\check{Z}_j} = 8\sqrt{2} c \gamma_2 \times \frac{j}{j-1} \times (j-1)\, 4^{j-1} \alpha_{j+1} \times \frac{1}{\delta_{r_{opt}}^2 R_0^2} \times \frac{R}{R_0}, \qquad \text{for } j = 7, \ldots, r_{opt} - 1.$$

But $(j-1)\, 4^{j-1} \alpha_{j+1} = \widetilde{\beta}_{j+1}$ (see (2.7.22)) and according with (2.7.23), it is $\widetilde{\beta}_s \leq \beta_s$, for $s = 3, \ldots, r_{opt}$. Therefore the quotient above may bounded as,

$$\frac{\check{Z}_{j+1}}{\check{Z}_j} \leq 16\sqrt{2} \times \frac{c\gamma_2}{g R_0^2} \times \beta_{j+1} \frac{g}{\delta_{r_{opt}}^2} \times \frac{R}{R_0} \leq \frac{4\sqrt{2}}{\mathrm{e}} \times \beta_{j+1} \frac{g}{\delta_{r_{opt}}^2} \times \frac{R}{R_0},$$
$$\text{for } j = 7, \ldots, r_{opt} - 2, \tag{2.7.34}$$

where, in the last step, we have used that by its definition –equation (2.5.16)– in proposition 2.16, is $g := \max\{1, \mathrm{e}d + 2a\mathrm{e}\}$, so $g \geq 2a\mathrm{e} = \frac{4c\mathrm{e}\gamma_2}{R_0^2}$ (we recall that $a = \frac{2c\gamma_2}{R_0^2}$ is one of the coefficients exacted in proposition 2.28, see (2.7.26)). Now, by comparison of the last term in the r. h. s. of (2.7.34) with (2.6.25) one may see that the following bounds for the quotients

$$\frac{\breve{Z}_{j+1}}{\breve{Z}_j} \leq \frac{4/c_5}{(j+1)^{3\tau j}}, \qquad 7 \leq j \leq r_{opt} - 1$$

hold if $0 < R < R_0$ is sufficiently small. In particular, as c_5 is independent of R, this will imply that the sum $\sum_j^{r_{opt}-1} \breve{Z}_j$, appearing between the braces on the r. h. s. in (2.7.33), may be bounded by a constant also independent of R if

$$\breve{Z}_7 = 2^4 c^6\, \widetilde{\eta}_7 \frac{\alpha_3 \cdots \alpha_7}{\delta_{r_{opt}}^2 R_0^{10}} \times \frac{R}{R_0}$$

does. But, with the same arguments than those employed to "control" $\xi_3\xi$ in section 2.6.1 (see (2.6.27)), one may check that, for $R < R_0$ small enough, $\breve{Z}_7$ can be made smaller than any prefixed positive constant. Hence, in (2.7.33),

$$\breve{c}_6 + \sum_{j=7}^{r_{opt}-1} \breve{Z}_j \leq c_9$$

with $c_9 > 0$ independent of R, for $0 < R < R_0$ small enough. This ends the proof of the last statement and thus of the theorem. □

Chapter 3

Persistence of the 2D-family of bifurcated invariant tori

3.1 Overview of the chapter

This chapter is devoted to the discussion of the persistence of the bifurcated 2-D invariant tori studied in section 1.9. There are different possible situations where persistence can be investigated. Here, we shall focus on the case of direct bifurcation. Therefore, throughout the chapter $a > 0$ and preservation of bifurcated normal *elliptic* tori will be investigated. After the setting of the problem (section 3.2), we devise a new parametrization of the unperturbed (i. e., those coming from the "formal" normal form) tori. In particular the new parameters $\boldsymbol{\zeta}^* = (\xi, \eta)$ are introduced in such a way that, if $\xi > 0$ the corresponding tori are real, and complex if $\xi < 0$ (see theorem 3.1).

The first difficulty one has to face in this problem is treated in section 3.2.2 and has to do with the choice of the suitable set of parameters to characterize the tori of the family along the iterative KAM process. Let us mention that we have three frequencies to control: the two intrinsic ones, Ω_1, Ω_2 of the quasi-periodic motion and the normal one (the positive imaginary part of λ_+), but just two parameters to control them ξ, η. So, we have to deal with the so called "lack of parameters" problem (see Moser, 1967; Sevryuk, 1999).

When applying applying KAM techniques for low-dimensional tori, one typically sets a diffeomorphism between some neighborhood of the origin in the parameter space (ξ, η) and a vicinity of (ω_1, ω_2) in the space of intrinsic frequencies (Ω_1, Ω_2). Hence, the characteristic exponents $\lambda_\pm$ may be put also as a function of the intrinsic frequencies. For elliptic tori, besides the non-degeneracy of these frequencies, one needs to ask the normal frequencies to "move" as a function of $\boldsymbol{\Omega}$, this forces to impose suitable "transversal" conditions in the denominators of the KAM process (see Sevryuk, 1999; Jorba and Villanueva, 1997a). In our case, for $\boldsymbol{\zeta}^* = (\xi, \eta)$ in a small neighborhood of the origin the invariant tori will be elliptic when $a > 0$ (and $\xi > 0$), as follows easily from the expression for $\lambda_\pm$ (see (3.2.19)). However, the typical transversal condition (3.2.21) does not work, since the derivatives of $\lambda_{pm}(\boldsymbol{\Omega})$ are not defined for $\boldsymbol{\Omega} = \boldsymbol{\omega}$ (the elliptic invariant tori are too close to parabolic). This pit is overcome taking as *basic frequencies* for the unperturbed tori not $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$, the intrinsic frequencies, but $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$ with $\mu = |\lambda_+|$ and then the first component of the intrinsic frequencies, Ω_1, as a function of $\boldsymbol{\Lambda}$, i. e.: $\Omega_1 = \Omega_1(\mathbf{\Lambda})$.

In other words: we "label" the (elliptic) invariant tori with their normal frequency and second intrinsic frequency. It is checked that, with this parametrization (shown explicitly in lemmas 3.7 and 3.8), the small divisors do change in the normal directions, so one can proceed with to the KAM iterative scheme, which –due to the forementioned proximity of parabolic tori–, involves a more tricky control on the different terms of the Hamiltonian appearing at each successive step.

The main result of the chapter, theorem 3.9 is proved along the rest of the chapter. We follow those ideas in Jorba and Villanueva (1997a). See section 3.2.3 for the details on the methodology, whereas the iterative KAM process is described in sections 3.3 and 3.4. Nevertheless, the control of the "regularity" with respect to $\boldsymbol{\Lambda}$ of the successive transformations (which is important to estimate the measure of the "good" and "bad" basic frequencies) is carried out here, not through the Lipschitz dependence (see also Jorba and Villanueva, 1997b), but using the original idea in Arnol'd (1963a,b) so we consider analytic dependence with respect to $\boldsymbol{\Lambda}$. This forces us to consider a KAM process with an ultra-violet cut-off (see section 3.3 for details). We point out that, in spite of this analytic dependence, the limit Hamiltonian will be defined in a Cantor set of the basic frequencies $\boldsymbol{\Lambda}$ and so the regularity is no longer analytic; in fact, is C^∞ in the sense of Whitney (see Broer, Huitema and Sevryuk, 1996, for an account on Whitney-regularity). Estimates on the measure of the complementary of the Cantor set of good frequencies are derived in section 3.5, while an outline on the Whitney-smoothness is exposed in section 3.5.2.

3.2 Setting of the problem

After the nonlinear reduction process described along chapter 1, the initial Hamiltonian is transformed through a symplectic change into a new complex Hamiltonian (see theorem 1.24),

$$\begin{aligned} H(\theta_1, \mathbf{q}, I_1, \mathbf{p}) = \omega_1 I_1 + \mathrm{i}\omega_2 \left(q_1 p_1 + q_2 p_2\right) + q_2 p_1 + \mathcal{Z}_{r_{opt}}(-q_1 p_2, I_1, \mathrm{i}(q_1 p_1 + q_2 p_2)) \\ + \mathfrak{R}^{(r_{opt})}(\theta_1, q_1, q_2, I_1, p_1, p_2). \end{aligned} \tag{3.2.1}$$

This Hamiltonian is defined in the complex domain $\mathcal{D}(7\rho/8, R\exp(-\rho/8))$, with $0 < R < 1$ and $0 < \rho < 1$. However, substitutions $7\rho/8 \mapsto \rho$, $R\exp(-\rho/8) \mapsto R$ will only change the constants c_5, c_6 in the expression (2.6.33) for $\check{r}$ and the bound (2.6.34) for $\mathfrak{R}^{(r_{opt})}$ (also, the term $(1 - R/R_0)^{-1}$ should be substituted by $(1 - \mathrm{e}^{\rho/8} R/R_0)^{-1}$ there). Moreover, this re-scaling will modify, by a factor $\mathrm{e}^{6\rho/8}$, the constant c_9 in the bound (2.7.31). See proposition 2.19 and theorem 2.29 in the previous chapter. Therefore, in the sequel, we shall consider (3.2.1) defined in $\mathcal{D}(\rho, R)$, for $0 < R < R_0$ small enough and $0 < \rho < \rho_0$.

In particular, in (3.2.1), $\mathcal{Z}_{r_{opt}}(u, v, w)$ is a polynomial of degree $\lfloor r_{opt}/2 \rfloor$ with real coefficients, beginning with quadratic terms and such that:

$$\mathcal{Z}_{r_{opt}}(-q_1 p_2, I_1, \mathrm{i}(q_1 p_1 + q_2 p_2)) = \sum_{s \geq 3}^{r_{opt}} Z_s(q_1, q_2, I_1, p_1, p_2), \quad (\text{with } Z_s \equiv 0, \text{ if } s \text{ is odd}).$$

The order $r_{opt} = \lfloor \check{r} \rfloor$, with $\check{r} = \check{r}(R)$ (see (2.6.33)) is taken, for a fixed R small enough, to give the *optimal* order of the normal form in the sense discussed in the last chapter (theorem 2.29). We shall use for $\mathcal{Z}_{r_{opt}}$ the expansion (1.8.2), keeping the same names

$a, b, c, \ldots$ for its coefficients. $\mathfrak{R}^{(r_{opt})}$ is the remainder of the normal form, and bounds of its size have just been obtained as a function of R (see (2.7.30) in theorem 2.29). Therefore, if R is small enough, $\mathfrak{R}^{(r_{opt})}$ can be thought of as a perturbation of the normal form

$$Z^{(r)} = \omega_1 I_1 + \mathrm{i}\omega_2(q_1p_1 + q_2p_2) + q_2p_1 + \mathcal{Z}_r(-q_1p_2, I_1, \mathrm{i}(q_1p_1 + q_2p_2)),$$

which was proved to be integrable (see section 1.8). Furthermore in section 1.9, chapter 1 we found a two-parameter family of invariant tori of $Z^{(r)}$. To show them up, though, it was convenient the introduction of a new angle θ_2 and its conjugate action I_2, through the symplectic transformation

$$\begin{aligned} q_1 &= \sqrt{q}\mathrm{e}^{\mathrm{i}\theta_2}, & p_1 &= \sqrt{q}\left(p - \mathrm{i}\frac{I_2}{2q}\right)\mathrm{e}^{-\mathrm{i}\theta_2}, \\ q_2 &= \sqrt{q}\left(p + \mathrm{i}\frac{I_2}{2q}\right)\mathrm{e}^{\mathrm{i}\theta_2}, & p_2 &= -\sqrt{q}\mathrm{e}^{-\mathrm{i}\theta_2}, \end{aligned} \tag{3.2.2}$$

this is the complex counterpart of the real change given by (1.9.1), (see (1.6.1) for the relations between real and complex coordinates). We shall take the new canonical variables, $(\boldsymbol{\theta}, q, \boldsymbol{I}, p) \in \mathbb{C}^2 \times \mathbb{C} \times \mathbb{C}^2 \times \mathbb{C}$, in a domain $\mathfrak{D}(M, \rho, R)$ defined by the inequalities:

$$|\mathrm{Im}\,\boldsymbol{\theta}| \le \min\{\rho, \ln 2\}, \quad M \le |q| \le R^2/4, \quad |I_1| \le R^2, \quad \left|\frac{I_2}{2\sqrt{q}}\right| \le R/4, \quad |p| \le \frac{1}{2}, \tag{3.2.3}$$

(M a positive quantity to be adjusted later). $\mathfrak{D}(M, \rho, R)$ has been specified in such a way that if $(\boldsymbol{\theta}, q, \boldsymbol{I}, p) \in \mathfrak{D}(M, \rho, R)$, then the coordinates $(\theta_1, \boldsymbol{q}, I_1, \boldsymbol{p}) \in \mathcal{D}(\rho, R)$. So, with the change above, Hamiltonian (3.2.1) transforms to (see (1.9.2)):

$$\mathcal{H}(\boldsymbol{\theta}, q, \boldsymbol{I}, p) = \langle \boldsymbol{\omega}, \boldsymbol{I} \rangle + qp^2 + \frac{I_2^2}{4q} + \mathcal{Z}(q, \boldsymbol{I}) + \mathcal{R}(\boldsymbol{\theta}, q, \boldsymbol{I}, p), \tag{3.2.4}$$

where $\boldsymbol{\omega}^* = (\omega_1, \omega_2)$ (we remark that this Hamiltonian is real when all the coordinates $\boldsymbol{\theta}$, q, $\boldsymbol{I}$ and p are). Here, the subscript (and superscript) r has been dropped from $\mathcal{Z}$ (and $\mathcal{R}$), and will no longer be written explicitly. In what follows, we shall refer the part of $\mathcal{H}$ which comes from the normal form (and consequently does not depend on the angles, $\boldsymbol{\theta}$), as the *unperturbed* Hamiltonian, $\mathcal{H}_0$, so

$$\mathcal{H}_0(q, \boldsymbol{I}, p) := \mathcal{H}(\boldsymbol{\theta}, q, \boldsymbol{I}, p) - \mathcal{R}(\boldsymbol{\theta}, q, \boldsymbol{I}, p). \tag{3.2.5}$$

On the other hand, the polynomial $\mathcal{Z}$ can be split as the sum $\mathcal{Z} = \mathcal{Z}_2 + \mathcal{Z}_3$, where $\mathcal{Z}_2$ contains the monomials of $\mathcal{Z}$ of degree 2, whereas $\mathcal{Z}_3$ holds the rest of the terms, i. e.,

$$\mathcal{Z}_2(q, \boldsymbol{I}) = \frac{1}{2}(aq^2 + bI_1^2 + cI_2^2) + dqI_1 + eqI_2 + fI_1I_2, \tag{3.2.6}$$

$$\mathcal{Z}_3(q, \boldsymbol{I}) = \sum_{3 \le i+j+k \le \lfloor r/2 \rfloor} f_{i,j,k} q^i I_1^j I_2^k, \tag{3.2.7}$$

where all the coefficients are real. Furthermore, it is

$$\mathcal{Z}_3(q, I_1, I_2) = Z_{\ge 6}^{(r_{opt})}(q_1, q_2, I_1, p_1, p_2),$$

where $Z_{\geq 6}^{(r_{opt})}$ is the sum:

$$Z_{\geq 6}^{(r_{opt})} = \sum_{s=6}^{r_{opt}} Z_s, \tag{3.2.8}$$

(see theorem 2.29). Then, using the auxiliary norm $||\cdot||_{\widetilde{R}}$, introduced in definition 2.5 and from the relations between the coefficients of the polynomials $Z_{\geq 6}^{(r_{opt})}$ and $\mathcal{Z}_3$, it can be checked that,

$$\|\mathcal{Z}_3\|_{R^2} \leq \|Z_{\geq 6}^{(r_{opt})}\|_{R^2} = |Z_{\geq 6}^{(r_{opt})}|_{0,R}.$$

Then, in view of the bounds for $Z_{\geq 6}^{(r_{opt})}$ derived in theorem 2.29 and further application of the Cauchy's inequalities (as in lemma A.2) it follows that there exists constants $\widehat{\mathcal{Z}}_3$, $\widehat{\mathcal{Z}}_{3;1}$, $\widehat{\mathcal{Z}}_{3;2}$ and $\widehat{\mathcal{Z}}_{3;3}$, independent of R (and of r), such that for any R small enough,

$$\|\mathcal{Z}_3\|_{R^2} \leq \widehat{\mathcal{Z}}_3 R^6, \qquad \|\partial_i \mathcal{Z}_3\|_{R^2/2} \leq \widehat{\mathcal{Z}}_{3;1} R^4, \tag{3.2.9}$$

$$\|\partial^2_{i,j} \mathcal{Z}_3\|_{R^2/2} \leq \widehat{\mathcal{Z}}_{3;2} R^2, \qquad \|\partial^3_{i,j,k} \mathcal{Z}_3\|_{R^2/2} \leq \widehat{\mathcal{Z}}_{3;3}, \quad \text{with } i,j,k = 1,2,3. \tag{3.2.10}$$

3.2.1 An alternative parametrization

Theorem 1.27 gave a real parametrization for the above referred family of invariant tori of $\mathcal{H}_0$ (the unperturbed Hamiltonian in the new coordinates). However, for the purposes of the present chapter (more precisely, to control the real character of the invariant tori), it should be convenient to introduce a new parametrization. From the Hamiltonian equations of $\mathcal{H}_0$ (see (1.9.3)), we can state a result similar to that of theorem 1.27.

Theorem 3.1. *If the coefficient d in* (3.2.6) *is $d \neq 0$, then there exists an analytic function $\mathcal{I} : \Gamma \subset \mathbb{R}^2 \to \mathbb{R}$, with Γ a neighborhood of* $(0,0)$, *defined implicitly by the equation*

$$\eta^2 = \partial_1 \mathcal{Z}_r \left(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta\right), \qquad \textit{with} \;\; \mathcal{I}(0,0) = 0, \tag{3.2.11}$$

and such that, for $(\xi,\eta) \in \Gamma$, the two-dimensional torus

$$\mathcal{T}_{\xi,\eta} = \{(\boldsymbol{\theta}, q, \boldsymbol{I}, p) \in \mathbb{T}^2 \times \mathbb{C} \times \mathbb{C}^2 \times \mathbb{C} : \\ q = \xi, \;\; I_1 = \mathcal{I}(\xi,\eta), \;\; I_2 = 2\xi\eta, \;\; p = 0\} \tag{3.2.12}$$

is invariant under the flow of (1.9.3) *with parallel dynamics (see definition 1.26, in chapter 2) determined by the vector $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$ of intrinsic frequencies:*

$$\begin{aligned} \Omega_1(\xi,\eta) &= \omega_1 + \partial_2 \mathcal{Z}_r(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta), \\ \Omega_2(\xi,\eta) &= \omega_2 + \eta + \partial_3 \mathcal{Z}_r(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta); \end{aligned} \tag{3.2.13}$$

Moreover, for $\xi > 0$, the corresponding tori are real.

Proof. It follows directly by substitution in the Hamiltonian differential equations. Here we only stress that the condition $d \neq 0$ is the necessary hypothesis for the implicit function $\mathcal{I}$ to exist in a neighborhood of $(0,0)$, since $\partial^2_{1,2} \mathcal{Z}_r(0,0,0) = d$. $\square$

Hence, $\boldsymbol{\zeta}^* = (\xi, \eta)$ are the parameters of the family of tori, so they "label" the concrete invariant torus of $\mathcal{H}_0$ filled up by quasi-periodic solutions. Furthermore, we have introduced the vector $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$ of *basic* or *intrinsic* frequencies which changes along the family according to (3.2.13). In fact,

$$\Omega_i(\boldsymbol{\zeta}) = \partial_{I_i}\mathcal{H}_0|_{\mathcal{T}_{\boldsymbol{\zeta}}}, \qquad i = 1, 2,$$

with the following notation: for a given function $g : \mathbb{C}^6 \to \mathbb{C}$, analytic and 2π-periodic with respect to its two first arguments, we shall denote by $g|_{\mathcal{T}_{\boldsymbol{\zeta}}}$ the restriction of g to $\mathcal{T}_{\boldsymbol{\zeta}}$, the invariant torus of the family corresponding to the parameter $\boldsymbol{\zeta}$. On the other hand, it is straightforward to derive formally the expansions of the function $\mathcal{I}$ in power series of the parameters ξ, η. Thus, explicit computation up to second order yields the series,

$$\mathcal{I}(\xi,\eta) = -\tfrac{a}{d}\xi - \tfrac{1}{d}\left(3f_{3,0,0} - \tfrac{2af_{2,1,0}}{d} + \tfrac{a^2 f_{1,2,0}}{d^2}\right)\xi^2 - \tfrac{2e}{d}\xi\eta + \tfrac{1}{d}\eta^2 + \mathrm{O}_3(\xi,\eta), \tag{3.2.14}$$

and by substitution in (3.2.13), one obtains for the frequencies

$$\begin{aligned}\Omega_1(\xi,\eta) = \omega_1 &+ \left(d - \frac{ab}{d}\right)\xi + \\ &+ \left(-\frac{3b}{d}f_{3,0,0} - \frac{a^2 b}{d^3}f_{1,2,0} + \frac{2ab}{d^2}f_{2,1,0} + f_{2,1,0} - \frac{2a}{d}f_{1,2,0} + \frac{3a^2}{d^2}f_{0,3,0}\right)\xi^2 + \\ &+ \left(-\frac{2eb}{d} + 2f\right)\xi\eta + \frac{b}{d}\eta^2 + \mathrm{O}_3(\xi,\eta) \end{aligned} \tag{3.2.15a}$$

$$\begin{aligned}\Omega_2(\xi,\eta) = \omega_2 &+ \left(e - \frac{af}{d}\right)\xi + \eta + \\ &+ \left(-\frac{3f}{d}f_{3,0,0} - \frac{a^2 f}{d^3}f_{1,2,0} + \frac{2af}{d^2}f_{2,1,0} + f_{2,0,1} - \frac{a}{d}f_{1,1,1} + \frac{a^2}{d^2}f_{0,2,1}\right)\xi^2 + \\ &+ \left(2c - \frac{2ef}{d}\right)\xi\eta + \frac{f}{d}\eta^2 + \mathrm{O}_3(\xi,\eta). \end{aligned} \tag{3.2.15b}$$

Another point here to remark is the *qualitative* character of theorem 3.1, in the sense that it does not provide any idea on how large the neighborhood Γ could be. However, as the next lemma shows, one can use the fixed point theorem to determine Γ.

Lemma 3.2. *Assume* (3.2.9), (3.2.10) *and, as in theorem 3.1,* $d \neq 0$. *Then if R is small enough, there exist a constant* $0 < \widehat{c} < 1$, *independent of R, and a real analytic function, $\mathcal{I}$, defined on the set*

$$\Gamma^*(R) = \left\{(\xi,\eta) \in \mathbb{C}^2 : |\xi| \le \frac{\widehat{c}}{4}R^2, |\eta| \le \frac{\widehat{c}}{2}R\right\}, \tag{3.2.16}$$

satisfying identically (3.2.11), *with* $\mathcal{I}(\mathbf{0}) = 0$ *and* $|\mathcal{I}(\xi,\eta)| \le R^2/2$ *for all* $(\xi,\eta) \in \Gamma^*(R)$.

Proof. Let us denote $\boldsymbol{\zeta}^* = (\xi,\eta)$ and for a fixed $\boldsymbol{\zeta} \in \Gamma^*$ and consider,

$$\mathcal{F}(I;\boldsymbol{\zeta}) = \frac{1}{d}\left(\eta^2 - a\xi - 2e\xi\eta - \partial_1\mathcal{Z}_3(\xi, I, 2\xi\eta)\right).$$

It is clear that $\mathcal{I}(\boldsymbol{\zeta}) = \mathcal{F}(\mathcal{I}(\boldsymbol{\zeta}), \boldsymbol{\zeta})$. We want to check: (i) $|\mathcal{F}| \le R^2/2$ if $|I| \le R^2/2$, $|\xi| \le \frac{\widehat{c}}{4}R^2$ and (ii) $\mathcal{F}$ is a contraction with respect to I, uniform for $\boldsymbol{\zeta} \in \Gamma$. Then, suppose first that $|I| \le R^2/2$ and $\boldsymbol{\zeta} \in \Gamma^*$; taking norms in the expression above, we get,

$$\begin{aligned} |\mathcal{F}| &\le \frac{1}{|d|}\left(\frac{\widehat{c}^2}{4}R^2 + \frac{\widehat{c}}{4}|a|R^2 + \frac{\widehat{c}^2}{4}|e|R^3 + \widehat{\mathcal{Z}}_{3;1}R^4\right) \\ &\le \left(\frac{1+|a|+|e|}{4|d|}\widehat{c} + \frac{1}{|d|}\widehat{\mathcal{Z}}_{3;1}R^2\right)R^2 \end{aligned}$$

(assuming $R<1$, $\widehat{c}<1$ and hence $\widehat{c}^2 \le \widehat{c}$). Therefore, if we take $\widehat{c} = \min\left\{1, \frac{|d|}{1+|a|+|d|}\right\}$, it turns out that

$$|\mathcal{F}| \le \left(\frac{1}{4} + \frac{1}{|d|}\widehat{\mathcal{Z}}_{3;1}R^2\right)R^2 \le \frac{1}{2}R^2,$$

if R is sufficiently small. To assure the contractive character, we apply the mean value theorem,

$$\begin{aligned} |\mathcal{F}(I',\boldsymbol{\zeta}) - \mathcal{F}(I,\boldsymbol{\zeta})| &\le \frac{|\partial_1\mathcal{Z}_3(\xi, I', 2\xi\eta) - \partial_1\mathcal{Z}_3(\xi, I, 2\xi\eta)|}{|d|} \\ &\le \left(\sup_{\substack{\boldsymbol{\zeta}\in\Gamma^* \\ |I|\le R^2/2}} \left|\partial^2_{1,2}\mathcal{Z}_3\right|\right)\frac{|I'-I|}{|d|} \\ &\le \frac{\widehat{\mathcal{Z}}_{3;2}}{|d|}\,|I'-I|\,R^2 \\ &\le \frac{1}{2}\,|I'-I|\,, \end{aligned}$$

for R small enough. □

Remark 3.3. Note, in particular, from its construction, it follows that, $\mathcal{I}$ is real when ξ and η are. Then, one might consider its restriction $\mathcal{I} : \Gamma \to \mathbb{R}$, where Γ is the *real* domain:

$$\Gamma(R) = \left\{(\xi,\eta) \in \mathbb{R}^2 : |\xi| \le \frac{\widehat{c}}{4}R^2, |\eta| \le \frac{\widehat{c}}{2}R\right\}.$$

In what follows we shall specify whether the parameters (ξ,η) are taken in Γ or in its complex extension Γ^*. ♠

Remark 3.4. (On the nondegeneracy of the basic frequencies). When applying KAM methods, it is usual to select an invariant torus of the unperturbed Hamiltonian with "good" intrinsic frequencies and to ask for the persistence of this torus under small perturbations, or, more precisely, if the perturbed Hamiltonian has an invariant torus with the same frequencies close to the initial one. A way to achieve this, is to assume that the space of parameters describing the invariant tori (for example, the actions in the usual context) maps diffeomorphically onto the space of frequencies. Particularly, it is (locally) guaranteed if,

$$\det\left(\partial_{\boldsymbol{\zeta}}\boldsymbol{\Omega}(\mathbf{0})\right) \ne 0, \qquad (3.2.17)$$

which is the standard (or Kolmogorov) nondegeneracy condition [(1)]. From the "frequency map" given by the expansions (3.2.15a) and (3.2.15b), one checks that,

$$\det\left(\partial_{\boldsymbol{\zeta}}\boldsymbol{\Omega}\right) = d - \frac{ab}{d} + \mathrm{O}_1(\boldsymbol{\zeta}).$$

So, $\boldsymbol{\Omega}(\boldsymbol{\zeta})$ should be a (local) diffeomorphism, provided $d - ab/d \neq 0$. Although this is the classic approach, we shall be forced to choose a different set of parameters to govern the successive approximations to the family of perturbed tori. This will be discussed in the next section. ♠

Remark 3.5. If, instead of the symplectic coordinates introduced by the transformation (3.2.2), we consider the real coordinates defined by the change (1.9.1), then it is clear that,

$$\begin{aligned}
&\theta_1 = \Omega_1(\boldsymbol{\zeta})t + \theta_1^{(0)}, && x_1 = \sqrt{2\xi}\cos(\Omega_2(\boldsymbol{\zeta})t + \theta_2^{(0)}), && x_2 = -\sqrt{2\xi}\sin(\Omega_2(\boldsymbol{\zeta})t + \theta_2^{(0)}),\\
&I_1 = \mathcal{I}(\boldsymbol{\zeta}), && y_1 = -\eta\sqrt{2\xi}\sin(\Omega_2(\boldsymbol{\zeta})t + \theta_2^{(0)}), && y_2 = -\eta\sqrt{2\xi}\cos(\Omega_2(\boldsymbol{\zeta})t + \theta_2^{(0)}),
\end{aligned}$$

(with $\boldsymbol{\zeta}^* = (\xi, \eta)$) is a two-parameter family of quasi-periodic solutions of the (real) unperturbed Hamiltonian

$$\begin{aligned}
H_0(\theta_1, x_1, x_2, I_1, y_1, y_2) = {} & \omega_1 I_1 + \omega_2(x_2 y_1 - x_1 y_2)\\
& + \tfrac{1}{2}(y_1^2 + y_2^2) + \mathscr{Z}_r\left(\frac{1}{2}(x_1^2 + x_2^2), I_1, x_2 y_1 - x_1 y_2\right),
\end{aligned}$$

(see (1.8.1)). So, as we are interested in *real* tori, it is clear that it is necessary to impose the condition $\xi \geq 0$ and this assumption is kept throughout. ♠

3.2.2 Linear behavior of the unperturbed tori: normal frequencies and lack of parameters

Again, consider the system (1.9.3) and the family $\{\mathcal{T}_{\boldsymbol{\zeta}}\}_{\boldsymbol{\zeta}\in\Gamma}$ of invariant tori. Hence, around one of these tori, the first variational equations in the normal directions are given by the linear system $\dot{\boldsymbol{Z}} = M_{\boldsymbol{\zeta}}\boldsymbol{Z}$, with the 2×2 matrix

$$M_{\xi,\eta} = \begin{pmatrix} 0 & 2\xi \\ -\frac{2\eta^2}{\xi} - a - \partial_{1,1}^2\mathscr{Z}_3(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta) & 0 \end{pmatrix}, \tag{3.2.18}$$

and then, the characteristic exponents (or normal eigenvalues) of the unperturbed (real) torus are,

$$\lambda_{\pm}(\xi,\eta) = \pm\sqrt{-4\eta^2 - 2a\xi - 2\xi\partial_{1,1}^2\mathscr{Z}_3(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta)}, \quad (\xi,\eta)\in\Gamma(R),\ \xi > 0, \tag{3.2.19}$$

and from the inequalities (3.2.9) and (3.2.10) it follows that, inside the square root, the last term is of fourth order in R, so we write symbolically: $\xi\partial_{1,1}^2\mathscr{Z}_3(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta) = \mathrm{O}(R^4)$,

[(1)] We allow some abuse of notation: through this chapter, $\boldsymbol{\Omega}$ will denote both, the vector of intrinsic frequencies (images of the frequency map for some values of the parameters) and the frequency map itself. The actual meaning should be clear by the context.

whereas $4\eta^2 + 2a\xi = \mathrm{O}(R^2)$. This implies that –for R small enough–, the hyperbolicity of the tori of the family is determined by the sign of the first two terms. In particular, if the coefficient a is positive, then the family only holds elliptic invariant tori, whilst for $a < 0$ both, elliptic and hyperbolic invariant tori will be present together with parabolic ones for, if one considers the equation $4\eta^2 + 2a\xi + 2\xi\partial_{1,1}\mathcal{Z}_3(\xi, \mathcal{I}(\xi,\eta), 2\xi\eta) = 0$, just the implicit function theorem applied at $(\xi,\eta) = (0,0)$ shows the existence –in the space of parameters (ξ,η)–, of a path, $\xi = \mathfrak{g}(\eta)$, giving rise to parabolic tori[(2)]. For an invariant torus, with reducible normal variational directions, it is usual to define its vector of *normal frequencies* as the positive imaginary parts of the characteristic exponents. One of the problems intrinsically linked to the perturbation of elliptic invariant tori is the so called "lack of parameters". In fact, this is a common difficulty in the theory of quasi-periodic motions in dynamical systems (see Moser (1967), and also Sevryuk (1999)). As an illustration of this fact, suppose we select $\boldsymbol{\zeta_0} \in \Gamma$ such that the corresponding torus is elliptic with frequencies $\boldsymbol{\Omega}^* = (\Omega_1, \Omega_2)$ and $\mu = |\lambda_+|$, satisfy the Diophantine conditions (Melnikov nonresonance conditions),

$$|\langle \boldsymbol{k}, \boldsymbol{\Omega}(\boldsymbol{\zeta_0})\rangle + \ell\mu(\boldsymbol{\zeta_0})| \geq \gamma|\boldsymbol{k}|^{-\tau}, \quad \forall \boldsymbol{k} \in \mathbb{Z}^n \setminus \{\boldsymbol{0}\}, \quad \forall \ell \in \mathbb{Z} \text{ with } |\ell| \leq 2, \tag{3.2.20}$$

and γ, τ positive constants. Here, $\mu(\boldsymbol{\zeta})$ is the normal frequency as defined above (i. e., the positive imaginary part of $\lambda_\pm(\boldsymbol{\zeta})$). Since (3.2.20) "controls" the small divisors appearing when one seeks elliptic reducible invariant tori, one could naively expect (as in the case of maximal dimensional tori) that the perturbed system has a torus with exactly the same frequencies (both normal and intrinsic). However, this conjecture fails, since the intrinsic and the normal frequencies are not functionally independent. Indeed: just assume the perturbation depends only on the actions, say, $\mathcal{R}(\boldsymbol{\theta}, q, \boldsymbol{I}, p) = \widetilde{\mathcal{R}}(\boldsymbol{I})$. This will preserve the family of invariant tori, but changing slightly their intrinsic frequencies to $\widetilde{\boldsymbol{\Omega}}(\boldsymbol{\zeta})$. Therefore, under the nondegeneracy condition on $\boldsymbol{\Omega}(\boldsymbol{\zeta})$ discussed in remark 3.4, the perturbed system will have an invariant torus exactly with the same intrinsic frequencies, since $\widetilde{\boldsymbol{\Omega}}(\boldsymbol{\zeta_1}) = \boldsymbol{\Omega}(\boldsymbol{\zeta_0})$ for some $\boldsymbol{\zeta_1} \in \Gamma$ close to $\boldsymbol{\zeta_0}$, but clearly, $\mu(\boldsymbol{\zeta_1}) \neq \mu(\boldsymbol{\zeta_0})$ in general. Thus, generically, one cannot construct a perturbed torus with a fixed set of (Diophantine) intrinsic and normal frequencies, for the system does not contain enough internal parameters to control them all simultaneously. All that one can expect is to build perturbed tori with only a given subset of frequencies previously fixed. But to succeed, we ask the small denominators to change when these selected frequencies do. This can be guaranteed adding suitable nondegeneracy conditions on those remaining frequencies. In the standard formulation (see paragraph below) one chooses naturally the intrinsic frequencies to pick out the invariant tori, while the normal frequencies are thought to depend on these basic ones.

Consider the general case in which an n-degrees of freedom unperturbed Hamiltonian has an r-parameter family of r-dimensional invariant tori with r nondegenerate (in Kolmogorov' sense) intrinsic frequencies. Assume these tori are reducible in the normal directions, which we suppose elliptic. Abusing notation, let $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\boldsymbol{\zeta}) \in \mathbb{R}^r$ and $\boldsymbol{\lambda}(\boldsymbol{\zeta}) \in \mathbb{C}^{2(n-r)}$, with $\lambda_{j+n-r}(\boldsymbol{\zeta}) = -\lambda_j(\boldsymbol{\zeta})$ for $j = 1, \ldots, n-r$, be the vectors with intrinsic frequencies and the characteristic exponents respectively, where $\lambda_j = \mathrm{i}\mu_j$ being

[(2)]Of course, the same can be done when $a < 0$, but then ξ, as a function of η, will take (at least for R sufficiently small) only negative values, against the condition for real invariant tori.

$\mu_j \in \mathbb{R}^+$ the j-th normal frequency for $j = 1, \ldots, n-r$. By the assumed nondegeneracy of $\boldsymbol{\Omega}(\boldsymbol{\zeta})$ at $\boldsymbol{\zeta} = \mathbf{0}$, the normal frequencies can be put as functions of the basic ones in a neighborhood of this point. Now, working directly with $\boldsymbol{\Omega}$ as the parameters of the family, one looks up the divisors of the KAM process $\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\Omega}\rangle + \langle \boldsymbol{\ell}, \boldsymbol{\lambda}(\boldsymbol{\Omega})\rangle$, with $\boldsymbol{\ell} \in \mathbb{Z}^{2(n-r)}$, $0 \le |\boldsymbol{\ell}|_1 \le 2$ and ask them to "move" as a function of $\boldsymbol{\Omega}$. Thus, it is enough to impose the "transversality condition",

$$\mathrm{Im}\,(\mathrm{grad}_{\boldsymbol{\Omega}}\langle \boldsymbol{\ell}, \boldsymbol{\lambda}(\boldsymbol{\Omega})\rangle|_{\boldsymbol{\Omega}=\boldsymbol{\omega}}) \notin \mathbb{Z}^r, \quad \text{and} \quad \ell_j \ne \ell_{j+n-r}, \quad \text{for } j = 1, \ldots, n-r \tag{3.2.21}$$

where $\boldsymbol{\omega}$ is the value of the frequency map at $\boldsymbol{\zeta} = \mathbf{0}$, i. e., $\boldsymbol{\Omega}(\mathbf{0}) = \boldsymbol{\omega}$ (see Jorba and Villanueva, 1997a). These, however, do not work to our particular problem. To realize easily, assume the coefficient a in (3.2.19) is positive, so the characteristic exponents will be purely imaginary (in a sufficiently small neighborhood of $\boldsymbol{\zeta} = \mathbf{0}$) and hence an elliptic unperturbed family of invariant tori unfolds. Next, a glance to (3.2.15a) and (3.2.15b) shows that the first order expansion, at $\boldsymbol{\Omega} = \boldsymbol{\omega}$, of the inverse of the frequency map must be,

$$\begin{aligned}
\xi &= \frac{d}{d^2 - ab}(\Omega_1 - \omega_1) + \mathrm{O}_2(\boldsymbol{\Omega} - \boldsymbol{\omega}),\\
\eta &= \frac{af - ed}{d^2 - ab}(\Omega_1 - \omega_1) + (\Omega_2 - \omega_2) + \mathrm{O}_2(\boldsymbol{\Omega} - \boldsymbol{\omega}),
\end{aligned}$$

(indeed, we assume $d^2 - ab \ne 0$). Now, substitution in the above expression for the characteristic exponents, gives for the normal frequency:

$$\mu = +\sqrt{-\frac{2ad}{d^2 - ab}(\Omega_1 - \omega_1) + \mathrm{O}_2(\boldsymbol{\Omega} - \boldsymbol{\omega})}.$$

So the derivative involved in the condition (3.2.21) is not defined at all when $\boldsymbol{\omega} = \boldsymbol{\Omega}$. Hence, it becomes necessary to choose an alternative set of parameters to characterize the perturbed invariant tori. From (3.2.15b) and (3.2.19), it can be seen that ξ and η may be expressed as a function of the other two frequencies, Ω_2 and μ. In fact, substitution and some additional algebra lead to the expansions,

$$\begin{aligned}
\xi &= \frac{\mu^2}{2a} - \frac{2}{a}(\Omega_2 - \omega_2)^2 + \mathrm{O}_3(\mu, \Omega_2 - \omega_2),\\
\eta &= \Omega_2 - \omega_2 + \left(\frac{f}{2d} - \frac{e}{2a}\right)\mu^2 + \left(\frac{2e}{a} - \frac{3f}{d}\right)(\Omega_2 - \omega_2)^2 + \mathrm{O}_3(\mu, \Omega_2 - \omega_2),
\end{aligned}$$

and then, substitution in (3.2.15a) yields:

$$\Omega_1(\mu, \Omega_2) = \omega_1 + \left(\frac{d}{2a} - \frac{b}{2d}\right)\mu^2 + \left(\frac{3b}{d} - \frac{2d}{a}\right)(\Omega_2 - \omega_2)^2 + \mathrm{O}_3(\mu, \Omega_2 - \omega_2) \tag{3.2.22}$$

(we recall that the assumptions $a > 0$ and $d \ne 0$ were made). Now, we can formulate an alternative to the nondegeneracy conditions (3.2.21) suited to our problem: let us denote $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$; now, the derivatives of the denominators involved in the present KAM process w. r. t. these frequencies, are

$$\mathrm{grad}_{\boldsymbol{\Lambda}}\,(k_1\Omega_1(\boldsymbol{\Lambda}) + k_2\Omega_2 + \ell\mu)|_{\boldsymbol{\Lambda}^*=(0,\omega_2)} = \begin{pmatrix} \ell \\ k_2 \end{pmatrix}, \qquad k_1, k_2, \ell \in \mathbb{Z}, \text{ with } |\ell| \le 2,$$

so the divisors will change with $\boldsymbol{\Lambda}$ whenever the integer vector on the r. h. s. is not the null vector, but in such case, necessarily $k_1 \neq 0$ and hence the moduli of the divisors will be bounded from below (for R small enough, since $\Omega_1(\boldsymbol{\Lambda})$ does, as seen from (3.2.22)). In few words: if Ω_1 is taken as function of μ and Ω_2, we can guarantee the correct behavior of the small divisors required for the KAM method.

Remark 3.6. From now on we will refer μ and Ω_2 as the *basic frequencies* labeling the invariant tori of the system, whereas the denomination of *intrinsic* frequencies will be reserved for Ω_1, Ω_2. ♠

On the other hand, the size of the domain where μ and Ω_2 move, can be derived from the lemmas below. Before, let Γ^* and $\mathcal{I}$, be the set and the function introduced in lemma 3.2 respectively. In addition, we define the sets:

$$\begin{aligned} \mathcal{M}^*(R) &= \left\{(\mu,\eta)\in\mathbb{C}^2 : |\mu|\le \check{c}R,\ |\eta|\le\check{c}R\right\},\\ \mathcal{N}^*(R) &= \left\{(\mu,\Omega_2)\in\mathbb{C}^2 : |\mu|\le \frac{1}{4}\check{c}R,\ |\Omega_2-\omega_2|\le\frac{1}{4}\check{c}R\right\},\\ \widetilde{\Gamma}^*(R) &= \left\{(\xi,\eta)\in\mathbb{C}^2 : |\xi|\le \frac{\widehat{c}}{8}R^2,\ |\eta|\le\check{c}R\right\}, \end{aligned} \tag{3.2.23}$$

with $0<\check{c}\le\frac{\widehat{c}}{2}$ (clearly, $\widetilde{\Gamma}^*\subset\Gamma^*$), and also the notation:

$$\boldsymbol{\zeta}^* = (\xi,\eta), \qquad \boldsymbol{\sigma}^* = (\mu,\eta), \qquad \boldsymbol{\Lambda}^* = (\mu,\Omega_2).$$

With these conventions, we can state:

Lemma 3.7. *There exists a real analytic function, g, defined in $\mathcal{M}^*$, such that if $\check{c}\le\min\left\{\frac{\widehat{c}}{2},\sqrt{\frac{a\widehat{c}}{40}}\right\}$ and for R small enough $(\xi,\mu,\eta)=(g(\boldsymbol{\sigma}),\boldsymbol{\sigma}^*)$ satisfies identically the equation,*

$$\mu^2 = 4\eta^2 + 2a\xi + 2\xi\partial^2_{1,1}\mathcal{Z}_3(\xi,\mathcal{I}(\boldsymbol{\zeta}),2\xi\eta). \tag{3.2.24}$$

Moreover, it is $|g(\mu,\eta)|\le\widehat{c}R^2/8$.

Proof. One proceeds exactly as in the proof of lemma 3.2. We define:

$$\mathcal{F}(\xi;\mu,\eta) = \frac{\mu^2}{2a} - \frac{2\eta^2}{a} - \frac{1}{a}\xi\partial^2_{1,1}\mathcal{Z}_3(\xi,\mathcal{I}(\xi,\eta),2\xi\eta).$$

It can be seen straightforward that $g(\boldsymbol{\sigma})=\mathcal{F}(g(\boldsymbol{\sigma});\boldsymbol{\sigma}^*)$. We must check:

(i) for $|\mu|\le\check{c}R$, $|\eta|\le\check{c}R$ and $|\xi|\le\widehat{c}R^2/8$ with $0<\check{c}\le\widehat{c}/2<1$ is $|\mathcal{F}|\le\frac{\widehat{c}}{8}R^2$, if R is sufficiently small.

(ii) $\mathcal{F}$ is a contraction with respect to ξ uniform in $\boldsymbol{\sigma}$.

The first is obviously true:

$$\begin{aligned} |\mathcal{F}| &\le \frac{1}{2a}\check{c}^2R^2 + \frac{2}{a}\check{c}^2R^2 + \frac{\widehat{c}}{8a}\widehat{\mathcal{Z}}_{3;2}R^4\\ &\le \frac{\widehat{c}}{8}R^2, \end{aligned}$$

taking, $\check{c} \leq \min\left\{\frac{\widehat{c}}{2}, \sqrt{\frac{a\widehat{c}}{40}}\right\}$, and R small enough. Now, to prove the second item, we introduce:

$$\begin{aligned} \mathcal{I}_\eta &= \mathcal{I}(\xi,\eta), & \mathcal{I}'_\eta &= \mathcal{I}(\xi',\eta), \\ \mathcal{F}_{\boldsymbol{\sigma}} &= \mathcal{F}(\xi;\boldsymbol{\sigma}), & \mathcal{F}'_{\boldsymbol{\sigma}} &= \mathcal{F}(\xi';\boldsymbol{\sigma}), \end{aligned}$$

apply the mean value theorem,

$$\begin{aligned} |\mathcal{F}'_{\boldsymbol{\sigma}} - \mathcal{F}_{\boldsymbol{\sigma}}| &\leq \frac{1}{a}|\xi'\partial^2_{1,1}\mathcal{Z}_3(\xi',\mathcal{I}'_\eta,2\xi'\eta) - \xi\partial^2_{1,1}\mathcal{Z}_3(\xi,\mathcal{I}_\eta,2\xi\eta)| \\ &\leq \frac{1}{a}\Bigg(\sup_{|q|,|I_1|,|I_2|\leq R^2/2}|\partial^3_{1,1,1}\mathcal{Z}_3| + \sup_{|q|,|I_1|,|I_2|\leq R^2/2}|\partial^3_{1,1,2}\mathcal{Z}_3| \times \sup_{\boldsymbol{\zeta}\in\widetilde{\Gamma}^*}|\partial_\xi\mathcal{I}|+ \\ &\qquad + 2|\eta|\sup_{|q|,|I_1|,|I_2|\leq R^2/2}|\partial^3_{1,1,3}\mathcal{Z}_3|\Bigg)|\xi'-\xi| \times |\xi|+ \\ &\quad + \frac{1}{a}|\xi'-\xi||\partial^2_{1,1}\mathcal{Z}_3(\xi',\mathcal{I}'_\eta,2\xi'\eta)| \\ &\leq \frac{1}{a}\left(\frac{\widehat{c}}{8}\widehat{\mathcal{Z}}_{3;3} + \frac{1}{2}\widehat{\mathcal{Z}}_{3;3} + \frac{\check{c}\,\widehat{c}}{4}\widehat{\mathcal{Z}}_{3;3}R + \widehat{\mathcal{Z}}_{3;2}\right)|\xi'-\xi|R^2 \\ &\leq \frac{1}{2}|\xi'-\xi|, \end{aligned}$$

for R small enough, where we have used the Cauchy's inequalities to bound the norm of the derivative $\partial_\xi\mathcal{I}$, according with:

$$\sup_{\boldsymbol{\zeta}\in\widetilde{\Gamma}^*}|\partial_\xi\mathcal{I}| \leq \frac{R^2/2}{(\widehat{c}/4-\widehat{c}/8)R^2},$$

(see lemma 3.2) and this ends the proof of the lemma. □

Lemma 3.8. *For $\widehat{c}$, $\check{c}$ and and R small enough, there exists a real analytic function h, defined in $\mathcal{N}^*$, such that the equation,*

$$\Omega_2 = \omega_2 + \eta + eg(\boldsymbol{\sigma}) + f\mathcal{I}(g(\boldsymbol{\sigma}),\eta) + 2c\eta g(\boldsymbol{\sigma}) + \partial_3\mathcal{Z}_3(g(\boldsymbol{\sigma}),\mathcal{I}(g(\boldsymbol{\sigma}),\eta),2\eta g(\boldsymbol{\sigma})),$$

is satisfied identically by $(\boldsymbol{\Lambda},\eta) = (\boldsymbol{\Lambda},h(\boldsymbol{\Lambda}))$. Furthermore $|h(\boldsymbol{\Lambda})| \leq \check{c}R/2$, for all $\boldsymbol{\Lambda}\in\mathcal{N}^*$.

Proof. It relies on the fixed point theorem, and is formally identical to those of lemmas 3.2 and 3.7. □

Thus, the (unperturbed) invariant tori described in theorem 3.1, can be parameterized by (μ,Ω_2) belonging to a neighborhood of $(0,\omega_2)$. Nonetheless, we shall see that only those tori with (μ,Ω_2) in a Cantorian subset will survive when the whole Hamiltonian (in the sense discussed above) is considered. This is the result of theorem 3.9 stated below, but previously, we shall introduce:

(1) ρ_0, $R^{(0)}$:

$$\rho_0 = \frac{1}{2}\min\{\rho, \ln 2\}, \qquad R^{(0)} = 2M^\alpha(R), \tag{3.2.25}$$

where $M(R)$ is the bound for $\mathfrak{R}$, i. e., the term corresponding to the remainder in the Hamiltonian (3.2.1) –and computed in the previous chapter–, whereas α is a fixed (real) quantity $0 < \alpha < 1$.

(2) The domains $\mathcal{D}_*(\widetilde{\rho}, \widetilde{R})$,

$$\mathcal{D}_*(\widetilde{\rho}, \widetilde{R}) = \{(\boldsymbol{\theta}, x, \boldsymbol{I}, y) \in \mathbb{C}^2 \times \mathbb{C}^1 \times \mathbb{C}^2 \times \mathbb{C}^1 : \\ |\mathrm{Im}\, \boldsymbol{\theta}| \le \widetilde{\rho},\ |\boldsymbol{z}| \le \widetilde{R},\ |\widehat{\boldsymbol{I}}| \le (\widetilde{R})^2\}, \tag{3.2.26}$$

with $\boldsymbol{z}^* = (x, y)$ (so $\mathcal{D}_*(\widetilde{\rho}, \widetilde{R}) = \mathcal{D}_{2,1}(\widetilde{\rho}, \widetilde{R})$),

(3) For a given $A > 0$, the sets $\mathcal{U}(A)$, $\mathcal{W}(A)$ and $\mathcal{V}(A)$ given by:

$$\mathcal{U}(A) = \left\{\boldsymbol{\Lambda} \in \mathbb{R}^2 : |\Omega_2 - \omega_2| \le A,\ |\mu| \le A\right\}, \tag{3.2.27}$$

$$\mathcal{W}(A) = \Big\{\boldsymbol{\Lambda} \in \mathcal{U}(A) : \Omega_2 = \omega_2 + e\xi + \eta + 2c\xi\eta + f\mathcal{I}(\boldsymbol{\zeta}) + \partial_3 \mathcal{Z}_r(\xi, \mathcal{I}(\boldsymbol{\zeta}), 2\xi\eta), \\ \mu = \sqrt{4\eta^2 + 2a\xi + 2\xi\partial_{1,1}\mathcal{Z}_3(\xi, \mathcal{I}(\boldsymbol{\zeta}), 2\xi\eta)},\ \text{with } \boldsymbol{\zeta} \in \Gamma(R),\ \xi > 0\Big\}, \tag{3.2.28}$$

$$\mathcal{V}(A) = \left\{\boldsymbol{\zeta} \in \mathbb{R}^2 : |\xi| \le A,\ |\eta| \le \frac{\check{c}}{4}R\right\}. \tag{3.2.29}$$

With the preceding notation and the results of the previous sections, we can enunciate the following theorem, whose proof extends up to the end of the present chapter.

Theorem 3.9. *Consider the real Hamiltonian $\mathcal{H}$ (equation (3.2.4)), in the domain defined by the inequalities (3.2.3), $\mathcal{D}(M, \rho, R)$, where $\mathcal{Z} = \mathcal{Z}_2 + \mathcal{Z}_3$, with $\mathcal{Z}_2$ and $\mathcal{Z}_3$ are given by (3.2.6) and (3.2.7) respectively. Assume, in addition, that the real coefficients a, d appearing in the development of $\mathcal{Z}_2$ are $a > 0$ and $d \neq 0$. Then, for $0 < R < 1$ small enough, there exists a Cantor subset $\mathfrak{A} \subset \mathbb{R}^2$, such that for any $\boldsymbol{\Lambda}^* = (\mu, \Omega_2) \in \mathfrak{A}$, the Hamiltonian system $\mathcal{H}$ has an invariant real 2-dimensional torus with parallel flow, normal frequency given by μ and the second component of its vector of intrinsic frequencies equal to Ω_2. Furthermore, these invariant tori are elliptic (i. e., with purely imaginary characteristic exponents) and reducible in their normal directions (that is, the first variational equations can be reduced to a system with constant coefficients in those coordinates which account for the normal flow).*

The Cantorian set $\mathfrak{A}$ is characterized in the following form:

(i) for any $R > 0$ small enough, and $\boldsymbol{\Lambda} \in \mathfrak{A}$, the Hamiltonian $\mathcal{H}$ can be reduced, through a symplectic change, Ψ, depending on $\boldsymbol{\Lambda}$ and defined on the set $\mathcal{D}_*(\rho_0, R^{(0)} \exp(-3\rho_0/8))$, into the (family of) Hamiltonians $H_{\boldsymbol{\Lambda}}^{(\infty)} = \mathcal{H} \circ \Psi$, given by:

$$H_{\boldsymbol{\Lambda}}^{(\infty)} = \phi^{(\infty)}(\boldsymbol{\Lambda}) + \langle \boldsymbol{\Omega}^{(\infty)}(\boldsymbol{\Lambda}), \boldsymbol{I}\rangle + \frac{1}{2}\langle \boldsymbol{z}, \mathcal{B}(\boldsymbol{\Lambda})\boldsymbol{z}\rangle \\ + \frac{1}{2}\langle \boldsymbol{I}, \mathcal{C}^{(\infty)}(\boldsymbol{\theta}; \boldsymbol{\Lambda})\boldsymbol{I}\rangle + \langle \boldsymbol{z}, \mathcal{E}(\boldsymbol{\Lambda})\boldsymbol{I}\rangle + H_*^{(\infty)}(\boldsymbol{\theta}, x, \boldsymbol{I}, y; \boldsymbol{\Lambda}), \tag{3.2.30}$$

Here, the 2×2 *matrices* $\mathcal{E}$, $\mathcal{B}$ *and the first component of the basic frequency vector* $\Omega_1^{(\infty)}$ *depends on the chosen* $\boldsymbol{\Lambda} \in \mathfrak{A}$ *whilst* $H_*^{(\infty)}$ *contains terms of degree higher than two (in the actions* $\boldsymbol{I}^* = (I_1, I_2)$ *and in the normal variables* $\boldsymbol{z}^* = (x, y)$*).*

(ii) For $0 < R < 1$ *small enough, if we put* $\mathfrak{A}(R) = \mathcal{W}(\frac{1}{8}\check{c}R) \cap \mathfrak{A}$ *–see* (3.2.28) *for the definition of* $\mathcal{W}(A)$*– then,*

$$\mathit{meas}\left(\mathcal{W}\left(\frac{1}{8}\check{c}R\right) \setminus \mathfrak{A}(R)\right) \leq \ \mathit{constant} \ \cdot M^{\alpha/2}(R),$$

being the constant on the right hand side independent of R.

(iii) $(\boldsymbol{\theta}, \boldsymbol{\Lambda}) \in \mathbb{T}^2 \times \mathfrak{A} \mapsto \Psi(\boldsymbol{\theta}, 0, \mathbf{0}, 0; \boldsymbol{\Lambda})$ *is a parametrization of a Whitney regular Cantorian manifold of invariant tori of* $\mathcal{H}$ *which can be completed to a* $\mathcal{C}^\infty$ *regular in* $\boldsymbol{\Lambda}$ *and analytic in* $\boldsymbol{\theta}$ *manifold, in such a way that the measure of the extension of the Cantorian manifold to this regular-analytic manifold is of the same order than the measure of gaps coming from the elimination of (basic) frequencies in the KAM process. Nevertheless, only those points on the Cantorian manifold will correspond to tori of the Hamiltonian* $\mathcal{H}$.

Remark 3.10. In addition, we note here that, for every $\boldsymbol{\Lambda} \in \mathfrak{A}$, the claimed invariant torus corresponds to $\boldsymbol{z} = \boldsymbol{I} = \mathbf{0}$, and, indeed, the flow of $H_{\boldsymbol{\Lambda}}^{(\infty)}$ restricted to that torus, is parallel with frequency given by $\boldsymbol{\Omega}^{(\infty)}$. It is also obvious, from (3.2.30), that the first variational equations in the normal directions, without varying the actions, i. e. taking $\boldsymbol{I} = \mathbf{0}$ fixed, are $\dot{\boldsymbol{z}} = J_1 \mathcal{B} \boldsymbol{z}$. This last precision on the variation of $\boldsymbol{I}$ can be skipped if one first removes the coupled quadratic term $\langle \boldsymbol{z}, \mathcal{E}(\boldsymbol{\Lambda})\boldsymbol{I}\rangle$, by means of an appropriate symplectic change. In particular, we can consider the transformation $(\boldsymbol{\theta}, x, \boldsymbol{I}, y) \to (\boldsymbol{\varphi}, X, \boldsymbol{K}, Y)$, depending on $\boldsymbol{\Lambda}$ and defined through:

$$\boldsymbol{\theta} = \boldsymbol{\varphi} + \boldsymbol{\Phi}(\boldsymbol{Z}; \boldsymbol{\Lambda}), \quad x = X + \mathcal{X}(\boldsymbol{K}; \boldsymbol{\Lambda}), \quad \boldsymbol{I} = \boldsymbol{K}, \quad y = Y + \mathcal{Y}(\boldsymbol{K}; \boldsymbol{\Lambda}), \tag{3.2.31}$$

where $\boldsymbol{Z}^* = (X, Y)$, and the functions $\boldsymbol{\Phi}$, $\mathcal{X}$, $\mathcal{Y}$ are defined by:

$$\begin{aligned}
\boldsymbol{\Phi}(\boldsymbol{Z}; \boldsymbol{\Lambda}) &= \frac{X}{\lambda_+}\boldsymbol{w}(\boldsymbol{\Lambda}) + \frac{\xi(\boldsymbol{\Lambda})Y}{\lambda_+^2}\boldsymbol{w}(\boldsymbol{\Lambda}), \\
\mathcal{X}(\boldsymbol{K}; \boldsymbol{\Lambda}) &= \frac{\xi(\boldsymbol{\Lambda})}{\lambda_+^2}\langle \boldsymbol{w}(\boldsymbol{\Lambda}), \boldsymbol{K}\rangle, \\
\mathcal{Y}(\boldsymbol{K}; \boldsymbol{\Lambda}) &= -\frac{1}{\lambda_+}\langle \boldsymbol{w}(\boldsymbol{\Lambda}), \boldsymbol{K}\rangle.
\end{aligned} \tag{3.2.32}$$

We observe that this transformation is symplectic, since as can be immediately seen, is obtained as the time-unit flow of the (family of) Hamiltonians,

$$\mathcal{K}_{\boldsymbol{\Lambda}} = \frac{1}{\lambda_+}\left(x + \frac{\xi(\boldsymbol{\Lambda})}{\lambda_+}y\right)\langle \boldsymbol{w}(\boldsymbol{\Lambda}), \boldsymbol{I}\rangle.$$

Here, ξ is considered as a function of the basic frequencies, $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$ –see lemmas 3.7 and 3.8–, $\lambda_+ = \mathrm{i}\mu$ (see (3.2.19)) whereas $\boldsymbol{w}^* = (\mathcal{E}_{1,1}, \mathcal{E}_{2,1})$, (i. e., the components of $\boldsymbol{w}$ are those of the first column of the matrix $\mathcal{E}(\boldsymbol{\Lambda})$, see (3.2.38)). Then, the change specified by (3.2.31) and (3.2.32) transforms the (family of) Hamiltonians (3.2.30) to[3]:

$$\breve{H}^{(\infty)}_{\boldsymbol{\Lambda}} = \phi^{(\infty)}(\boldsymbol{\Lambda}) + \langle \boldsymbol{\Omega}^{(\infty)}(\boldsymbol{\Lambda}), \boldsymbol{K}\rangle + \frac{1}{2}\langle \boldsymbol{Z}, \mathcal{B}(\boldsymbol{\Lambda})\boldsymbol{Z}\rangle \\ + \frac{1}{2}\langle \boldsymbol{K}, \breve{\mathcal{C}}^{(\infty)}(\boldsymbol{\varphi};\boldsymbol{\Lambda})\boldsymbol{K}\rangle + \breve{H}^{(\infty)}_{*}(\boldsymbol{\varphi}, X, \boldsymbol{K}, Y; \boldsymbol{\Lambda}), \tag{3.2.33}$$

being,

$$\breve{\mathcal{C}}^{(\infty)}(\boldsymbol{\varphi};\boldsymbol{\Lambda}) = \mathcal{C}^{(\infty)}(\boldsymbol{\varphi};\boldsymbol{\Lambda}) + \frac{2\xi(\boldsymbol{\Lambda})}{\lambda_+^2}A(\boldsymbol{\Lambda}),$$

$$\breve{H}^{(\infty)}_{*}(\boldsymbol{\varphi}, X, \boldsymbol{K}, Y; \boldsymbol{\Lambda}) = H^{(\infty)}_{*}(\boldsymbol{\varphi} + \boldsymbol{\Phi}(\boldsymbol{Z};\boldsymbol{\Lambda}), X + \mathcal{X}(\boldsymbol{Z};\boldsymbol{\Lambda}), \boldsymbol{K}, Y + \mathcal{Y}(\boldsymbol{Z};\boldsymbol{\Lambda}); \boldsymbol{\Lambda}) + \\ + \frac{1}{2}\left\langle \boldsymbol{K}, \left(\mathcal{C}^{(\infty)}(\boldsymbol{\varphi} + \boldsymbol{\Phi}(\boldsymbol{Z};\Lambda);\boldsymbol{\Lambda}) - \mathcal{C}^{(\infty)}(\boldsymbol{\varphi};\boldsymbol{\Lambda})\right)\boldsymbol{K}\right\rangle,$$

and where (to simplify) we have introduced the matrix $A(\boldsymbol{\Lambda})$, as:

$$A_{i,j}(\boldsymbol{\Lambda}) = w_i(\boldsymbol{\Lambda})w_j(\boldsymbol{\Lambda}), \qquad i,j = 1,2.$$

Therefore, the family of Hamiltonians (3.2.33) have, for each $\boldsymbol{\Lambda}$ in the cantor set $\mathfrak{A}$ (see theorem 3.9), an invariant torus at the origin (more precisely at $\boldsymbol{Z} = \boldsymbol{K} = \boldsymbol{0}$), with parallel flow determined by the vector of intrinsic frequencies $\boldsymbol{\Omega}^{(\infty)}(\boldsymbol{\Lambda})$. Furthermore, in the normal directions, the variational equations around these tori are $\dot{\boldsymbol{Z}} = J_1\mathcal{B}\boldsymbol{Z}$, and hence they are normally reducible. ♠

In what follows, and up to the end of this chapter, we shall develop a proof for theorem 3.9. The starting point is to pick values (supposed "good") of μ and Ω_2 and expand the Hamiltonian (3.2.4) around the corresponding unperturbed torus. This will lead to a more suitable (from the point of view of the applicability of KAM iterative schemes) Hamiltonian.

3.2.3 Expansion of the Hamiltonian around the unperturbed invariant tori

To apply KAM methods, it is convenient for us to put the Hamiltonian (3.2.4) into a more suitable form. In particular the proof (of theorem 3.9) follows similar ideas than those in Jorba and Villanueva (1997a): we replace the initial Hamiltonian (see (3.2.4)) by a family of Hamiltonians, $H^{(0)}_{\boldsymbol{\Lambda}}$, having as a parameter the vector of basic frequencies $\boldsymbol{\Lambda}$. The Hamiltonian $H^{(0)}_{\boldsymbol{\Lambda}}$ is obtained from $\mathcal{H}$ by placing at the origin the (unperturbed) invariant torus of $\mathcal{H}_0$ (see (3.2.5)) with vector of basic frequencies $\boldsymbol{\Lambda}$, and arranging the corresponding normal variational equations (for the unperturbed part $\mathcal{H}_0$) to diagonal form. So, if we put $\mathcal{R} \equiv 0$ in (3.2.4), then $H^{(0)}_{\boldsymbol{\Lambda}}$ constitutes a family of analytic Hamiltonians having at the origin an invariant torus with vector of basic frequencies $\boldsymbol{\Lambda}$ (Ω_1 is known as a function of $\boldsymbol{\Lambda}$ through (3.2.22)).

[3] To check this transformation, it is useful to take into account that the matrix $\mathcal{E}(\boldsymbol{\Lambda})$ can be expressed as $\mathcal{E}(\boldsymbol{\Lambda}) = \begin{pmatrix} w_1(\boldsymbol{\Lambda}) & -\frac{\xi(\boldsymbol{\Lambda})}{\lambda_+}w_1(\boldsymbol{\Lambda}) \\ w_2(\boldsymbol{\Lambda}) & -\frac{\xi(\boldsymbol{\Lambda})}{\lambda_+}w_2(\boldsymbol{\Lambda}) \end{pmatrix}$. See (3.2.38).

Our purpose is to perform a sequence of canonical transformations on $H_{\boldsymbol{\Lambda}}^{(0)}$, depending on the parameter $\boldsymbol{\Lambda}$, and to analyze for which values of $\boldsymbol{\Lambda}$ (in a Cantor set) we can overcome the influence of the non-integrable remainder $\mathcal{R}$ to obtain a limit Hamiltonian having also an invariant torus at the origin with vector of basic frequencies $\boldsymbol{\Lambda}$ (of course, with a different Ω_1). The "regularity" with respect to $\boldsymbol{\Lambda}$ of the successive transformed Hamiltonians is important, because it will be used to control the measure of "bad" and "good" parameters $\boldsymbol{\Lambda}$ along the iterative process. In Jorba and Villanueva (1997a) it was used Lipschitz dependence (which is enough for measure purposes, see also Jorba and Villanueva, 1997b). However, we have preferred here to follow the original idea in Arnol'd (1963a,b) and consider analytic dependence with respect to $\boldsymbol{\Lambda}$. This forces us to consider a KAM process with an ultra-violet cut-off (see section 3.3 for details). We point out that, in spite of this analytic dependence, the limit Hamiltonian will be defined in a Cantor set of the basic frequencies $\boldsymbol{\Lambda}$ and so the regularity is no longer analytic; in fact, is C^∞ in the sense of Whitney (see Broer, Huitema and Sevryuk, 1996, for a more precise description).

In few words, to begin we first select a frequency, expand the Hamiltonian around the corresponding initial torus and, second, we introduce a linear canonical change to diagonalize (the variational equations in the normal directions). These two steps are equivalent to directly apply the symplectic transformation $(\theta_1, \theta_2, x, \widehat{I}_1, \widehat{I}_2, y) \to (\theta_1, \theta_2, q, I_1, I_2, p)$, given by:

$$q = \xi + x - \frac{\xi}{\lambda_+} y, \quad I_1 = \mathcal{I}(\boldsymbol{\zeta}) + \widehat{I}_1, \quad I_2 = 2\xi\eta + \widehat{I}_2, \quad p = \frac{\lambda_+}{2\xi} x + \frac{1}{2} y, \tag{3.2.34}$$

where $\lambda_+ = \mathrm{i}\mu$ (see (3.2.19)) and, though it has not been written explicitly in the equations above, the parameters ξ, η must be thought of as functions of the basic frequencies μ and Ω_2. Note also, that the "diagonalizing" change,

$$\begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\xi}{\lambda_+} \\ \frac{\lambda_+}{2\xi} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{3.2.35}$$

is actually a complexification of the real Hamiltonian (3.2.4): the normal variables q, p become complex even though x and y are real, for the transformation involves λ_+, which is a purely imaginary quantity. Even though the tori we finally obtain will be real tori (see remark 3.29 for more details). In this way, one obtains a family of Hamiltonians which can be cast into,

$$\begin{aligned} H_{\boldsymbol{\Lambda}}^{(0)}(\boldsymbol{\theta}, x, \widehat{\boldsymbol{I}}, y) = {} & \phi^{(0)}(\boldsymbol{\Lambda}) + \langle \boldsymbol{\Omega}^{(0)}, \widehat{\boldsymbol{I}} \rangle + \frac{1}{2}\langle \boldsymbol{z}, \mathcal{B}(\boldsymbol{\Lambda})\boldsymbol{z} \rangle + \frac{1}{2}\langle \widehat{\boldsymbol{I}}, \mathcal{C}^{(0)}(\boldsymbol{\Lambda})\widehat{\boldsymbol{I}} \rangle \\ & + \langle \widehat{\boldsymbol{I}}, \mathcal{E}(\boldsymbol{\Lambda})\boldsymbol{z} \rangle + H_*^{(0)}(x, \widehat{\boldsymbol{I}}, y; \boldsymbol{\Lambda}) + \widehat{H}^{(0)}(\boldsymbol{\theta}, x, \widehat{\boldsymbol{I}}, y; \boldsymbol{\Lambda}). \end{aligned} \tag{3.2.36}$$

Here, $\boldsymbol{z}^* = (x, y)$, $\boldsymbol{\zeta} = \boldsymbol{\zeta}(\boldsymbol{\Lambda})$, $H_*^{(0)}$ holds terms of order greater than two (these are the terms in the normal form that we do not write explicitly), $\widehat{H}^{(0)}$ is the transformed of the remainder $\mathcal{R}$ (see (3.2.4)), whereas

$$\phi^{(0)} = \mathcal{H}_0|_{\mathcal{T}_{\boldsymbol{\zeta}(\boldsymbol{\Lambda})}}, \qquad \Omega_i^{(0)} = \partial_{I_i}\mathcal{H}_0|_{\mathcal{T}_{\boldsymbol{\zeta}(\boldsymbol{\Lambda})}}, \qquad \mathcal{C}_{i,j}^{(0)}(\boldsymbol{\Lambda}) = \partial^2_{I_i, I_j}\mathcal{H}_0\Big|_{\mathcal{T}_{\boldsymbol{\zeta}(\boldsymbol{\Lambda})}} \tag{3.2.37}$$

(for $i, j = 1, 2$), and the matrices $\mathcal{B}$, $\mathcal{E}$ given by,

$$\mathcal{B}(\boldsymbol{\Lambda}) = \begin{pmatrix} 0 & \lambda_+ \\ \lambda_+ & 0 \end{pmatrix}, \quad \mathcal{E}(\boldsymbol{\Lambda}) = \begin{pmatrix} \partial^2_{I_1,q}\mathcal{H}_0\big|_{\mathcal{T}_{\zeta(\boldsymbol{\Lambda})}} & -\frac{\xi(\boldsymbol{\Lambda})}{\lambda_+}\,\partial^2_{I_1,q}\mathcal{H}_0\big|_{\mathcal{T}_{\zeta(\boldsymbol{\Lambda})}} \\ \partial^2_{I_2,q}\mathcal{H}_0\big|_{\mathcal{T}_{\zeta(\boldsymbol{\Lambda})}} & -\frac{\xi(\boldsymbol{\Lambda})}{\lambda_+}\,\partial^2_{I_2,q}\mathcal{H}_0\big|_{\mathcal{T}_{\zeta(\boldsymbol{\Lambda})}} \end{pmatrix}. \tag{3.2.38}$$

We point out here that, if one considers $H^{(0)}_{\boldsymbol{\Lambda}}$ for a fixed $\boldsymbol{\Lambda}$ and skips the remainder $\widehat{H}^{(0)}$, then $\widehat{\boldsymbol{I}} = 0$ and $\boldsymbol{z} = 0$ corresponds to an invariant two-dimensional elliptic torus with vector of basic frequencies $\boldsymbol{\Lambda}$ and reducible normal variational flow given by the (complex) diagonal matrix $J_1\mathcal{B}$. However, we also remark that, here, the "neutral" and normal directions are coupled through the matrix $\mathcal{E}$. In fact, we can eliminate $\mathcal{E}$ by means of the symplectic change described in remark 3.10 –equations (3.2.31) and (3.2.32)–, but due to the close-to degeneracy of the problem it is then difficult to control the domains where the change is well-defined. Hence, we shall keep $\mathcal{E}$ through the iterative process even though it complicates the solvability of the homological equations. Moreover, note that, for the change (3.2.34) to be well-defined and analytic, we need that $\lambda_+ \neq 0$. As we are assuming $a > 0$, then, using the reality condition $\xi > 0$ it is not difficult to check (see (3.2.19)) that, if R is small enough, then the expression inside the square root defining λ_+ is positive. However, as we want to work iteratively with analytic dependence with respect to the basic frequencies, we are forced to complexify ξ, and to allow values of ξ with $\mathrm{Re}\,\xi > 0$ and $\mathrm{Im}\,\xi$ small. This complicates the control of the non-vanishing character of λ_+, which needs additional assumptions on the size of $\mathrm{Re}\,\xi$ and $\mathrm{Im}\,\xi$.

Let, as in the statement of theorem 3.9, $M(R)$ be the bound for the remainder computed in chapter 2 (see theorem 2.29). Thus $|\mathfrak{R}|_{\rho,R} \leq M(R)$. Suppose now that we select a domain in the (complex) space of the basic frequencies μ, Ω_2, such that the corresponding values of the parameters ξ, η satisfy:

$$32M^\alpha \leq \mathrm{Re}\,\xi \leq \frac{\widehat{c}}{8}R^2, \qquad |\mathrm{Re}\,\eta| \leq \frac{\check{c}}{2}R, \tag{3.2.39}$$

and for their imaginary parts

$$|\mathrm{Im}\,\xi| \leq 4M^\alpha, \qquad |\mathrm{Im}\,\eta| \leq 4M^\alpha, \tag{3.2.40}$$

(these bounds are motivated by the sets introduced in page 104) where $0 < \alpha < 1$ is the exponent introduced in (3.2.25), and fixed later, in section 3.4. Then, in the domain defined by (3.2.39) and (3.2.40), we can control λ_+ in the following way.

Lemma 3.11. *Consider the definition of λ_+ in* (3.2.19) *with the coefficient $a > 0$ and values of ξ, η satisfying the restrictions* (3.2.39) *and* (3.2.40). *Then, if R is small enough, the square root defining λ_+ is (taking the appropriate determination) well-defined and the following bounds hold:*

$$\left|\frac{\xi}{\lambda_+}\right| \leq \chi_1|\xi|^{\frac{1}{2}}, \qquad \left|\frac{\eta}{\lambda_+}\right| \leq \chi_2, \qquad \chi_3 M^{\alpha/2} \leq |\lambda_+| \leq \chi_4 R, \tag{3.2.41}$$

where the constants χ_1, χ_2, χ_3 and χ_4 do not depend on R.

Proof. Let us define:

$$\breve{h}(\xi,\eta) := \partial_{1,1}\mathcal{Z}_3(\xi,\mathcal{I}(\xi,\eta),2\xi\eta),$$

$$\breve{g}(\xi,\eta) := 4\eta^2 + 2a\xi + 2\xi\breve{h}(\xi,\eta).$$

Then, we have:

$$\lambda_+ = +\mathrm{i}\sqrt{\breve{g}(\xi,\eta)},$$

where $\sqrt{\cdots}$ means now the principal determination of the square root. Moreover:

$$\begin{aligned}\mathrm{Re}\,\breve{g}(\xi,\eta) &= 4\left(\mathrm{Re}\,\eta\right)^2 + 2a\left(\mathrm{Re}\,\xi\right) + 2\mathrm{Re}\,\left(\breve{h}(\mathrm{Re}\,\xi,\mathrm{Re}\,\eta)\right)\mathrm{Re}\,\xi \\ &\quad - 4\left(\mathrm{Im}\,\eta\right)^2 + 2\mathrm{Re}\,\left(\breve{h}(\xi,\eta) - \breve{h}(\mathrm{Re}\,\xi,\mathrm{Re}\,\eta)\right)\mathrm{Re}\,\xi - 2\mathrm{Im}\,\left(\breve{h}(\xi,\eta)\right)\mathrm{Im}\,\xi,\end{aligned}$$

and

$$\mathrm{Im}\,\breve{g}(\xi,\eta) = 8\mathrm{Re}\,\eta\,\mathrm{Im}\,\eta + 2a\mathrm{Im}\,\xi + 2\breve{h}(\xi,\eta)\mathrm{Im}\,\xi + 2\mathrm{Im}\,\left(\breve{h}(\xi,\eta) - \breve{h}(\mathrm{Re}\,\xi,\mathrm{Re}\,\eta)\right)\mathrm{Re}\,\xi.$$

So, using that (see (3.2.10)):

$$\|\partial^2_{i,j}\mathcal{Z}_3\|_{R^2/2} \le \widehat{\mathcal{Z}}_{3;2}R^2, \qquad \|\partial^3_{i,j,k}\mathcal{Z}_3\|_{R^2/2} \le \widehat{\mathcal{Z}}_{3;3}, \quad \text{with } i,j,k = 1,2,3,$$

and that,

$$|\mathcal{I}(\xi,\eta)| \le \frac{R^2}{2}, \qquad \text{if } \ (\xi,\eta)\in\Gamma^*$$

(see lemma 3.2), one has –for R small enough–, the following inequalities:

$$|\mathrm{Re}\,\breve{g}(\xi,\eta)| \ge a\mathrm{Re}\,\xi + 4\left(\mathrm{Re}\,\eta\right)^2 - c_{10}RM^\alpha(R),$$

$$|\mathrm{Im}\,\breve{g}(\xi,\eta)| \le 2a|\mathrm{Im}\,\xi| + c_{11}RM^\alpha(R),$$

where the constants c_{10} and c_{11} are independent of R. Thus, it is clear that

$$\sqrt{\breve{g}(\xi,\eta)} = \sqrt{\mathrm{Re}\,\breve{g}(\xi,\eta)} \times \sqrt{1 + \mathrm{i}\tfrac{\mathrm{Im}\,\breve{g}(\xi,\eta)}{\mathrm{Re}\,\breve{g}(\xi,\eta)}},$$

where $|\frac{\mathrm{Im}\,\breve{g}}{\mathrm{Re}\,\breve{g}}| \le \frac{1}{2}$, if one takes R sufficiently small. From here:

$$\left|\frac{\xi}{\lambda_+}\right| = |\xi|^{1/2}\sqrt{\frac{|\xi|}{|\breve{g}|}} \le |\xi|^{1/2}\sqrt{\frac{|\mathrm{Re}\,\xi| + |\mathrm{Im}\,\xi|}{\mathrm{Re}\,\breve{g}(\xi,\eta)}} \le |\xi|^{1/2}\sqrt{\frac{\mathrm{Re}\,\xi + |\mathrm{Im}\,\xi|}{a\mathrm{Re}\,\xi - c_{10}RM^\alpha(R)}} \le \chi_1|\xi|^{1/2}$$

again, if R small enough. In the same way we can derive the inequality $|\frac{\eta}{\lambda_+}| \le \chi_2$ and also,

$$\begin{aligned}|\breve{g}(\xi,\eta)| &\ge |\mathrm{Re}\,\breve{g}(\xi,\eta)| \\ &\ge 4(\mathrm{Re}\,\eta)^2 + a\,\mathrm{Re}\,\xi - c_{10}RM^\alpha(R) \\ &\ge \mathit{constant}\cdot M^\alpha(R) \qquad \text{(if } R \text{ small enough)}\end{aligned}$$

and

$$|\breve{g}(\xi,\eta)| \le \mathit{constant}'\cdot R^2, \qquad \text{(if } R \text{ small enough)}$$

from which inequalities $|\lambda_+| \ge \chi_3 M^{\alpha/2}$ and $|\lambda_+| \le \chi_4 R$ follow. This closes the proof of lemma 3.11. □

Remark 3.12. We note that, if one takes $\boldsymbol{\zeta}^* = (\xi, \eta) \in \Gamma(R)$, with $\xi > 0$ (and $a > 0$) the inequalities,

$$\left|\frac{\xi}{\lambda_+}\right| \leq \chi_1 \xi^{\frac{1}{2}}, \qquad \left|\frac{\eta}{\lambda_+}\right| \leq \chi_2, \qquad |\lambda_+| \leq \chi_4 R \tag{3.2.42}$$

can equally be proved, provided $0 < R < 1$ is taken sufficiently small. We stress that no lower bound on ξ are required for the above inequalities to hold. If further, one asks for example $\xi > 16 M^\alpha(R)$, then,

$$|\lambda_+| \geq \chi_3 M^{\alpha/2}(R) \tag{3.2.43}$$

(again, for R small enough). ♠

Let, from now on, $M^{(0)}(R)$ and $\rho^{(0)}$ denote:

$$M^{(0)}(R) \equiv M(R), \qquad \rho^{(0)} = \rho_0, \tag{3.2.44}$$

with ρ_0 the one given in (3.2.25). Then, we can formulate the next lemma, which states that the norm of the "perturbative" term $\widehat{H}^{(0)}$ of the transformed Hamiltonian (3.2.36) can –in $\mathcal{D}_*(\rho^{(0)}, R^{(0)})$– be estimated using the same bounds obtained in chapter 2 for the remainder $\mathfrak{R}$ in the "original" Hamiltonian (3.2.1).

Lemma 3.13. *Taking* $\boldsymbol{\zeta} \in \widetilde{\Gamma}^*(R)$ *satisfying the restrictions* (3.2.39), (3.2.40) *and assuming* R *small enough such that*

$$M^\alpha(R) \leq \frac{1}{4} R, \tag{3.2.45}$$

the remainder $\widehat{H}^{(0)}$ *of the Hamiltonian* $H^{(0)}$ *in* (3.2.36) *is also bounded by* $M^{(0)}(R) \equiv M(R)$*, i. e.:*

$$|\widehat{H}^{(0)}|_{\rho^{(0)}, R^{(0)}} \leq M^{(0)}(R), \tag{3.2.46}$$

provided $R < 1$ *is sufficiently small.*

Proof. Taking into account the symplectic changes (3.2.2) and (3.2.34) one may write the coordinates q_1, q_2, I_1, p_1, p_2 in function of θ_2, x, $\widehat{I}_2$, y and also depending on the parameters ξ and η. Writing them up explicitly,

$$\begin{aligned}
q_1 &= \mathrm{e}^{\mathrm{i}\theta_2}\sqrt{\xi + x - \frac{\xi}{\lambda_+}y}, \\
q_2 &= \mathrm{e}^{\mathrm{i}\theta_2}\sqrt{\xi + x - \frac{\xi}{\lambda_+}y}\left(\frac{\lambda_+}{2\xi}x + \frac{1}{2}y\right) + \frac{\mathrm{i}\mathrm{e}^{\mathrm{i}\theta_2}(2\xi\eta + \widehat{I}_2)}{2\sqrt{\xi + x - \frac{\xi}{\lambda_+}y}}, \\
I_1 &= \mathcal{I}(\boldsymbol{\zeta}) + \widehat{I}_1, \\
p_1 &= \mathrm{e}^{-\mathrm{i}\theta_2}\sqrt{\xi + x - \frac{\xi}{\lambda_+}y}\left(\frac{\lambda_+}{2\xi}x + \frac{1}{2}y\right) - \frac{\mathrm{i}\mathrm{e}^{-\mathrm{i}\theta_2}(2\xi\eta + \widehat{I}_2)}{2\sqrt{\xi + x - \frac{\xi}{\lambda_+}y}}, \\
p_2 &= -\mathrm{e}^{-\mathrm{i}\theta_2}\sqrt{\xi + x - \frac{\xi}{\lambda_+}y}.
\end{aligned} \tag{3.2.47}$$

As ξ satisfies the restrictions (3.2.39), and (3.2.40) we have:

$$\begin{aligned}|\sqrt{\xi}| &\leq \left|\sqrt{\operatorname{Re}\xi}\right| \times \left|\sqrt{1+\mathrm{i}\frac{\operatorname{Im}\xi}{\operatorname{Re}\xi}}\right| \\ &\leq \sqrt{\operatorname{Re}\xi} \times \sqrt{1+\frac{|\operatorname{Im}\xi|}{|\operatorname{Re}\xi|}} \leq \frac{\widehat{c}^{1/2}}{\sqrt{8}} R\sqrt{1+\frac{1}{8}} = \frac{3}{8}\widehat{c}^{1/2}R,\end{aligned}$$

and then, from (3.2.41) in lemma 3.11, one derives the bounds

$$\left|x-\frac{\xi}{\lambda_+}y\right|_{0,R^{(0)}} \leq 2\left(1+\frac{3}{8}\widehat{c}^{\frac{1}{2}}\chi_1 R\right) M^\alpha(R), \tag{3.2.48a}$$

$$\begin{aligned}\left|\frac{\lambda_+}{2\xi}x+\frac{1}{2}y\right|_{0,R^{(0)}} &\leq \frac{\chi_4}{64M^\alpha}2RM^\alpha+\frac{1}{2}2M^\alpha = \left(\frac{\chi_4}{32}+\frac{M^\alpha(R)}{R}\right)R \\ &\leq \frac{1}{4}\left(\frac{\chi_4}{8}+1\right)R,\end{aligned} \tag{3.2.48b}$$

where, in the second, it was used that

$$|\xi| \geq |\operatorname{Re}\xi| \geq 32M^\alpha, \tag{3.2.49}$$

–see (3.2.41)–, together with the assumption $M^\alpha(R) \leq \frac{1}{4}R$. With the bounds (3.2.48a,b), and (3.2.49) it is straightforward now to apply lemma A.9 of appendix A to estimate the square roots appearing in transformation (3.2.47). Indeed, assuming also R small enough to make for example $\frac{3}{8}\widehat{c}^{1/2}\chi_1 R < 1$, one has:

$$\begin{aligned}\left|\sqrt{\xi+x-\frac{\xi}{\lambda_+}y}\right|_{0,R^{(0)}} &\leq \frac{3}{8}\widehat{c}^{1/2}\left(2-\sqrt{1-\frac{1}{16}\left(1+\frac{1}{8}\widehat{c}^{1/2}\chi_1 R\right)}\right)R \\ &\leq \frac{3}{8}\widehat{c}^{1/2}\left(2-\sqrt{\frac{7}{8}}\right)R,\end{aligned}$$

and

$$\begin{aligned}\left|\frac{1}{\sqrt{\xi+x-\frac{\xi}{\lambda_+}y}}\right|_{0,R^{(0)}} &\leq \frac{1/|\sqrt{\xi}|}{\sqrt{1-\frac{1}{16}\left(1+\frac{1}{8}\widehat{c}^{1/2}\chi_1 R\right)}} \\ &\leq \frac{1}{|\sqrt{\xi}|}\times\sqrt{\frac{8}{7}}.\end{aligned}$$

In addition, we shall use that

$$\begin{aligned}\left|\mathrm{e}^{\mathrm{i}\theta_2}\right|_{\rho^{(0)},0} &= \mathrm{e}^{\rho^{(0)}} \leq \mathrm{e}^{\frac{1}{2}\ln 2} = \sqrt{2}, \\ 2\left|\sqrt{\xi}\right||\eta| + \frac{\widehat{I}_2}{|\sqrt{\xi}|} &\leq \frac{3}{4}\widehat{c}^{1/2}R\left(\frac{\check{c}}{2}R+4M^\alpha(R)\right)+\frac{4M^{2\alpha}}{4\sqrt{2}M^{\alpha/2}} \\ &\leq \left\{\frac{3}{4}\widehat{c}^{1/2}\left(\frac{\check{c}}{2}R+4M^\alpha\right)+\frac{1}{4\sqrt{2}}M^{\alpha/2}\right\}R,\end{aligned}$$

(since $\rho^{(0)} = \frac{1}{2}\min\{\rho, \ln 2\} \le \frac{1}{2}\ln 2$, see (3.2.25)), to get following bounds for $|q_j|_{\rho^{(0)},R^{(0)}}$, $|p_j|_{\rho^{(0)},R^{(0)}}$, $j = 1, 2$ and $|I_1|_{\rho^{(0)},R^{(0)}}$:

$$|q_1|_{\rho^{(0)},R^{(0)}}, |p_2|_{\rho^{(0)},R^{(0)}} \le \frac{3\sqrt{2}}{8}\widehat{c}^{1/2}\left(2 - \sqrt{\frac{7}{8}}\right) R < R,$$

(we recall that is $\widehat{c} < 1$)

$$\begin{aligned} |q_2|_{\rho^{(0)},R^{(0)}}, |p_1|_{\rho^{(0)},R^{(0)}} \le & \left(\frac{\chi_4}{32} + \frac{M^\alpha}{R}\right) R^2 \\ & + \sqrt{\frac{8}{14}}\left\{\frac{3}{4}\widehat{c}^{1/2}\left(\frac{\check{c}}{2}R + 4M^\alpha\right) + \frac{1}{4\sqrt{2}}M^{\alpha/2}\right\} R < R, \\ & \qquad \text{(for } R \text{ small enough)} \end{aligned}$$

$$\begin{aligned} |I_1|_{\rho^{(0)},R^{(0)}} & \le |\mathcal{I}(\boldsymbol{\zeta})| + 4M^{2\alpha}(R) \\ & \le \frac{R^2}{2} + 4M^{2\alpha}(R) < R^2, \qquad \text{(for } R \text{ small enough)} \end{aligned}$$

where, in the last inequality it has been taken into account that $|\mathcal{I}(\boldsymbol{\zeta})| \le \frac{R^2}{2}$ whenever $\boldsymbol{\zeta}^* = (\xi, \eta) \in \Gamma^*(R)$ –see lemma 3.2– and $\widetilde{\Gamma}^*(R) \subset \Gamma^*(R)$. Therefore, we have seen that for $R < 1$ small enough,

$$|\boldsymbol{q}|_{\rho^{(0)},R^{(0)}} \le R, \qquad |I_1|_{\rho^{(0)},R^{(0)}} \le R^2, \qquad |\boldsymbol{p}|_{\rho^{(0)},R^{(0)}} \le R,$$

so, immediate application of lemma A.4 of appendix A leads to the inequality,

$$|\widehat{H}^{(0)}|_{\rho^{(0)},R^{(0)}} \le |\mathfrak{R}|_{\rho,R}$$

and hence, for $\widehat{H}^{(0)}$ we have,

$$|\widehat{H}^{(0)}|_{\rho^{(0)},R^{(0)}} \le M^{(0)}(R),$$

which was the desired result. $\square$

Some useful bounds

Prior to set up the iterative scheme, it is convenient to bound the different elements of the Hamiltonian $H^{(0)}$. To account for its dependence on the set of the frequencies, we need to introduce an appropriate norm. Given a function $f(\boldsymbol{\phi})$ defined for $\boldsymbol{\phi} \in \widetilde{\mathcal{E}}$, $\widetilde{\mathcal{E}} \subset \mathbb{C}^n$ for some n, and with values in $\mathbb{C}$, $\mathbb{C}^n$ or $\mathbb{M}_{n_1,n_2}(\mathbb{C})$, we define $||f||_{\widetilde{\mathcal{E}}} = \sup_{\boldsymbol{\phi}\in\widetilde{\mathcal{E}}} |f(\boldsymbol{\phi})|$. This definition can be extended to analytic functions depending on the parameter $\boldsymbol{\phi}$. So, if $f(\boldsymbol{\theta}, \boldsymbol{\phi})$ or $f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\phi})$ are, for every $\boldsymbol{\phi} \in \widetilde{\mathcal{E}}$, analytic, 2π-periodic on $\boldsymbol{\theta}$ and defined on the complex strip $\{\boldsymbol{\theta} \in \mathbb{C}^r : |\mathrm{Im}\,\boldsymbol{\theta}| \le \rho\}$ or $\mathcal{D}_{r,m}(\rho, R)$ (see notation in section A.1 of appendix A), we may introduce, $||f||_{\widetilde{\mathcal{E}}} = \sup_{\boldsymbol{\phi}\in\widetilde{\mathcal{E}}} |f|_\rho$ and $||f||_{\widetilde{\mathcal{E}}} = \sup_{\boldsymbol{\phi}\in\widetilde{\mathcal{E}}} |f|_{\rho,R}$ respectively. On the other hand, it is clear that all the properties of the norms $|\cdot|_\rho$, $|\cdot|_{\rho,R}$ (see appendix A) are translated directly to $||\cdot||_{\widetilde{\mathcal{E}},\rho}$ and $||\cdot||_{\widetilde{\mathcal{E}},\rho,R}$.

Let $\breve{\mathcal{E}}^{(0)} \subseteq \mathcal{N}^*(R)\backslash\{(\mu, \Omega_2) \in \mathcal{N}^* : \operatorname{Re}\mu \leq 0\}$ be the set of (complex) basic frequencies such that if $\boldsymbol{\Lambda}^* = (\mu, \Omega_2) \in \breve{\mathcal{E}}^{(0)}$, then the corresponding ξ, η through the map:

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} g(\mu, h(\mu, \Omega_2)) \\ h(\mu, \Omega_2) \end{pmatrix} \tag{3.2.50}$$

satisfy the restrictions (3.2.39) and (3.2.40). Note: here g and h are those of lemmas 3.7 and 3.8 respectively.

Therefore, if the components of the matrix $\mathcal{C}^{(0)}$ in (3.2.37) are developed explicitly, the following estimates can be gleaned:

$$\begin{aligned} ||\mathcal{C}^{(0)}_{1,1}||_{\breve{\mathcal{E}}^{(0)},\rho^{(0)}} &\leq |b| + \widehat{\mathcal{Z}}_{3;2}R^2, \\ ||\mathcal{C}^{(0)}_{1,2}||_{\breve{\mathcal{E}}^{(0)},\rho^{(0)}} &\leq |f| + \widehat{\mathcal{Z}}_{3;2}R^2, \\ ||\mathcal{C}^{(0)}_{2,2}||_{\breve{\mathcal{E}}^{(0)},\rho^{(0)}} &\leq M^{-\alpha}(1 + |c|M^\alpha + \widehat{\mathcal{Z}}_{3;2}R^2M^\alpha), \end{aligned} \tag{3.2.51}$$

and similarly for the components of $\mathcal{E}$, appearing in (3.2.38),

$$\begin{aligned} ||\mathcal{E}_{1,1}||_{\breve{\mathcal{E}}^{(0)}} &\leq |d| + \widehat{\mathcal{Z}}_{3;2}R^2, \\ ||\mathcal{E}_{1,2}||_{\breve{\mathcal{E}}^{(0)}} &\leq \chi_1\widehat{c}^{1/2}(|d| + \widehat{\mathcal{Z}}_{3;2}R^2)R, \\ ||\mathcal{E}_{2,1}||_{\breve{\mathcal{E}}^{(0)}} &\leq M^{-\alpha}(R + |e|M^\alpha + \widehat{\mathcal{Z}}_{3;2}R^2M^\alpha), \\ ||\mathcal{E}_{2,2}||_{\breve{\mathcal{E}}^{(0)}} &\leq \chi_1(1 + |e|R + \widehat{\mathcal{Z}}_{3;2}R^3), \end{aligned} \tag{3.2.52}$$

Remark 3.14. We have used the norm $||\cdot||_{\mathcal{E}^{(0)},\rho^{(0)}}$ for the constant matrix $\mathcal{C}^{(0)}$, because the matrices replacing $\mathcal{C}^{(0)}$ along the successive steps of the iterative procedure will depend on $\boldsymbol{\theta}$. ♠

Next, we are going to compute estimates of $H_*^{(0)}$. Before, let us denote by $Z_3(x, \widehat{\boldsymbol{I}}, y)$, the transformed –through the change (3.2.34)–, of the term $\mathcal{Z}_3(q, \boldsymbol{I})$ in Hamiltonian (3.2.4). Then, using lemma A.4 and the properties of the norm $||\cdot||_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}$, it can be shown that, for $R < 1$ small enough, $||Z_3(x, \widehat{\boldsymbol{I}}, y)||_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}$ is bounded by a quantity independent of R. Moreover, if $Q = x - \frac{\xi}{\lambda_+}y$, then the terms of order higher than two in the quotient $\frac{I_2^2}{4q}$ can be expressed as,

$$\begin{aligned} &\xi\eta^2\left(\left(1 + \frac{Q}{\xi}\right)^{-1} - 1 + \frac{Q}{\xi} + \left(\frac{Q}{\xi}\right)^2\right) + \\ &\qquad + \eta\widehat{I}_2\left(\left(1 + \frac{Q}{\xi}\right)^{-1} - 1 + \frac{Q}{\xi}\right) + \frac{\widehat{I}_2^2}{4\xi}\left(\left(1 + \frac{Q}{\xi}\right)^{-1} - 1\right), \end{aligned} \tag{3.2.53}$$

but, $\|Q/\xi\|_{\mathcal{E}^{(0)},R^{(0)}} \leq (1 + \chi_1 R)/8$, and if R is sufficiently small, $\|1 + Q/\xi\|_{\mathcal{E}^{(0)},R^{(0)}} \geq 3/4$; hence (3.2.53) is, for R small enough, bounded by a quantity independent of R. Of course the same can be said for the terms in $H_*^{(0)}$ coming from the product qp^2 in (3.2.4) and consequently for $H_*^{(0)}$. To summarize, the foregoing arguments show the existence –for R

sufficiently small–, of positive R-independent constants: $\widehat{\alpha}_1 < \widehat{\alpha}_2$, $\widehat{m}_1^{(0)}$, $\widehat{m}_2$, and $\widehat{\nu}^{(0)}$ such that,

$$\widehat{\alpha}_1 \left(M^{(0)}\right)^{\alpha} \leq |\lambda_+| \leq \widehat{\alpha}_2, \tag{3.2.54}$$

$$||\mathcal{C}^{(0)}||_{\check{\mathcal{E}}^{(0)},\rho^{(0)}} \leq \widehat{m}_1^{(0)}(M^{(0)})^{-\alpha}, \tag{3.2.55}$$

$$||\mathcal{E}||_{\check{\mathcal{E}}^{(0)}} \leq \widehat{m}_2(M^{(0)})^{-\alpha}, \tag{3.2.56}$$

$$||H_*^{(0)}||_{\check{\mathcal{E}}^{(0)},\rho^{(0)},R^{(0)}} \leq \widehat{\nu}^{(0)}. \tag{3.2.57}$$

3.3 The iterative scheme

Now, we are going to describe the iterative procedure used to construct those quasi-periodic solutions (in particular two-dimensional tori) we are interested in. It is based on the early idea of Kolmogorov, see Kolmogorov (1979); Benettin et al. (1984), or chapter 5, sec. §2 in Arnol'd (1988) for an overall description –though the method outlined in this last reference actually corresponds to the Arnol'd (1963a,b) construction–. Thus, taking the initial Hamiltonian $H^{(0)}$, we apply a sequence of canonical changes given by the time-one flow of a suitable generating function S (see section B.3), in such a way that it removes from the Hamiltonian those terms obstructing the existence of the invariant reducible 2D-torus, with a vector of basic frequencies given by $\boldsymbol{\Lambda}$. But that is possible because, quoting Arnol'd (1988)–: "this procedure of successive coordinate transformation actually possesses the remarkable property of quadratic convergence". Let us now show how these general ideas are developed in the problem on hand: for this purpose we proceed to describe a generic step of the iterative process. Consider thus a Hamiltonian $H^{(0)}$ of the form (3.2.36), and a generating function S (to be determined). Furthermore, we shall split $\widehat{H}^{(0)}$ –i. e., the piece of $H^{(0)}$ holding the "obstructing terms"—, into:

$$\widehat{H}^{(0)} = \widehat{H}^{(0)}_{<N,\boldsymbol{\theta}} + \widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}},$$

where $\widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}}$ stands for the terms in Taylor-Fourier expansion of $\widehat{H}^{(0)}$ with harmonics of order $|\boldsymbol{k}|_1 \geq N$. To bound the contribution of this truncation in the size of $\widehat{H}^{(1)} = \widehat{H}^{(0)} \circ \psi_1^S$ (see below), we point that, for $0 < \delta < \rho^{(0)}$,

$$\|\widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}}\|_{\check{\mathcal{E}}^{(0)},\rho^{(0)}-\delta,R^{(0)}} \leq \|\widehat{H}^{(0)}\|_{\check{\mathcal{E}}^{(0)},\rho^{(0)},R^{(0)}} \exp(-\delta N). \tag{3.3.1}$$

(which follows from the exponential decay of the Fourier coefficients). If, *plus encore*, we define

$$\widetilde{H}^{(0)} = H^{(0)} - \widehat{H}^{(0)}, \qquad \check{H}_1^{(0)} = \{H^{(0)} - \widehat{H}^{(0)}, S\}, \qquad \check{H}_2^{(0)} = \{\widehat{H}^{(0)}, S\}. \tag{3.3.2}$$

Then, by corollary B.22:

$$H^{(1)} = H^{(0)} - \widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}} + \{\widetilde{H}^{(0)}, S\} + \widehat{H}^{(1)}, \tag{3.3.3}$$

being,

$$\widehat{H}^{(1)} = \widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}} + \int_0^1 \left(\check{H}_2^{(0)} + (1-t)\{\check{H}_1^{(0)}, S\}\right) \circ \psi_t^S \, dt \tag{3.3.4}$$

and we take N large enough to make $\widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}}$ "of the same order" as the integral term. On the other hand, the generating function S, must be such that the $H^{(1)}$ will take the same form as $H^{(0)}$ in (3.2.36), i. e., it must satisfy

$$\begin{aligned} H^{(0)} - \widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}} + \left\{\widetilde{H}^{(0)}, S\right\} = \\ = \phi^{(1)} + \langle \boldsymbol{\Omega}^{(1)}, \boldsymbol{I}\rangle + \frac{1}{2}\langle \boldsymbol{z}, \mathcal{B}\boldsymbol{z}\rangle + \frac{1}{2}\left\langle \boldsymbol{I}, \mathcal{C}^{(1)}(\boldsymbol{\theta})\boldsymbol{I}\right\rangle + \langle \boldsymbol{I}, \mathcal{E}\boldsymbol{z}\rangle + H^{(1)}_*, \end{aligned} \tag{3.3.5}$$

(the hat in the action variables has been dropped). Moreover, note that the matrices $\mathcal{B}$, $\mathcal{E}$ are left invariant, whereas $\boldsymbol{\Omega}^{(1)*} = (\Omega^{(1)}_1, \Omega^{(0)}_2)$. Now, to write the homological equations in order to determine the generating function, S we expand the Hamiltonian $H^{(0)}$ in the form,

$$\begin{aligned} H^{(0)} &= a_*(\boldsymbol{\theta}) + \langle \boldsymbol{b}(\boldsymbol{\theta}), \boldsymbol{z}\rangle + \langle \boldsymbol{c}(\boldsymbol{\theta}), \boldsymbol{I}\rangle + \frac{1}{2}\langle \boldsymbol{z}, B(\boldsymbol{\theta})\boldsymbol{z}\rangle + \langle \boldsymbol{I}, E(\boldsymbol{\theta})\boldsymbol{z}\rangle \\ &+ \frac{1}{2}\langle \boldsymbol{I}, C(\boldsymbol{\theta})\boldsymbol{I}\rangle + \mho(\boldsymbol{\theta}, x, \boldsymbol{I}, y), \end{aligned} \tag{3.3.6}$$

where we have not written explicitly neither the $\boldsymbol{\Lambda}$-dependence nor the (0) superscript in the different parts of the Hamiltonian. From this last expansion, we introduce the following notations $[H^{(0)}]_{(\boldsymbol{z},\boldsymbol{z})} = B$, $[H^{(0)}]_{(\boldsymbol{I},\boldsymbol{z})} = E$ and, given a function $f(\boldsymbol{\theta}, x, \boldsymbol{I}, y)$,

$$\langle f\rangle_{\boldsymbol{\theta}} = \tfrac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} f(\boldsymbol{\theta}, x, \boldsymbol{I}, y)\, d\theta_1 d\theta_2, \tag{3.3.7}$$

$$\{f\}_{\boldsymbol{\theta}} = f - \langle f\rangle_{\boldsymbol{\theta}}. \tag{3.3.8}$$

From the decomposition of Hamiltonian $H^{(0)}$ in (3.2.36), we have that $\{a_*\}_{\boldsymbol{\theta}}$, $\boldsymbol{b}$, $\boldsymbol{c} - \boldsymbol{\Omega}$, $B - \mathcal{B}$, $C - \mathcal{C}^{(0)}$ and $E - \mathcal{E}$ are all of order $\mathrm{O}(\widehat{H}^{(0)})$. So, if starting from Hamiltonian $H^{(0)}$, we were able to remove $\{a_*\}_{\boldsymbol{\theta}}$, $\boldsymbol{b}$, $\{c_1\}_{\boldsymbol{\theta}}$, $c_2 - \Omega_2$, keeping the matrix $\mathcal{B}$ unchanged, then we shall obtain an invariant torus with the same basic frequency $\boldsymbol{\Lambda}$ (additionally, we want the matrix $\mathcal{E}$ to be held fixed). Hence, we shall transform $H^{(0)}$, iteratively, by the successive application of generating functions of the form

$$S(\boldsymbol{\theta}, x, \boldsymbol{I}, y) = \langle \boldsymbol{\xi}, \boldsymbol{\theta}\rangle + d_*(\boldsymbol{\theta}) + \langle \boldsymbol{e}(\boldsymbol{\theta}), \boldsymbol{z}\rangle + \langle \boldsymbol{f}(\boldsymbol{\theta}), \boldsymbol{I}\rangle + \frac{1}{2}\langle \boldsymbol{z}, G(\boldsymbol{\theta})\boldsymbol{z}\rangle + \langle \boldsymbol{I}, F(\boldsymbol{\theta})\boldsymbol{z}\rangle, \tag{3.3.9}$$

where $\boldsymbol{\xi} \in \mathbb{C}^2$, $\langle d_*\rangle_{\boldsymbol{\theta}} = 0$, $\langle \boldsymbol{f}\rangle_{\boldsymbol{\theta}} = \boldsymbol{0}$ and G is a symmetric matrix. The transformed Hamiltonian is $H^{(1)} = H^{(0)} \circ \phi^S_1$. Now, we may expand $H^{(1)}$ in identical way as $H^{(0)}$, keeping the same name of the variables, but adding the superscript (1) to a_*, $\boldsymbol{b}$, $\boldsymbol{c}$, B, C, E and $\mho$. Then, we ask $\{a^{(1)}_*\}_{\boldsymbol{\theta}} = 0$, $\boldsymbol{b}^{(1)} = 0$, $\{c^{(1)}_1\}_{\boldsymbol{\theta}} = 0$, $c^{(2)}_2 - \Omega_2 = 0$, $E^{(1)} = \mathcal{E}$, $B^{(1)} = \mathcal{B}$. This is not possible in general, but can be achieved up to first order in the size of $\widehat{H}^{(0)}$. To show how, we write those conditions in terms of the initial Hamiltonian and the generating function, S. This leads to the following equations:

$$\{a_*\}_{\boldsymbol{\theta}} - \left\langle \left(\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right)^*, \boldsymbol{\Omega} \right\rangle = 0, \tag{eq$_1$}$$

$$\boldsymbol{b} - \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{\theta}}\boldsymbol{\Omega} - \mathcal{E}^*\boldsymbol{\xi} - \mathcal{E}^*\left(\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right)^* + \mathcal{B}J\boldsymbol{e} = 0, \tag{eq$_2$}$$

$$\begin{pmatrix} \{c_1\}_{\boldsymbol{\theta}} + \delta_* \\ c_2 - \Omega_2 \end{pmatrix} - \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}}\boldsymbol{\Omega} + \mathcal{C}^{(0)}\left(\boldsymbol{\xi} + \left(\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right)^*\right) + \mathcal{E}J\boldsymbol{e} = 0, \tag{eq$_3$}$$

$$\widetilde{B}(\boldsymbol{\theta}) - \mathcal{B} - \frac{\partial G}{\partial \boldsymbol{\theta}}\boldsymbol{\Omega} + \mathcal{B}JG - GJ\mathcal{B} = 0, \tag{eq$_4$}$$

$$\widetilde{E}(\boldsymbol{\theta}) - \mathcal{E} - \frac{\partial F}{\partial \boldsymbol{\theta}}\boldsymbol{\Omega} - FJ\mathcal{B} = 0, \tag{eq$_5$}$$

with the following convention:

$$\left(\frac{\partial G}{\partial \boldsymbol{\theta}}\boldsymbol{\Omega}\right)_{j,m} = \sum_{n=1,2} \frac{\partial G_{j,m}}{\partial \theta_n}\Omega_n,$$

(the same for $\left(\frac{\partial F}{\partial \boldsymbol{\theta}}\boldsymbol{\Omega}\right)_{j,m}$, $j, m = 1, 2$). (eq$_1$)–(eq$_5$) are the *homological* equations, where,

$$\delta_* \equiv -\left\langle \mathcal{C}^{(0)}_{1,1}\left(\xi_1 + \frac{\partial d_*}{\partial \theta_1}\right)\right\rangle_{\boldsymbol{\theta}} - \left\langle \mathcal{C}^{(0)}_{1,2}\left(\xi_2 + \frac{\partial d_*}{\partial \theta_2}\right)\right\rangle_{\boldsymbol{\theta}} - \mathcal{E}_{1,1}\langle e_2\rangle_{\boldsymbol{\theta}} + \mathcal{E}_{1,2}\langle e_1\rangle_{\boldsymbol{\theta}}, \tag{3.3.10}$$

and $\widetilde{B}$, $\widetilde{E}$ are the matrices given by,

$$\widetilde{B} = B - \mathcal{E}^*\left(\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{\theta}}\right)^* - \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{\theta}}\mathcal{E} - \left[\frac{\partial H_*^{(0)}}{\partial \boldsymbol{I}}\left(\boldsymbol{\xi} + \left(\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right)^*\right) - \frac{\partial H_*^{(0)}}{\partial \boldsymbol{z}}J\boldsymbol{e}\right]_{(\boldsymbol{z},\boldsymbol{z})}, \tag{3.3.11}$$

$$\widetilde{E} = E - \mathcal{C}^{(0)}\left(\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{\theta}}\right)^* - \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}}\mathcal{E} + \mathcal{E}JG - \left[\frac{\partial H_*^{(0)}}{\partial \boldsymbol{I}}\left(\boldsymbol{\xi} + \left(\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right)^*\right) - \frac{\partial H_*^{(0)}}{\partial \boldsymbol{z}}J\boldsymbol{e}\right]_{(\boldsymbol{I},\boldsymbol{z})}. \tag{3.3.12}$$

Remark 3.15. Note that, if S is defined from the solutions of these equations, (eq$_1$)–(eq$_5$), then it verifies (3.3.5) if we skip the term $\widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}}$. If we want (3.3.5) to hold with the cut-off in the expansion with respect to θ, we just have to replace the expression of $H^{(0)}$ in (3.3.6) by the analogous expression for $H^{(0)} - \widehat{H}^{(0)}_{\geq N,\boldsymbol{\theta}}$. Moreover, by the linearity (in the harmonics of S) of (eq$_1$)–(eq$_5$), it is easy to realize that (ignoring for a moment the small divisors involved), it should be $S = \mathrm{O}(\widehat{H}^{(0)})$. Then, the integral term of (3.3.4) is clearly of second order in $\widehat{H}^{(0)}$, whereas $H^{(0)}_{*\geq N,\boldsymbol{\theta}}$ can be made –taking a large enough number of harmonics, see (3.3.1)–, of the same order of smallness. Thus, $\widehat{H}^{(1)} = \mathrm{O}((\widehat{H}^{(0)})^2)$. Hence, the quadratic convergence of the method follows. It is precisely this property which makes possible the control of the small denominators (on a suitable nonresonance set). ♠

Prior to solve completely the homological equations, we want to discuss how to fix the vector $\boldsymbol{\xi}$, as this is the most involved item when solving such equations. We point out that the vector $\boldsymbol{\xi}$ is used to adjust the "averaged part" of the equations, assuring the compatibility of the system when $\boldsymbol{\xi}$ is chosen appropriately. As we want Ω_2 and μ not to change from one iterate to another, then $\boldsymbol{\xi}$ must satisfy the linear system formed by the averages of: (eq_2), the second component of (eq_3) and the first row, second column component of the (matricial) equation (eq_4). After developing those terms in the brackets $[\cdots]_{(\boldsymbol{z},\boldsymbol{z})}$ and $[\cdots]_{(\boldsymbol{I},\boldsymbol{z})}$ of (3.3.11) and (3.3.12) respectively, one obtains, explicitly,

$$\langle \boldsymbol{b}\rangle_{\boldsymbol{\theta}} - \mathcal{E}^*\boldsymbol{\xi} - \mathcal{B}J\langle \boldsymbol{e}\rangle_{\boldsymbol{\theta}} = 0, \tag*{$\overline{(\text{eq}_2)}$}$$

$$\begin{aligned}&\langle c_2\rangle_{\boldsymbol{\theta}} - \Omega_2 + \mathcal{E}_{2,1}\langle e_2\rangle_{\boldsymbol{\theta}} - \mathcal{E}_{2,2}\langle e_1\rangle_{\boldsymbol{\theta}}\\ &\qquad - \left\langle \mathcal{C}^{(0)}_{2,1}\left(\xi_1 + \frac{\partial d_*}{\partial\theta_1}\right)\right\rangle_{\boldsymbol{\theta}} - \left\langle \mathcal{C}^{(0)}_{2,2}\left(\xi_2 + \frac{\partial d_*}{\partial\theta_2}\right)\right\rangle_{\boldsymbol{\theta}} = 0,\end{aligned} \tag*{$\overline{(\text{eq}_3)}_2$}$$

$$\begin{aligned}&\langle B_{1,2}\rangle_{\boldsymbol{\theta}} - \mathcal{B}_{1,2} - \left\langle \frac{\partial^3 H^{(0)}_*}{\partial I_1\partial x\partial y}(\boldsymbol{\theta},\boldsymbol{0})\left(\xi_1 + \frac{\partial d_*}{\partial\theta_1}\right)\right\rangle_{\boldsymbol{\theta}} - \left\langle \frac{\partial^3 H^{(0)}_*}{\partial I_2\partial x\partial y}(\boldsymbol{\theta},\boldsymbol{0})\left(\xi_2 + \frac{\partial d_*}{\partial\theta_2}\right)\right\rangle_{\boldsymbol{\theta}}\\ &\quad + \left\langle \frac{\partial^3 H^{(0)}_*}{\partial x^2\partial y}(\boldsymbol{\theta},\boldsymbol{0})\right\rangle_{\boldsymbol{\theta}}\langle e_2\rangle_{\boldsymbol{\theta}} - \left\langle \frac{\partial^3 H^{(0)}_*}{\partial x\partial y^2}(\boldsymbol{\theta},\boldsymbol{0})\right\rangle_{\boldsymbol{\theta}}\langle e_1\rangle_{\boldsymbol{\theta}}\\ &\qquad + \left\langle \left\{\frac{\partial^3 H^{(0)}_*}{\partial x^2\partial y}(\boldsymbol{\theta},\boldsymbol{0})\right\}_{\boldsymbol{\theta}}\{e_2\}_{\boldsymbol{\theta}}\right\rangle_{\boldsymbol{\theta}} - \left\langle \left\{\frac{\partial^3 H^{(0)}_*}{\partial x\partial y^2}(\boldsymbol{\theta},\boldsymbol{0})\right\}_{\boldsymbol{\theta}}\{e_1\}_{\boldsymbol{\theta}}\right\rangle_{\boldsymbol{\theta}} = 0.\end{aligned} \tag*{$\overline{(\text{eq}_4)}_{1,2}$}$$

Hence, we get the linear system,

$$\langle \mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}\begin{pmatrix}\boldsymbol{\xi}\\ \langle \boldsymbol{e}\rangle_{\boldsymbol{\theta}}\end{pmatrix} = \mathfrak{h} \tag{3.3.13}$$

(note that, if solvable, the system also furnishes the values of $\langle e_1\rangle_{\boldsymbol{\theta}}$ and $\langle e_2\rangle_{\boldsymbol{\theta}}$) where the matrix $\mathcal{A}^{(0)}$ is given by,

$$\mathcal{A}^{(0)} = \begin{pmatrix} -\mathcal{E}_{1,1} & -\mathcal{E}_{2,1} & -\lambda_+ & 0\\ -\mathcal{E}_{1,2} & -\mathcal{E}_{2,2} & 0 & \lambda_+\\ -\mathcal{C}^{(0)}_{1,2}(\boldsymbol{\theta}) & -\mathcal{C}^{(0)}_{2,2}(\boldsymbol{\theta}) & -\mathcal{E}_{2,2} & \mathcal{E}_{2,1}\\ -\frac{\partial^3 H^{(0)}_*}{\partial I_1\partial x\partial y}(\boldsymbol{\theta},\boldsymbol{0}) & -\frac{\partial^3 H^{(0)}_*}{\partial I_2\partial x\partial y}(\boldsymbol{\theta},\boldsymbol{0}) & -\frac{\partial^3 H^{(0)}_*}{\partial x\partial y^2}(\boldsymbol{\theta},\boldsymbol{0}) & \frac{\partial^3 H^{(0)}_*}{\partial x^2\partial y}(\boldsymbol{\theta},\boldsymbol{0})\end{pmatrix} \tag{3.3.14}$$

and the components of the independent term $\mathfrak{h}$ are,

$$
\begin{aligned}
\mathfrak{h}_1 &= -\langle b_1\rangle_{\boldsymbol{\theta}},\\
\mathfrak{h}_2 &= -\langle b_2\rangle_{\boldsymbol{\theta}},\\
\mathfrak{h}_3 &= -\langle c_2\rangle_{\boldsymbol{\theta}} + \Omega_2 + \left\langle \mathcal{C}^{(0)}_{2,1}(\boldsymbol{\theta})\frac{\partial d_*}{\partial\theta_1} + \mathcal{C}^{(0)}_{2,2}(\boldsymbol{\theta})\frac{\partial d_*}{\partial\theta_2}\right\rangle_{\boldsymbol{\theta}},\\
\mathfrak{h}_4 &= -\langle B_{1,2}\rangle_{\boldsymbol{\theta}} + \mathcal{B}_{1,2} + \left\langle \frac{\partial H^{(0)}_*}{\partial I_1\partial x\partial y}(\boldsymbol{\theta},\mathbf{0})\left(\frac{\partial d_*}{\partial\theta_1}\right)\right\rangle_{\boldsymbol{\theta}} + \left\langle \frac{\partial H^{(0)}_*}{\partial I_2\partial x\partial y}(\boldsymbol{\theta},\mathbf{0})\left(\frac{\partial d_*}{\partial\theta_2}\right)\right\rangle_{\boldsymbol{\theta}}\\
&\quad + \left\langle\left\{\frac{\partial^3 H^{(0)}_*}{\partial x\partial y^2}(\boldsymbol{\theta},\mathbf{0})\right\}_{\boldsymbol{\theta}}\{e_1\}_{\boldsymbol{\theta}}\right\rangle_{\boldsymbol{\theta}} - \left\langle\left\{\frac{\partial^3 H^{(0)}_*}{\partial x^2\partial y}(\boldsymbol{\theta},\mathbf{0})\right\}_{\boldsymbol{\theta}}\{e_2\}_{\boldsymbol{\theta}}\right\rangle_{\boldsymbol{\theta}}.
\end{aligned}
$$

Remark 3.16. It is important to stress now that, to solve (3.3.13), the function d and both, $\{e_1\}_{\boldsymbol{\theta}}$, $\{e_2\}_{\boldsymbol{\theta}}$ are previously required (because they play in the r. h. s., $\mathfrak{h}$), but all them may be computed from the two first $(\mathrm{eq})_1$, $(\mathrm{eq})_2$ of the homological equations set. Once (3.3.13) is solved, and hence $\boldsymbol{\xi}$, $\langle e_1\rangle_{\boldsymbol{\theta}}$, $\langle e_1\rangle_{\boldsymbol{\theta}}$, exacted, all the terms involved in (eq_3), (eq_4) and (eq_5) become determined, and these equations may be solved in $\boldsymbol{f}$, G and F (see below). ♠

Remark 3.17. The matrix $\mathcal{A}^{(0)}$ (and the averaged equations below) has been written as a function of the angles. Actually, for this initial step, $\mathcal{C}^{(0)}$ does not depend explicitly on $\boldsymbol{\theta}$ and, on the other hand, $H^{(0)}_*$ in (3.2.36) comes from the the terms $qp^2 + \frac{I_2^2}{4q} + \mathcal{Z}(q,\boldsymbol{I})$ in (3.2.4), where $\boldsymbol{\theta}$ neither appears. However, as it has already been noted in remark 3.14 (but concerning $\mathcal{C}^{(0)}$), the matrices replacing $\mathcal{A}^{(0)}$ along the iterative procedure, will depend on $\boldsymbol{\theta}$. ♠

Hence, it will be necessary to see that the matrix $\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}$ of the averaged system (3.3.13) is not singular and (in order to bound the solutions of the homological equations later on) to derive suitable estimates for the norm of its inverse. In fact, we can expand $\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}} = \left\langle\mathcal{A}^{(0)}_0\right\rangle_{\boldsymbol{\theta}} + \left\langle\mathcal{A}^{(0)}_*\right\rangle_{\boldsymbol{\theta}}$, where $\left\langle\mathcal{A}^{(0)}_0\right\rangle_{\boldsymbol{\theta}}$ is the constant matrix,

$$
\left\langle\mathcal{A}^{(0)}_0\right\rangle_{\boldsymbol{\theta}} = \begin{pmatrix}
-d & \frac{\eta}{\xi} - e & -\lambda_+ & 0\\
\frac{\xi}{\lambda_+}d & -\frac{\eta}{\lambda} + \frac{\xi}{\lambda_+}e & 0 & \lambda_+\\
-f & -\frac{1}{2\xi} - c & -\frac{\eta}{\lambda_+} + \frac{\xi}{\lambda_+}e & -\frac{\eta}{\xi} + e\\
\frac{2\xi}{\lambda_+}f_{2,1,0} & \frac{2\xi}{\lambda_+}f_{2,0,1} + \frac{2\eta}{\xi\lambda_+} & \frac{1}{2} + \frac{6\eta^2}{\lambda_+^2} - 6f_{3,0,0}\left(\frac{\xi}{\lambda_+}\right)^2 & \frac{6\eta^2}{\lambda_+\xi} + \frac{\lambda_+}{2\xi} - \frac{6\xi}{\lambda_+}f_{3,0,0}
\end{pmatrix},
$$

and,

$$
\langle\mathcal{A}^{(0)}_*\rangle_{\boldsymbol{\theta}} = \begin{pmatrix}
\mathrm{O}(R^2) & \mathrm{O}(R^2) & 0 & 0\\
\frac{\xi}{\lambda_+}\mathrm{O}(R^2) & \frac{\xi}{\lambda_+}\mathrm{O}(R^2) & 0 & 0\\
\mathrm{O}(R^2) & \mathrm{O}(R^2) & \frac{\xi}{\lambda_+}\mathrm{O}(R^2) & \mathrm{O}(R^2)\\
\frac{\xi}{\lambda_+}\mathrm{O}(R^2) & \frac{\xi}{\lambda_+}\mathrm{O}(R^2) & \left(\frac{\xi}{\lambda_+}\right)^2\mathrm{O}(R^2) & \frac{\xi}{\lambda_+}\mathrm{O}(R^2)
\end{pmatrix},
$$

here, ξ, η must be considered functions of $\boldsymbol{\Lambda}$, and note also that the entries in $\langle\mathcal{A}^{(0)}_*\rangle_\theta$ are all of order[4] $\mathrm{O}(R)$. First, using the expansions above together with the estimates (3.2.54)–(3.2.56), and (3.2.39) we can state the bound,

$$\|\langle\mathcal{A}^{(0)}\rangle_\theta\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le \widehat{m}_3^{(0)}(M^{(0)})^{-\alpha}, \tag{3.3.15}$$

and next, computation of the determinant gives,

$$\det\langle\mathcal{A}^{(0)}\rangle_\theta = -\frac{1}{\lambda_+\xi}\left(-ad\xi + \xi\mathrm{O}(R)\right),$$

but it was assumed $d \neq 0$ and a, ξ positive, so the above expression leads to,

$$|\det\langle\mathcal{A}^{(0)}\rangle_\theta| \ge \frac{a|d|}{2|\lambda_+|},$$

for R sufficiently small. Therefore[5]: $\| \left\langle\mathcal{A}^{(0)}\right\rangle_{\boldsymbol{\theta}}^{-1} \|_{\mathcal{E}^{(0)}} \le \frac{4!(\widehat{m}_3^{(0)})^3 M^{-3\alpha}|\lambda_+|}{a|d|/2}$, so we can put:

$$\| \left\langle\mathcal{A}^{(0)}\right\rangle_{\boldsymbol{\theta}}^{-1} \|_{\mathcal{E}^{(0)}} \le \overline{m}_3^{(0)}(M^{(0)})^{-3\alpha}. \tag{3.3.16}$$

To solve the homological equations, we expand them in Fourier series and equate the corresponding coefficients so *formal* solutions are obtained. Next, we ought to derive bounds of these solutions, but, as we shall want to use them iteratively, it is worth clarifying which bounds change from one step to another, and which ones are independent of the step. Following Jorba and Villanueva (1997a), we take fixed positive constants, $\widehat{m}_1$, $\widehat{m}_3$, $\overline{m}_3$, $\widehat{\nu}$, as twice the corresponding initial ones $\widehat{m}_1^{(0)}$, $\widehat{m}_3^{(0)}$, $\overline{m}_3^{(0)}$, $\widehat{\nu}^{(0)}$. Furthermore, in the description of the process, $\widehat{N}$ will denote a quantity depending on $\widehat{\alpha}_1$, $\widehat{\alpha}_2$, $\widehat{m}_1$, $\widehat{m}_2$, $\widehat{m}_3$, $\overline{m}_3$, $\widehat{\nu}$, ρ_0 and τ (see the Diophantine conditions below). To simplify the notation, $\widehat{N}$ will be redefined along the description of the iterative step to fulfill a finite number of conditions, but, once determined, it does not vary in the next step; moreover, in the derivation of the bounds, those terms changing at every iteration are marked with the superscript "(0)". Hence, by comparison of (3.2.36) and (3.3.6) and using lemma A.2,

$$\|a_* - \phi^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le M^{(0)}, \qquad \|E - \mathcal{E}\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le \frac{2M^{(0)}}{(R^{(0)})^3}, \tag{3.3.17}$$

$$\|\boldsymbol{c} - \boldsymbol{\Omega}^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le \frac{M^{(0)}}{(R^{(0)})^2}, \qquad \|B - \mathcal{B}\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le \frac{3M^{(0)}}{(R^{(0)})^2}, \tag{3.3.18}$$

$$\|\boldsymbol{b}\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le \frac{M^{(0)}}{R^{(0)}}, \qquad \|C - \mathcal{C}^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)}} \le \frac{5M^{(0)}}{(R^{(0)})^4}. \tag{3.3.19}$$

Moreover, to prove the convergence of the expansion of S, we need to control the different small divisors involved in the process. With this purpose, we shall take, basic *complex*

[4] The symbol $\mathrm{O}(R^l)$ is here introduced in the usual "big O" sense, i. e., given a function f, defined on $\widehat{\mathcal{D}}(\rho, R)$, we say that f is $\mathrm{O}(R^l)$ if, for R sufficiently small, there exists a constant κ, independent of R, such that: $|f|_{\rho,R} \le \kappa R^l$.

[5] Given a $n \times n$ nonsingular matrix, A, let $\widetilde{A}$ to denote its adjoint matrix (i. e., the matrix whose i-th row j-th column element, $\widetilde{A}_{i,j}$, $i, j = 1, \dots, n$, is the adjoint of the corresponding entry of the matrix A. Then, $|A_{i,j}| \le (n-1)!\|A\|_\infty^{n-1}$, and the inverse, $A^{-1} = \frac{1}{\det A}\widetilde{A}^*$, so $|A^{-1}|_\infty \le \frac{n!\|A\|_\infty^{n-1}}{|\det A|}$.

frequencies $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$ in a set $\breve{\mathcal{E}}^{(1)} \subset \mathbb{C}^2$, which will be fully specified later, in section 3.5.1, but for now it is enough to characterize $\breve{\mathcal{E}}^{(1)}$ as the set of those frequencies in $\breve{\mathcal{E}}^{(0)}$ which, in addition, satisfy:

$$|\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\Omega}^{(0)}\rangle + \ell\lambda_+| \geq \frac{\gamma^{(0)}(R)}{|\boldsymbol{k}|_1^\tau}, \tag{3.3.20}$$

for all $\boldsymbol{k} \in \mathbb{Z}^2 \setminus \boldsymbol{0}$, $0 < |\boldsymbol{k}|_1 < 2N^{(0)}$; $\ell \in \mathbb{Z}$, $0 \leq |\ell| \leq 2$; with $\tau > 1$ and certain $\gamma^{(0)}(R) > 0$. We stress that only the first component of the vector $\boldsymbol{\Omega}$ changes at each step. One expects the measure of $\breve{\mathcal{E}}^{(0)} \setminus \breve{\mathcal{E}}^{(1)}$ to be of the same order of $\widehat{H}^{(0)}$, so we take $\gamma^{(0)} \equiv (M^{(0)})^\alpha$.

Remark 3.18. We shall accept that, if $\breve{\mathcal{E}}^{(j)}$ is the set of basic frequencies taken at the j-th step of the iterative scheme, then $\breve{\mathcal{E}}^{(j)} \subseteq \breve{\mathcal{E}}^{(j-1)}$ for $j = 1, 2, \ldots$ This will be justified in section 3.5.1 (see remark 3.25 there). ♠

Remark 3.19. Before going on, it is important to note that, though we take, $\widehat{H}^{(0)}_{<N^{(0)},\boldsymbol{\theta}}$ and not the whole $\widehat{H}$ (so only $N^{(0)}$ harmonics appear in a_*, $\boldsymbol{c}$, $\boldsymbol{b}$, C, B, E), the solutions d_*, $\boldsymbol{e}$ of (eq$_1$), (eq$_2$) are "recombined" in the last three equations with terms that –beyond the first step of the iterative process–, contain harmonics whose order can be proved to be, for R sufficiently small, lesser than $N^{(0)}$. Then, looking at the homological equations it is easy to realize that the order of the harmonics involved (and hence in the denominators of their solutions), can be as twice $N^{(0)}$. That is why, in (3.3.20), the $|\cdot|_1$ norm of the integer vector $\boldsymbol{k}$ is bounded by $2N^{(0)}$ and not simply by $N^{(0)}$. In any case, at each step, we deal only with a finite number of resonances, so the interior of the sets of basic frequencies $\breve{\mathcal{E}}^{(j)}$, defined for the first, second and so on steps –and satisfying (3.3.20) with the superscripts (1) and (0) changed to (j) and $(j-1)$ respectively for $j = 2, 3, \ldots$–, is not empty (as would happen if all the harmonics were considered). This method is known as the "ultraviolet cut off" (see Arnol'd, 1963a, for the first example). ♠

To bound the solutions of the different homological equations we use the lemma A.1, with $\delta^{(0)} = (M^{(0)})^\alpha$ as the value of δ in the different estimates. In order to simplify, we assume $\rho^{(0)} - K\delta^{(0)} \geq \rho_0/4$, where $K \in \mathbb{N}$ is a fixed integer to be determined at the end of the description of the present step. Furthermore, it will be also assumed that $(M^{(0)})^\alpha \leq R^{(0)} \leq 1$[6]. Then, one can solve (eq$_1$) – (eq$_5$) in the form:

(eq$_1$) For d_*, we have

$$d_*(\boldsymbol{\theta}) = \sum_{0<|\boldsymbol{k}|_1<N^{(0)}} \frac{\{a_*\}_{\boldsymbol{\theta}}}{\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\Omega}^{(0)}\rangle} \exp\left(\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\theta}\rangle\right),$$

then, by lemma A.1 with the Diophantine condition (3.3.20) in mind, one derives the bound,

$$\|d_*\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-\delta^{(0)}} \leq \left(\frac{\tau}{\delta^{(0)}\exp(1)}\right)^\tau \frac{\|\{a_*\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}}}{\gamma^{(0)}} \leq \widehat{N}(M^{(0)})^{1-\alpha-\alpha\tau}.$$

(eq$_2$) Let us introduce:

$$\widetilde{\boldsymbol{b}} = \left\{-\mathcal{E}^* \left(\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right)^* + \boldsymbol{b}\right\}_{\boldsymbol{\theta}}, \tag{3.3.21}$$

[6] More precisely, we take $R^{(0)} = 2(M^{(0)})^\alpha$ –see (3.2.25)–, and then at every successive step, n, one must check that: $(M^{(n)})^\alpha \leq R^{(n)} \leq 1$ (see section 3.4).

then,

$$\{e_j\}_{\boldsymbol{\theta}} = \sum_{0\le|\boldsymbol{k}|_1\le N^{(0)}} \frac{\widetilde{b}_{j,\boldsymbol{k}}}{\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\Omega}^{(0)}\rangle + (-1)^{j+1}\lambda_+} \exp\left(\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\theta}\rangle\right), \qquad \text{for } j=1,2,$$

hence in the same way as before,

$$\|\{\boldsymbol{e}\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \left(\frac{1}{\widehat{\alpha}_1}(M^{(0)})^{-\alpha} + \left(\frac{\tau}{\delta^{(0)}\exp(1)}\right)^{\tau}\frac{1}{\gamma^{(0)}}\right)\|\widetilde{\boldsymbol{b}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-2\delta^{(0)}}.$$

To bound $\widetilde{\boldsymbol{b}}$, we take norms in (3.3.21), use the first inequality in (3.3.19) and the third item of lemma A.1,

$$\begin{aligned}
\|\widetilde{\boldsymbol{b}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-2\delta^{(0)}} &\le \frac{M^{(0)}}{R^{(0)}} + 2\widehat{m}_2(M^{(0)})^{-\alpha}\frac{\|d_*\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-\delta^{(0)}}}{\delta^{(0)}\exp(1)} \\
&\le (M^{(0)})^{1-\alpha} + \frac{2\widehat{m}_2}{\exp(1)}\widehat{N}(M^{(0)})^{1-3\alpha-\alpha\tau} \\
&\le \left(1+\frac{2\widehat{m}_2\widehat{N}}{\exp(1)}\right)(M^{(0)})^{1-3\alpha-\alpha\tau}.
\end{aligned}$$

Therefore,

$$\|\{\boldsymbol{e}\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \left(\frac{1}{\widehat{\alpha}_1} + \left(\frac{\tau}{\exp(1)}\right)^{\tau}\right)\left(1+\frac{\widehat{m}_2\widehat{N}}{\exp(1)}\right)(M^{(0)})^{1-4\alpha-2\alpha\tau}$$

so we can write,

$$\|\{\boldsymbol{e}\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \Pi_1(M^{(0)})^{1-4\alpha-2\alpha\tau},$$

with the obvious definition of Π_1. In fact, we want to bound $\boldsymbol{e}$. As $\|\boldsymbol{e}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \|\{\boldsymbol{e}\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} + \|\langle\boldsymbol{e}\rangle_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}}$, it remains to bound $\|\langle\boldsymbol{e}\rangle_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}}$. Let us denote by $\boldsymbol{\zeta}^* = (\boldsymbol{\xi}, \langle\boldsymbol{e}\rangle_{\boldsymbol{\theta}})$ the solution of the averaged system (3.3.13). Bounding explicitly the components of the vector $\mathfrak{h}$,

$$\begin{aligned}
\|\mathfrak{h}_1\|_{\breve{\mathcal{E}}^{(1)}}, \|\mathfrak{h}_2\|_{\breve{\mathcal{E}}^{(1)}} &\le \frac{M^{(0)}}{R^{(0)}}, \\
\|\mathfrak{h}_3\|_{\breve{\mathcal{E}}^{(1)}} &\le \frac{M^{(0)}}{(R^{(0)})^2} + \left\|\mathcal{C}^{(0)}\left(\frac{\partial d_*}{\partial\boldsymbol{\theta}}\right)^*\right\|_{\breve{\mathcal{E}}^{(1)},0} \\
&\le (M^{(0)})^{1-2\alpha} + \widehat{m}_1(M^{(0)})^{-\alpha}\frac{\|d_*\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-\delta^{(0)}}}{(\rho^{(0)}-\delta^{(0)})\exp(1)} \\
&\le (M^{(0)})^{1-2\alpha} + \frac{\widehat{m}_1\widehat{N}}{\rho_0\exp(1)/4}(M^{(0)})^{1-2\alpha-\alpha\tau} \\
&\le \Pi_2(M^{(0)})^{1-2\alpha-\alpha\tau},
\end{aligned}$$

$$
\begin{aligned}
\|\mathfrak{h}_4\|_{\breve{\mathcal{E}}^{(1)}} &\le \frac{3M^{(0)}}{(R^{(0)})^2} + \frac{2\widehat{\nu}}{(R^{(0)})^4}\left\|\frac{\partial d_*}{\partial \boldsymbol{\theta}}\right\|_{\breve{\mathcal{E}}^{(1)},0} + \frac{2\widehat{\nu}}{(R^{(0)})^3}\|\{\boldsymbol{e}\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},0} \\
&\le \frac{3M^{(0)}}{(R^{(0)})^2} + \frac{2\widehat{\nu}\widehat{N}}{(R^{(0)})^4\exp(1)/4}(M^{(0)})^{1-2\alpha-\alpha\tau} + \frac{2\widehat{\nu}}{(R^{(0)})^3}\Pi_1(M^{(0)})^{1-4\alpha-2\alpha\tau} \\
&\le \Pi_3(M^{(0)})^{1-7\alpha-2\alpha\tau},
\end{aligned}
$$

we arrive to the estimate:

$$
\|\mathfrak{h}\|_{\breve{\mathcal{E}}^{(1)}} \le \Pi_4(M^{(0)})^{1-7\alpha-2\alpha\tau},
$$

where $\Pi_1,\dots,\Pi_4$ are quantities depending on $\widehat{\alpha}_1$, τ, $\widehat{m}_1$, $\widehat{m}_2$, ρ_0, $\widehat{\nu}$. Thus, by (3.3.16)

$$
\begin{aligned}
\|\boldsymbol{\zeta}\|_{\breve{\mathcal{E}}^{(1)}} &= \|\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}^{-1}\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}\boldsymbol{\zeta}\|_{\breve{\mathcal{E}}^{(1)}} \\
&\le \|\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}^{-1}\|_{\breve{\mathcal{E}}^{(1)}}\|\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}\boldsymbol{\zeta}\|_{\breve{\mathcal{E}}^{(1)}} \le \overline{m}_3\Pi_4(M^{(0)})^{1-10\alpha-2\alpha\tau},
\end{aligned}
$$

which, in particular, implies:

$$
\|\boldsymbol{\xi}\|_{\breve{\mathcal{E}}^{(1)}} \le \Pi_5(M^{(0)})^{1-10\alpha-2\alpha\tau}, \qquad \|\langle\boldsymbol{e}\rangle_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)}} \le \Pi_5(M^{(0)})^{1-10\alpha-2\alpha\tau}, \tag{3.3.22}
$$

(Π_5 depending on the same constants as $\Pi_1,\dots,\Pi_4$ plus $\overline{m}_3$). So, with $\widehat{N}$ conveniently redefined,

$$
\|\boldsymbol{e}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \widehat{N}(M^{(0)})^{1-10\alpha-2\alpha\tau}. \tag{3.3.23}
$$

(eq$_3$) We introduce $\widetilde{\boldsymbol{c}}$, and $\underline{\boldsymbol{c}}$ by:

$$
\begin{aligned}
\widetilde{\boldsymbol{c}} &= \begin{pmatrix}\{c_1\}_{\boldsymbol{\theta}} \\ c_2 - \Omega_2\end{pmatrix}, \\
\underline{\boldsymbol{c}} &= \{\widetilde{\boldsymbol{c}}\}_{\boldsymbol{\theta}} - \mathcal{C}^{(0)}\left(\boldsymbol{\xi} + \left(\frac{\partial d_*}{\partial\boldsymbol{\theta}}\right)^*\right) + \left\langle \mathcal{C}^{(0)}\left(\boldsymbol{\xi} + \left(\frac{\partial d_*}{\partial\boldsymbol{\theta}}\right)^*\right)\right\rangle_{\boldsymbol{\theta}} + \mathcal{E}J\boldsymbol{e} - \langle\mathcal{E}J\boldsymbol{e}\rangle_{\boldsymbol{\theta}},
\end{aligned}
$$

with this notation the solution of the third equation is:

$$
f_j(\boldsymbol{\theta}) = \sum_{0<|\boldsymbol{k}|_1<2N^{(0)}} \frac{\underline{c}_{j,\boldsymbol{k}}}{\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\Omega}^{(0)}\rangle}\exp\left(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\theta}\rangle\right), \qquad \text{for } j=1,2,
$$

To bound $\boldsymbol{f}$, first we have to bound $\underline{\boldsymbol{c}}$:

$$
\begin{aligned}
\|\underline{\boldsymbol{c}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} &\le \|\widetilde{\boldsymbol{c}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \\
&\quad + \|\mathcal{C}^{(0)}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}}\left(\|\boldsymbol{\xi}\|_{\breve{\mathcal{E}}^{(1)}} + \frac{\|d_*\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-2\delta^{(0)}}}{\delta^{(0)}\exp(1)}\right) + \|\mathcal{E}\|_{\breve{\mathcal{E}}^{(1)}}\|\{\boldsymbol{e}\}_{\boldsymbol{\theta}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \\
&\le (M^{(0)})^{1-2\alpha} + \widehat{m}_1\Pi_5(M^{(0)})^{1-11\alpha-2\alpha\tau} + \frac{\widehat{m}_1\widehat{N}}{\exp(1)}(M^{(0)})^{1-3\alpha-\alpha\tau} \\
&\quad + 2\widehat{m}_2\Pi_1(M^{(0)})^{1-5\alpha-2\alpha\tau}
\end{aligned}
$$

where the r. h. s can be gathered to give an estimate of the form,

$$
\|\underline{\boldsymbol{c}}\|_{\breve{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \Pi_6(M^{(0)})^{1-11\alpha-2\alpha\tau},
$$

(with the appropriate definition of Π_6 in terms of the others constants). From here, and again by lemma A.1,

$$\|\boldsymbol{f}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{(0)}} \leq \widehat{N}(M^{(0)})^{1-12\alpha-3\alpha\tau}.$$

(eq$_4$) Let us introduce now $\underline{B}$ as,

$$\underline{B} = \widetilde{B} - \mathcal{B},$$

hence, if $G = (G_{j,l})_{1\leq j,l\leq 2}$:

$$G_{j,l}(\boldsymbol{\theta}) = \sum_{0\leq|\boldsymbol{k}|_1\leq 2N^{(0)}} \frac{\underline{B}_{j,l,\boldsymbol{k}}}{\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\Omega}^{(0)}\rangle + 2(-1)^{j+1}\lambda_+\delta_{j,l}} \exp\left(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\theta}\rangle\right),$$

from this last sum we have to avoid the indices $(j,l,\boldsymbol{k}) = (1,2,\boldsymbol{0})$ and $(j,l,\boldsymbol{k}) = (2,1,\boldsymbol{0})$, because they lead to zero divisors, but the coefficients $\underline{B}_{1,2,\boldsymbol{0}}$, $\underline{B}_{2,1,\boldsymbol{0}}$ are also zero (because $\boldsymbol{\xi}$, $\langle\boldsymbol{e}\rangle_{\boldsymbol{\theta}}$ were adjusted precisely to fulfill this condition). Moreover, we note that G is a symmetric matrix. Hence, to bound G we need to bound $\underline{B}$ first. Directly form the definitions we have,

$$\begin{aligned}\|\widetilde{B} - \mathcal{B}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{(0)}} \leq{}& \|B - \mathcal{B}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{(0)}} + 3\|\mathcal{E}\|_{\check{\mathcal{E}}^{(1)}} \frac{\|\boldsymbol{e}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}}}{\delta^{(0)}\exp(1)} \\ &+ \frac{6\|H_*^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^4}\left(\|\boldsymbol{\xi}\|_{\check{\mathcal{E}}^{(1)}} + \frac{\|d\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}}}{\delta^{(0)}\exp(1)}\right) \\ &+ \frac{24\|H_*^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^3}\|\boldsymbol{e}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}}\end{aligned}$$

and then,

$$\begin{aligned}\|\widetilde{B} - \mathcal{B}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{(0)}} \leq{}& \frac{3M^{(0)}}{(R^{(0)})^2} + \frac{3\widehat{m}_2}{\exp(1)}\widehat{N}(M^{(0)})^{1-12\alpha-2\alpha\tau} \\ &+ 6\widehat{\nu}(M^{(0)})^{-4\alpha}\left(\Pi_5(M^{(0)})^{1-10\alpha-2\alpha\tau} + \frac{\widehat{N}}{\exp(1)}(M^{(0)})^{1-2\alpha-\alpha\tau}\right) \\ &+ 24\widehat{\nu}\widehat{N}(M^{(0)})^{1-13\alpha-2\alpha\tau},\end{aligned}$$

which, with the introduction of a new constant Π_7, allows us to write,

$$\|\widetilde{B} - \mathcal{B}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{0}} \leq \Pi_7(M^{(0)})^{1-14\alpha-2\alpha\tau},$$

and therefore (with $\widehat{N}$ conveniently redefined),

$$\begin{aligned}\|G\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-5\delta^{(0)}} &\leq 2\left(\frac{1}{\widehat{\alpha}_1}(M^{(0)})^{-\alpha} + \left(\frac{\tau}{\delta^{(0)}\exp(1)}\right)^{\tau}\frac{1}{\gamma^{(0)}}\right)\|\underline{B}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{(0)}} \\ &\leq \widehat{N}(M^{(0)})^{1-15\alpha-3\alpha\tau}.\end{aligned}$$

(eq$_5$) As before, we define,

$$\underline{E} = \widetilde{E} - \mathcal{E},$$

and then, the components of F are given by,

$$F_{j,l}(\boldsymbol{\theta}) = \sum_{0\le|\boldsymbol{k}|_1\le 2N^{(0)}} \frac{\underline{E}_{j,l,\boldsymbol{k}}}{\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\Omega}^{(0)}\rangle + (-1)^{l+1}\lambda_+} \exp\left(\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{\theta}\rangle\right), \qquad \text{for } j,l = 1,2.$$

As for the previous equation, we need to bound this numerator. From its definition, using (3.3.17), we get

$$\|\widetilde{E} - \mathcal{E}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-5\delta^{(0)}} \le \|E - \mathcal{E}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}}$$

$$+\|\mathcal{C}^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}} \frac{\|\boldsymbol{e}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}}}{\delta^{(0)}\exp(1)} + \|\mathcal{E}\|_{\check{\mathcal{E}}^{(1)}} \frac{\|\boldsymbol{f}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-4\delta^{(0)}}}{\delta^{(0)}\exp(1)} + \|\mathcal{E}\|_{\check{\mathcal{E}}^{(1)}}\|G\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-5\delta^{(0)}}$$

$$+\frac{8\|H_*^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^5}\left(\|\boldsymbol{\xi}\|_{\check{\mathcal{E}}^{(1)}} + \frac{\|d\|_{\check{\mathcal{E}}^{(1)}\rho^{(0)}-\delta^{(0)}}}{\delta^{(0)}\exp(1)}\right) + \frac{8\|H_*^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^4}\|\boldsymbol{e}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-3\delta^{(0)}},$$

and, by substitution of the bounds previously obtained,

$$\|\widetilde{E} - \mathcal{E}\|_{\check{\mathcal{E}}^{(1)}} \le \frac{2M^{(0)}}{(R^{(0)})^3} + \frac{\widehat{m}_1\widehat{N}}{\exp(1)}(M^{(0)})^{1-12\alpha-2\alpha\tau} + \frac{\widehat{m}_2\widehat{N}}{\exp(1)}(M^{(0)})^{1-14\alpha-3\alpha\tau}$$

$$+\widehat{m}_2\widehat{N}(M^{(0)})^{1-16\alpha-3\alpha\tau} + 8\widehat{\nu}(M^{(0)})^{-5\alpha}\left(\Pi_5(M^{(0)})^{1-10\alpha-2\alpha\tau} + \frac{\widehat{N}}{\exp(1)}(M^{(0)})^{1-2\alpha-\alpha\tau}\right)$$

$$+8\widehat{\nu}\widehat{N}(M^{(0)})^{1-14\alpha-2\alpha\tau},$$

but, if we introduce a quantity, Π_8 (depending, as all the previous ones $\Pi_1,\dots,\Pi_7$ on the forementioned constants $\widehat{\alpha}_1, \widehat{m}_1, \widehat{m}_2, \dots,$ which do not change with the step), the above inequality can be shorten as,

$$\|\underline{E}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-5\delta^{(0)}} \le \Pi_8(M^{(0)})^{1-16\alpha-3\alpha\tau}$$

and from here,

$$\|F\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-6\delta^{(0)}} \le \widehat{N}(M^{(0)})^{1-17\alpha-4\alpha\tau}.$$

Now, we use these estimates to bound the transformed Hamiltonian $H^{(1)}$, which, using definitions in (3.3.2) casts into:

$$H^{(1)} = \widetilde{H}^{(0)} + \widehat{H}^{(0)}_{<N^{(0)},\boldsymbol{\theta}} + \check{H}^{(0)}_1 + \widehat{H}^{(0)}_{\ge N^{(0)},\boldsymbol{\theta}} + \widehat{H}^{(1)}_{\int},$$

this is in fact the same expression as in (3.3.3), but in the term $\widehat{H}^{(1)}$ –see (3.3.4)–, we have separated the contribution due to the harmonics of order $\ge N^{(0)}$ and those of the integral term, so we put,

$$\widehat{H}^{(1)}_{\int} = \int_0^1 \left(\check{H}^{(0)}_2 + (1-t)\{\check{H}^{(0)}_1, S\}\right)\circ\psi^S_t\, dt, \tag{3.3.24}$$

and we shall use lemma A.3 to bound the different Poisson brackets involved in the expansions above. First, from the estimates on the solutions of (eq$_1$)-(eq$_5$), we derive,

$$\|\text{grad}\, S\|_{\check{\mathcal{E}}^{(0)},\rho^{(0)}-7\delta^{(0)},R^{(0)}} \le \widehat{N}(M^{(0)})^{1-18\alpha-4\alpha\tau}, \tag{3.3.25}$$

and after explicit computation one obtains,

$$\|\check{H}_1^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} \le \widehat{N}(M^{(0)})^{1-21\alpha-4\alpha\tau},$$

$$\|\{\check{H}_1^{(0)}, S\}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-8\delta^{(0)},R^{(0)}\exp(-2\delta^{(0)})} \le \widehat{N}(M^{(0)})^{2-42\alpha-8\alpha\tau}, \tag{3.3.26}$$

$$\|\check{H}_2^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} \le \widehat{N}(M^{(0)})^{2-20\alpha-4\alpha\tau}, \tag{3.3.27}$$

but, to bound $\widehat{H}_{\int}^{(1)}$ one needs to control the effect of the transformation ψ_t^S. so, if we assume the condition,

$$\|\text{grad}\, S\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)},R^{(0)}} \le (R^{(0)})^2\delta^{(0)}\exp(-1)/2, \tag{3.3.28}$$

then, by lemma A.8 and (A.2.6), ψ_t^S is well-defined from $\mathcal{D}_*(\rho^{(0)}-8\delta^{(0)}, R^{(0)}\exp(-\delta^{(0)}))$ to $\mathcal{D}_*(\rho^{(0)}-7\delta^{(0)}, R^{(0)})$, for any $-1<t<1$ and $\boldsymbol{\Lambda}\in\check{\mathcal{E}}^{(1)}$. More precisely, and from the second item of the lemma,

$$\|\psi_t^S - \text{Id}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-8\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} \le \|\text{grad}\, S\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)},R^{(0)}}, \tag{3.3.29}$$

for all $-1\le t\le 1$, and $\boldsymbol{\Lambda}\in\check{\mathcal{E}}^{(1)}$. We stress that, actually (3.3.28) holds whenever $\widehat{N}(M^{(0)})^{1-21\alpha-4\alpha\tau}\le\frac{1}{2\exp(1)}$, but this condition could be assured from the inductive restrictions. Hence, taking (3.3.26), (3.3.27) and (3.3.29) into account, application of lemma A.4 in (3.3.24) yields:

$$\|\widehat{H}_{\int}^{(1)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-9\delta^{(0)},R^{(0)}\exp(-3\delta^{(0)})} \le \widehat{N}(M^{(0)})^{2-42\alpha-8\alpha\tau}. \tag{3.3.30}$$

Now a look at (3.3.1) allows one to realize that, if, for example, we want $\widehat{H}^{(0)}_{\ge N^{(0)},\boldsymbol{\theta}}$ to have the same size as $\widehat{H}_{\int}^{(1)}$, then, we can ask for the condition,

$$\exp(-9\delta^{(0)}N^{(0)}) \le \widehat{N}(M^{(0)})^{1-42\alpha-8\alpha\tau},$$

which implies that $N^{(0)}$ must be as large as:

$$N^{(0)} \ge \frac{\ln\widehat{N} + (1-42\alpha-8\alpha\tau)\ln M^{(0)}}{-9\delta^{(0)}}. \tag{3.3.31}$$

Note: we assume that α is small enough to make $1-42\alpha-8\alpha\tau>0$ (see the definition of s in the next section). This fixes $N^{(0)}$, which of course depends on the iterate and tends monotonically to infinity (we should remove an increasing number of harmonics at each successive step). Hence,

$$\|\widehat{H}^{(1)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-9\delta^{(0)},R^{(0)}\exp(-3\delta^{(0)})} \le \widehat{N}(M^{(0)})^{2-42\alpha-8\alpha\tau}. \tag{3.3.32}$$

Moreover, from the bounds on the norm of $\check{H}_1^{(0)}$,

$$\begin{aligned} \|\phi^{(1)}-\phi^{(0)}\|_{\check{\mathcal{E}}^{(1)}} &\le \widehat{N}(M^{(0)})^{1-21\alpha-4\alpha\tau}, \\ \|\Omega_1^{(1)}-\Omega_1^{(0)}\|_{\check{\mathcal{E}}^{(1)}} &\le \widehat{N}(M^{(0)})^{1-23\alpha-4\alpha\tau}, \\ \|\mathcal{C}^{(1)}-\mathcal{C}^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)}} &\le \widehat{N}(M^{(0)})^{1-25\alpha-4\alpha\tau}, \\ \|\mathcal{A}^{(1)}-\mathcal{A}^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)}} &\le \widehat{N}(M^{(0)})^{1-25\alpha-4\alpha\tau}, \\ \|H_*^{(1)}-H_*^{(0)}\|_{\check{\mathcal{E}}^{(1)},\rho^{(0)}-7\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} &\le \widehat{N}(M^{(0)})^{1-21\alpha-4\alpha\tau}. \end{aligned} \tag{3.3.33}$$

Therefore, we take $K \ge 9$ and define,

$$\rho^{(1)} = \rho^{(0)} - K\delta^{(0)}, \qquad R^{(1)} = R^{(0)}\exp(-(K-6)\delta^{(0)}). \tag{3.3.34}$$

It is possible to rewrite the bounds (3.3.17)–(3.3.19) but now for the transformed Hamiltonian $H^{(1)}$ on $\mathcal{D}_*(\rho^{(1)}, R^{(1)})$. However, to iterate the scheme just described, it is worth checking that the bounds assumed on $H^{(0)}$ to define $\widehat{N}$, still hold on $H^{(1)}$. This is the subject of the next section.

3.4 Convergence of the iterative scheme

Looking at the bounds of the previous section we define $s = 2(1-25\alpha-4\alpha\tau)$ and (for $\tau > 1$ fixed), take $\alpha > 0$ small enough to make $s > 1$. Assuming $\widehat{N} \ge 1$, let us introduce $M^{(1)} = (\widehat{N}M^{(0)})^s$ and note that this is a bound for the norm of $\widehat{H}^{(1)}$ given in (3.3.32). Then (provided the hypothesis needed to iterate hold) one defines recursively, $M^{(j)} = (\widehat{N}M^{(j-1)})^s$, so $M^{(j)} = (\widehat{N}^{1+\frac{1}{s}+\cdots+\frac{1}{s^{j-1}}}M^{(0)})^{s^j}$. Also, it is convenient to define $\Pi \equiv \widehat{N}^{\frac{1}{1-1/s}}$. Then, since $\widehat{N}^{1+\frac{1}{s}+\cdots+\frac{1}{s^{j-1}}} \le \Pi$ the inequalities:

$$M^{(j)} \le (\Pi M^{(0)})^{s^j}, \tag{3.4.1a}$$

$$\widehat{N}(M^{(j)})^{s/2} \le (\Pi M^{(0)})^{\frac{s^{j+1}}{2}}, \tag{3.4.1b}$$

are easily checked to hold for $j = 0, 1, 2, \ldots$ Moreover, in order to derive easier bounds, we shall suppose $\Pi M^{(0)}(R) < 1$ and that the sequence $\{M^{(j)}\}_{j\in\mathbb{Z}_+}$ decreases monotonically(7), thus $\lim_{n\to\infty} M^{(n)} = 0$. Let us define now $\check{\mathcal{E}}^{(\infty)}(R)$ as the set of frequencies $\boldsymbol{\Lambda}$ for which all the steps are well-defined, and assume that for any $\boldsymbol{\Lambda} \in \check{\mathcal{E}}^{(\infty)}(R)$ the composition of the canonical transformations determined at each step of the iterative process, $\psi^{(\infty)} = \psi^{S^{(0)}} \circ \psi^{S^{(1)}} \circ \ldots$, is convergent, then, the limit Hamiltonian, $H^{(\infty)} = H^{(0)} \circ \psi^{(\infty)}$ will take the form (3.2.30), with $\boldsymbol{\Omega}^{(\infty)*} = (\Omega_1^{(\infty)}(\boldsymbol{\Lambda}), \Omega_2)$. Hence, we obtain, for any $\boldsymbol{\Lambda} \in \check{\mathcal{E}}^{(\infty)}$ a Hamiltonian with a two-dimensional invariant torus with linear quasi-periodic flow with frequency $\boldsymbol{\Omega}^{(\infty)}$.

To prove the validity of the inductive bounds, we first check that it is possible to define, recursively, j-th step constants: $\widehat{m}_1^{(j)}$, $\widehat{m}_3^{(j)}$, $\overline{m}_3^{(j)}$, $\widehat{\nu}^{(j)}$, replacing the initial ones, $\widehat{m}_1^{(0)}$, $\widehat{m}_3^{(0)}$, $\overline{m}_3^{(0)}$, $\widehat{\nu}^{(0)}$, but still bounded by $\widehat{m}_1$, $\widehat{m}_3$, $\overline{m}_3$, $\widehat{\nu}$ respectively. To do this, we

(7)both assumptions can be achieved taking R small enough. The first one is clear and, to justify the second, we remark that the quotient $M^{(j+1)}/M^{(j)} = \widehat{N}^{s^{j+1}}(M^{(0)})^{s^{j+1}-s^j}$, then it suffices to ask $M^{(0)} < \widehat{N}^{-\frac{s}{s-1}}$.

realize that the r. h. s. terms of the inequalities in (3.3.33) can be bounded by $\widehat{N}(M^{(0)})^{s/2}$ (and the same bound works also for (3.3.25)), then, assume (3.2.55),(3.2.57) and (3.3.15), (3.3.16) hold for $j = 1, 2, \dots, n-1$, i. e., assume:

$$||\mathcal{C}^{(j)}||_{\breve{\mathcal{E}}^{(j)},\rho^{(j)}} \leq \widehat{m}_1^{(j)}(M^{(j)})^{-\alpha}, \tag{3.4.2}$$

$$||H_*^{(j)}||_{\breve{\mathcal{E}}^{(j)},\rho^{(j)},R^{(j)}} \leq \widehat{\nu}^{(j)}. \tag{3.4.3}$$

$$\|\mathcal{A}^{(j)}\|_{\breve{\mathcal{E}}^{(j)},\rho^{(j)}} \leq \widehat{m}_3^{(j)}(M^{(j)})^{-\alpha}, \tag{3.4.4}$$

$$\| \left\langle \mathcal{A}^{(j)} \right\rangle_{\boldsymbol{\theta}}^{-1} \|_{\breve{\mathcal{E}}^{(j)}} \leq \overline{m}_3^{(j)}(M^{(j)})^{-3\alpha}, \tag{3.4.5}$$

with $\widehat{m}_1^{(j)} \leq \widehat{m}_1$, $\widehat{m}_3^{(j)} \leq \widehat{m}_3$, $\overline{m}_3^{(j)} \leq \overline{m}_3$, $\widehat{\nu}^{(j)} \leq \widehat{\nu}$ for $j = 1, \dots, n-1$, being $\rho^{(j)} = \rho^{(j-1)} - K\delta^{(j-1)}$ and $\breve{\mathcal{E}}^{(j)} \subseteq \breve{\mathcal{E}}^{(j-1)}$ –the set holding the "good" frequencies at the j-th step–. These assumptions allow n-iterations of the process and, at the end of the j-th ($1 \leq j \leq n$) iteration, one derives the analogous estimates to those in (3.3.33), but with (j) replacing super-index (1) and ($j-1$) replacing super-index (0) therein. Thus, consider, at the end of the n-th step, the corresponding –in the recursive sense just described–, to the third inequality in (3.3.33). Therefore:

$$\begin{aligned}
\|\mathcal{C}^{(n)}\|_{\breve{\mathcal{E}}^{(n)},\rho^{(n)}} &\leq \|\mathcal{C}^{(n-1)}\|_{\breve{\mathcal{E}}^{(n-1)},\rho^{(n-1)}} + \|\mathcal{C}^{(n)} - \mathcal{C}^{(n-1)}\|_{\breve{\mathcal{E}}^{(n)},\rho^{(n-1)}-7\delta^{(n-1)}} \\
&\leq \|\mathcal{C}^{(n-1)}\|_{\breve{\mathcal{E}}^{(n-1)},\rho^{(n-1)}} + \widehat{N}(M^{(n-1)})^{s/2} \\
&\leq \|\mathcal{C}^{(n-2)}\|_{\breve{\mathcal{E}}^{(n-2)},\rho^{(n-2)}} + \widehat{N}(M^{(n-2)})^{s/2} + \widehat{N}(M^{(n-1)})^{s/2} \\
&\leq \widehat{m}_1^{(0)}(M^{(0)})^{-\alpha} + \sum_{j=0}^{n-1} \widehat{N}(M^{(j)})^{s/2} \\
&\leq \widehat{m}_1^{(0)}(M^{(0)})^{-\alpha} + \sum_{j\geq 0} (\Pi M^{(0)})^{\frac{s^{j+1}}{2}},
\end{aligned}$$

(where (3.4.1b) is applied in the last inequality) and the sum $\sum_{j\geq 0}(\Pi M^{(0)})^{\frac{s^{j+1}}{2}}$ is not only convergent, but also tends to zero as R does. Whence, for R small enough, it can be made less than $\widehat{m}_1^{(0)}$, so the rightmost term in the inequalities above will, in turn, be smaller than,

$$\widehat{m}_1^{(0)}(M^{(0)})^{-\alpha} + \widehat{m}_1^{(0)}(M^{(0)})^{-\alpha}(M^{(0)})^{\alpha} < 2\widehat{m}_1^{(0)}(M^{(0)})^{-\alpha} = \widehat{m}_1(M^{(0)})^{-\alpha}$$

(here, the assumption, $M^{(0)} < 1$ has been used). This closes the induction and shows that: $\widehat{m}_1^{(n)} \leq \widehat{m}_1$ for all $n \in \mathbb{Z}_+$. Similar analysis, applied to the fourth and fifth in (3.3.33) –as obtained at the end of the n-th iterate–, will lead to $\widehat{m}_3^{(j)} \leq \widehat{m}_3$, $\widehat{\nu}^{(j)} \leq \widehat{\nu}$ respectively. However, $\overline{m}_3^{(j)} \leq \overline{m}_3$ does not follow in the same way. To check it out, we shall proceed from the following result, concerning nonsingular matrices.

Lemma 3.20. *Let A be a $n \times n$ nonsingular matrix with complex components and $\|\cdot\|$ denote a norm in $\mathbb{C}^n$ and its associated matrix norm. Furthermore, if $\varrho > 0$. Then:*

$$\|A^{-1}\| \leq \varrho \Leftrightarrow \|A\boldsymbol{u}\| \geq \varrho^{-1}\|\boldsymbol{u}\|,$$

for all $\boldsymbol{u} \in \mathbb{C}^n$.

Proof. Given $\boldsymbol{u} \in \mathbb{C}^n$: $\|\boldsymbol{u}\| = \|A^{-1}A\boldsymbol{u}\| \leq \|A^{-1}\|\|A\boldsymbol{u}\| \Rightarrow \|A\boldsymbol{u}\| \geq \frac{\|\boldsymbol{u}\|}{\|A^{-1}\|} \geq \varrho^{-1}\|\boldsymbol{u}\|$. To prove the converse, take the vector $\boldsymbol{w} = A^{-1}\boldsymbol{v}$ with $\|\boldsymbol{v}\| = 1$ and such that: $\|A^{-1}\boldsymbol{v}\| = \|A^{-1}\|$. Hence $1 = \|\boldsymbol{v}\| = \|A\boldsymbol{w}\| \geq \varrho^{-1}\|A^{-1}\boldsymbol{v}\| = \varrho^{-1}\|A^{-1}\|$, which implies: $\|A^{-1}\| \leq \varrho$. □

For any $\boldsymbol{u} \in \mathbb{C}^4$ we have,

$$\begin{aligned}\|\langle\mathcal{A}^{(1)}\rangle_{\boldsymbol{\theta}}\boldsymbol{u}\|_{\breve{\mathcal{E}}^{(1)}} &\geq \|\langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}}\boldsymbol{u}\|_{\breve{\mathcal{E}}^{(1)}} - \|(\langle\mathcal{A}^{(1)}\rangle_{\boldsymbol{\theta}} - \langle\mathcal{A}^{(0)}\rangle_{\boldsymbol{\theta}})\boldsymbol{u}\|_{\breve{\mathcal{E}}^{(1)}} \geq \\ &\geq \left((\overline{m}_3^{(0)})^{-1} - \widehat{N}(M^{(0)})^{s/2}\right)(M^{(0)})^{3\alpha}\|\boldsymbol{u}\| \quad (3.4.6)\end{aligned}$$

(we stress that the r. h. s. of the fourth inequality of (3.3.33) can be bounded by $\widehat{N}(M^{(0)})^{s/2}(M^{(0)})^{3\alpha}$). Applying the lemma above,

$$\|\langle\mathcal{A}^{(1)}\rangle_{\boldsymbol{\theta}}^{-1}\|_{\breve{\mathcal{E}}^{(1)}} \leq \frac{\overline{m}_3^{(0)}}{1 - \overline{m}_3\widehat{N}(M^{(0)})^{s/2}}(M^{(0)})^{-3\alpha},$$

(where we have used that $\overline{m}_3^{(0)} \leq \overline{m}_3$). If we define

$$\overline{m}_3^{(1)} \equiv \frac{\overline{m}_3^{(0)}}{1 - \overline{m}_3\widehat{N}(M^{(0)})^{s/2}},$$

then, it is clear that, for R small enough $\overline{m}_3^{(1)} \leq \overline{m}_3$. Assume now, for $j = 1, \ldots, n-1$:

(i) $\|\langle\mathcal{A}^{(j)}\rangle_{\boldsymbol{\theta}}^{-1}\|_{\breve{\mathcal{E}}^{(j)}} \leq \overline{m}_3^{(j)}(M^{(0)})^{-3\alpha}$ $\left(\text{and hence: } \leq \overline{m}_3(M^{(0)})^{-3\alpha}\right)$.

(ii) $\overline{m}_3^{(j)} \equiv \overline{m}_3^{(0)} \prod_{l=0}^{j-1} \frac{1}{1-\overline{m}_3\widehat{N}(M^{(l)})^{s/2}}$.

(iii) Our already stated assumption: $\overline{m}_3^{(j)} \leq \overline{m}_3$.

Then, proceeding as in (3.4.6), for $\boldsymbol{u} \in \mathbb{C}^n$ we have:

$$\begin{aligned}\|\langle\mathcal{A}^{(n)}\rangle_{\boldsymbol{\theta}}\boldsymbol{u}\|_{\breve{\mathcal{E}}^{(n)}} &\geq \left(\left(\overline{m}_3^{(n-1)}\right)^{-1}(M^{(0)})^{3\alpha} - \widehat{N}(M^{(n-1)})^{s/2}(M^{(n-1)})^{3\alpha}\right)\|\boldsymbol{u}\| \\ &\geq \left(\left(\overline{m}_3^{(n-1)}\right)^{-1} - \widehat{N}(M^{(n-1)})^{s/2}\right)(M^{(0)})^{3\alpha}\|\boldsymbol{u}\| \quad (3.4.7)\end{aligned}$$

(using that $M^{(n-1)} \leq M^{(0)}$), and by lemma 3.20:

$$\begin{aligned}\|\langle\mathcal{A}^{(n)}\rangle_{\boldsymbol{\theta}}^{-1}\|_{\breve{\mathcal{E}}^{(n)}} &\leq \frac{\overline{m}_3^{(n-1)}}{1 - \overline{m}_3\widehat{N}(M^{(n-1)})^{s/2}}(M^{(0)})^{-3\alpha} \\ &\leq \overline{m}_3^{(0)}\left(\prod_{j=0}^{n-1}\frac{1}{1-\overline{m}_3\widehat{N}(M^{(j)})^{s/2}}\right)(M^{(0)})^{-3\alpha}.\end{aligned}$$

To complete the induction, we must check that: $\overline{m}_3^{(n)} \equiv \overline{m}_3^{(0)}\left(\prod_{j=0}^{n-1}\frac{1}{1-\overline{m}_3\widehat{N}(M^{(j)})^{s/2}}\right) \leq \overline{m}_3$. Since $\widehat{N}(M^{(j)})^{s/2} \leq (\Pi M^{(0)})^{\frac{s^{j+1}}{2}}$ (see (3.4.1b)), we have:

$$\overline{m}_3^{(n)} \leq \overline{m}_3^{(0)}\left(\prod_{j\geq 0}\frac{1}{1-\overline{m}_3(\Pi M^{(0)})^{\frac{s^{j+1}}{2}}}\right).$$

It turns out that the product on the right hand side converges and, actually tends to 1 when R goes to zero (as follows from the convergence of $\sum_{j\geq 0}(\Pi M^{(0)})^{\frac{s^{j+1}}{2}}$, whose sum goes to zero when R does). The former assertion is derived straightforward from the next lemma:

Lemma 3.21. *The real product:*

$$\prod_{j\geq 1}\frac{1}{1-\gamma_* x^{s^j}}, \qquad \textit{with } \gamma_*>0,\ s>1,$$

converges for $0\leq x<\inf\left\{1,(\frac{1}{2\gamma_*})^{1/s}\right\}$, *and*

$$\lim_{x\to 0^+}\prod_{j\geq 1}\frac{1}{1-\gamma_* x^{s^j}}=1.$$

Proof. Since $\prod_{j\geq 1}\frac{1}{1-\gamma_* x^{s^j}}=\exp\left\{-\sum_{j\geq 1}\ln(1-\gamma_* x^{s^j})\right\}$, the convergence of the sum inside the exponential implies the convergence of the product. But,

$$-\ln(1-\gamma_* x^{s^j})\leq 2\gamma_* x^{s^j} \Leftrightarrow \gamma_* x^{s^j}\leq 1-\exp\left(-2\gamma_* x^{s^j}\right)$$

and the last inequality holds (see (A.2.6)) because $0<2\gamma_* x^{s^j}\leq 2\gamma_* x^s\leq 1$ for all $j\geq 1$. On the other hand, the sum $\sum_{j\geq 1}x^{s^j}$ converges and tends to zero when $x\to 0$, so the product will tend to 1 when $R\to 0$. ☐

Hence, identifying $\gamma_*\equiv\overline{m}_3$ and $x\equiv(\Pi M^{(0)})^{s/2}$, it is clear that $\overline{m}_3^{(n)}\leq\overline{m}_3$ (for R sufficiently small).

Next, we need to check that $\rho^{(n)}\geq\rho_0/4$, $R^{(n)}\geq(M^{(n)})^\alpha$. From the inductive definitions, $\rho^{(n)}=\rho^{(n-1)}-K\delta^{(n-1)}$, $R^{(n)}=R^{(n-1)}\exp(-(K-6)\delta^{(n-1)})$, $n=2,3,\ldots$, is

$$\rho^{(n)}=\rho^{(0)}-K(\delta^{(0)}+\delta^{(1)}+\cdots+\delta^{(n-1)}),$$

and, as we take $\delta^{(n)}=(M^{(n)})^\alpha$, if we use $M^{(j)}\leq(\Pi M^{(0)})^{s^j}$ (see (3.4.1a)),

$$\sum_{j=0}^{n-1}\delta^{(j)}=\sum_{j=0}^{n-1}(M^{(j)})^\alpha\leq(\Pi M^{(0)})^\alpha+\sum_{j\geq 0}(\Pi M^{(0)})^{\alpha s^{j+1}}\leq 2(\Pi M^{(0)})^\alpha$$

provided R sufficiently small. Then, as K is a fixed number, the desired upper bound on $\rho^{(n)}$ follows. Identically, we have,

$$\begin{aligned}R^{(n)}&=R^{(0)}\exp\left(-(K-6)(\delta^{(0)}+\cdots+\delta^{(n-1)}\right)\\&\geq R^{(0)}\exp\left(-\frac{\rho_0}{4}\right)\geq R^{(0)}\exp\left(-\frac{1}{4}\right)\geq\frac{R^{(0)}}{2}=(M^{(0)})^\alpha\geq(M^{(n)})^\alpha,\end{aligned}\qquad(3.4.8)$$

at least for R small enough.

3.4.1 Convergence of the change of variables

Finally, to prove the well defined character of the limit Hamiltonian, it only remains to check the convergence of the sequence $\{\breve{\psi}^{(n)}\}_{n\in\mathbb{Z}_+}$, being $\breve{\psi}^{(n)} = \psi^{(0)} \circ \cdots \circ \psi^{(n)}$, where, for simplicity, we write $\psi^{(j)} = \psi_1^{S^{(j)}}$, $j \geq 0$ and also introduce $\rho'_n = \rho^{(n)} - \rho_0/8$, $R'_n = R^{(n)} \exp(-\frac{\rho_0}{8})$, for $n \geq 1$. Furthermore, we have already introduced the set $\breve{\mathcal{E}}^{(\infty)} = \bigcap_{j\geq 0} \breve{\mathcal{E}}^{(j)}$, as the set of frequencies $\boldsymbol{\Lambda}$ where all the transformations are well defined. We pick a fixed $\boldsymbol{\Lambda} \in \breve{\mathcal{E}}^{(\infty)}$, though, actually, the results will be valid for the whole set $\breve{\mathcal{E}}^{(\infty)}$. In order to prove the Cauchy character (and hence its convergence) of $\{\breve{\psi}^{(n)}\}_{n\in\mathbb{Z}_+}$, we shall look for (suitable) estimates of $\|\breve{\psi}^{(n+1)} - \breve{\psi}^{(n)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}}$. Next lemma gathers some results useful for such purpose.

Lemma 3.22. *The following bounds hold,*

(i) $\|(\psi^{(\nu)} - \mathrm{Id}) \circ \psi^{(\nu+1)} \circ \cdots \circ \psi^{(n)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+1},R'_{n+1}} \leq \|\psi^{(\nu)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{\nu+1},R'_{\nu+1}}$, *for* $\nu \geq 0$.

(ii) $\|\psi^{(l)} \circ \cdots \circ \psi^{(n)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+1},R'_{n+1}} \leq (\delta^{(l)} + \delta^{(l+1)} + \cdots + \delta^{(n)})(R^{(l)})^2$, *for* $l \geq 1$.

(iii) Let us define, $G_l^{(1)} = \psi^{(l)} \circ \cdots \circ \psi^{(n)}$, $G_l^{(2)} = \psi^{(l)} \circ \cdots \circ \psi^{(n+1)}$, *with components:* $G_l^{(j)*} = (\boldsymbol{\theta}^* + \boldsymbol{\Theta}_l^{(j)*}, \boldsymbol{\mathcal{X}}_l^{(j)*}, \boldsymbol{\mathcal{I}}_l^{(j)*}, \boldsymbol{\mathcal{Y}}_l^{(j)*})$, *for* $j = 1, 2$, $n \geq l \geq 1$ *and being* $\boldsymbol{\Theta}_l^{(j)}$, $\boldsymbol{\mathcal{I}}_l^{(j)}$, $\boldsymbol{\mathcal{X}}_l^{(j)}$, $\boldsymbol{\mathcal{Y}}_l^{(j)}$ *analytic* 2π*-periodic functions on* $\boldsymbol{\theta}$ *defined in* $\mathcal{D}(\rho'_{n+2}, R'_{n+2})$. *Then:*

$$\|\boldsymbol{\Theta}_l^{(j)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} \leq \rho^{(l)} - \rho'_{n+2} - \frac{\rho_0}{32}(R^{(0)})^2,$$
$$\|\boldsymbol{\mathcal{I}}_l^{(j)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} \leq (R^{(l)})^2 - \frac{\rho_0}{32}(R^{(0)})^2,$$
$$\|\boldsymbol{\mathcal{X}}_l^{(j)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}}, \|\boldsymbol{\mathcal{Y}}_l^{(j)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} \leq R^{(l)} - \frac{\rho_0}{32}R^{(0)}.$$

Proof. (i) It is enough to consider $\psi^{(\nu+1)} \circ \cdots \circ \psi^{(n)}$, as a function from $\mathcal{D}(\rho^{(n+1)}, R^{(n+1)})$ to $\mathcal{D}(\rho^{(\nu+1)}, R^{(\nu+1)})$ and then apply lemma A.4 identifying the functions f and F there with $\psi^{(\nu)} - \mathrm{Id}$ and $(\psi^{(\nu)} - \mathrm{Id}) \circ \psi^{(\nu+1)} \circ \cdots \circ \psi^{(n)}$, respectively. To prove (ii), use the decomposition:

$$\|\psi^{(l)} \circ \cdots \circ \psi^{(n)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+1},R'_{n+1}} \leq$$
$$\leq \sum_{\nu=l}^{n-1} \|(\psi^{(\nu)} - \mathrm{Id}) \circ \psi^{(\nu+1)} \circ \cdots \circ \psi^{(n)}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+1},R'_{n+1}} +$$
$$+\|\psi^{(n)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+1},R'_{n+1}},$$

apply the previous item to every term in the sum, i. e.,

$$\|\psi^{(l)} \circ \cdots \circ \psi^{(n)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{n+1},R'_{n+1}} \leq \sum_{\nu=l}^{n} \|\psi^{(\nu)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{\nu+1},R'_{\nu+1}},$$

and next, the recursive bounds (3.3.28) and (3.3.29) to obtain:

$$\sum_{\nu=l}^{n} \|\psi^{(\nu)} - \mathrm{Id}\|_{\breve{\mathcal{E}}^{(\infty)},\rho'_{\nu+1},R'_{\nu+1}} \leq (R^{(l)})^2(\delta^{(l)} + \ldots \delta^{(n+1)}),$$

which is the desired result. From this last item, one has:

$$\begin{aligned}
\|\boldsymbol{\Theta}_l^{(j)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} &\le (R^{(l)})^2(\delta^{(l)}+\cdots+\delta^{(n+1)}),\\
\|\boldsymbol{\mathcal{I}}_l^{(j)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} &\le (R^{(l)})^2(\delta^{(l)}+\cdots+\delta^{(n+1)})+(R'_{n+2})^2,\\
\|\boldsymbol{\mathcal{Z}}_l^{(j)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} &\le (R^{(l)})^2(\delta^{(l)}+\cdots+\delta^{(n+1)})+R'_{n+2},
\end{aligned}$$

where $\boldsymbol{\mathcal{Z}}_l^{(j)*}=(\boldsymbol{\mathcal{X}}_l^{(j)*},\boldsymbol{\mathcal{Y}}_l^{(j)*})$, $j=1,2$. Thus, to prove (iii) we only need to check that,

$$\rho^{(l)}-\rho'_{n+2}-\frac{\rho}{32}(R^{(0)})^2 \ge (R^{(l)})^2(\delta^{(l)}+\cdots+\delta^{(n+1)}), \tag{3.4.9}$$

$$(R^{(l)})^2-\frac{\rho_0}{32}(R^{(0)})^2 \ge (R^{(l)})^2(\delta^{(l)}+\cdots+\delta^{(n+1)})+(R'_{n+2})^2, \tag{3.4.10}$$

$$R^{(l)}-\frac{\rho_0}{32}R^{(0)} \ge (R^{(l)})^2(\delta^{(l)}+\cdots+\delta^{(n+1)})+R'_{n+2}.$$

Let us show the third one: since by (A.2.6) is $\rho_0/16\le 1-\exp(-\rho_0/8)$ (for $\rho_0/16<1$) and (see (3.4.8)) $R^{(\nu)}\ge R^{(0)}/2$ with $\nu=0,1,\dots$ for R small enough, we may write,

$$\begin{aligned}
R^{(l)}-\frac{\rho_0}{32}R^{(0)} &\ge R^{(l)}-R^{(n+2)}+\exp(-\rho_0/8)R^{(n+2)}\\
&\ge R^{(l)}(1-\exp(-2(\delta^{(l)}+\cdots+\delta^{(n+1)})))+R'_{n+2},
\end{aligned} \tag{3.4.11}$$

(where we have used that $K-6\ge 2$), but $\delta^{(l)}+\cdots+\delta^{(n+1)}\le\sum_{j\ge 0}\delta^{(j)}$, and it was shown that this sum is (for R small enough) less that $\rho_0/4$. Hence $2(\delta^{(l)}+\cdots+\delta^{(n+1)})<1$, so if we apply (A.2.6) to (3.4.11) the desired result follows (since $R^{(l)}\ge(R^{(l)})^2$). The inequalities (3.4.9) and (3.4.10) are derived in a similar way, so we shall not carry out them explicitly. This completes the proof of the last item and so those of the lemma. □

The difference $\check{\psi}^{(n+1)}-\check{\psi}^{(n)}$ can be written as:

$$\begin{aligned}
&\psi^{(0)}\circ\dots\psi^{(n+1)}-\psi^{(0)}\circ\dots\psi^{(n)}=\\
&\quad=(\psi^{(0)}-\mathrm{Id})\circ\psi^{(1)}\circ\cdots\circ\psi^{(n+1)}-(\psi^{(0)}-\mathrm{Id})\circ\psi^{(1)}\circ\cdots\circ\psi^{(n)}+\\
&\qquad+\psi^{(1)}\circ\cdots\circ\psi^{(n+1)}-\psi^{(1)}\circ\cdots\circ\psi^{(n)},
\end{aligned} \tag{3.4.12}$$

Consider the first term on the r. h. s., taking:

$$\begin{aligned}
&\varrho=\rho^{(1)}, &&\varrho_0=\rho'_{n+2}, &&\delta=\frac{\rho_0}{32}(R^{(0)})^2,\\
&R=R^{(1)}, &&R_0=R'_{n+2}, &&\chi=\frac{\rho_0}{16}R^{(0)}, &&\sigma=\frac{\rho_0}{32}(R^{(0)})^2,
\end{aligned}$$

item (iii) of 3.22 allows us to apply lemma A.5 of appendix A to give,

$$\begin{aligned}
&\|(\psi^{(0)}-\mathrm{Id})\circ\psi^{(1)}\circ\cdots\circ\psi^{(n+1)}-(\psi^{(0)}-\mathrm{Id})\circ\psi^{(1)}\circ\cdots\circ\psi^{(n)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}}\le\\
&\le\left(\tfrac{1}{\frac{\rho_0}{32}(R^{(0)})^2}+\tfrac{2}{\frac{\rho_0}{32}(R^{(0)})^2}+\tfrac{2}{\frac{\rho_0}{32}R^{(0)}}\right)\|\psi^{(1)}\circ\cdots\circ\psi^{(n+1)}-\psi^{(1)}\circ\cdots\circ\psi^{(n)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}}\times\\
&\qquad\times\|\psi^{(0)}-\mathrm{Id}\|_{\check{\mathcal{E}}^{(\infty)},\rho^{(1)},R^{(1)}}
\end{aligned}$$

and hence, using again (3.3.28) and (3.3.29) to bound $\|\psi^{(0)} - \mathrm{Id}\|_{\check{\mathcal{E}}^{(\infty)},\rho^{(1)},R^{(1)}}$, we arrive to,

$$\|\check{\psi}^{(n+1)} - \check{\psi}^{(n)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} \leq \left(1 + \Delta_{\rho_0}(\Pi M^{(0)})^{\frac{s}{2}-2\alpha}\right) \times \\ \times \|\psi^{(1)} \circ \cdots \circ \psi^{(n+1)} - \psi^{(1)} \circ \cdots \circ \psi^{(n)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}},$$

where Δ_{ρ_0} depends only on ρ_0. Now, the difference $\psi^{(1)} \circ \cdots \circ \psi^{(n+1)} - \psi^{(1)} \circ \cdots \circ \psi^{(n)}$ can be decomposed in the same way as in (3.4.12) and then bounded by the application of lemmas 3.22 and A.5. Therefore, the described process can be applied recursively, and so, taking α (and R) small enough, one obtains:

$$\begin{aligned}\|\check{\psi}^{(n+1)} - \check{\psi}^{(n)}\|_{\check{\mathcal{E}}^{(\infty)},\rho'_{n+2},R'_{n+2}} &\leq \prod_{j=0}^{n}\left(1 + \Delta_{\rho_0}(\Pi M^{(0)})^{\frac{s^{j+1}}{2}-2\alpha}\right)(\Pi M^{(0)})^{\frac{s^{n+2}}{2}} \\ &\leq 2(\Pi M^{(0)})^{\frac{s^{n+2}}{2}},\end{aligned}$$

where we have used the convergent character of $\sum_{j\geq 0}(\Pi M^{(0)})^{\frac{s^{j+1}}{2}}$, and the fact that this sum goes to zero when R does. Then, it is easy to prove (with the same ideas worked in lemma 3.21) that the product $\prod_{j=0}^{n}\left(1 + \Delta_{\rho_0}(\Pi M^{(0)})^{\frac{s^{j+1}}{2}-2\alpha}\right)$ tends to 1 when $R \to 0$. From the last bound it is now clear that, if $\nu > r \geq 0$, then

$$\|\check{\psi}^{(\nu)} - \check{\psi}^{(r)}\|_{\check{\mathcal{E}}^{(\infty)},\rho_0/8,R_0\exp(-3\rho_0/8)} \leq \sum_{j\geq r} 2(\Pi M^{(0)})^{\frac{s^{j+2}}{2}},$$

bound that goes to zero as $r, \nu \to \infty$. This allows us to conclude that the limit transformation, $\check{\psi}^{(\infty)}$, is well-defined and goes from $\mathcal{D}_*(\rho_0/8, R^{(0)}\exp(-3\rho_0/8))$ to $\mathcal{D}_*(\rho^{(0)}, R^{(0)})$.

3.5 Bounds on the measure

We have shown the existence, under the above specified conditions, of invariant reducible tori for the basic frequencies $\boldsymbol{\Lambda}^* = (\mu, \Omega_2) \in \check{\mathcal{E}}^{(\infty)}$. Now, our purpose is to obtain bounds for the measure (of the complementary) of a set of *real* frequencies, $\mathcal{E}^{(\infty)} \subseteq \check{\mathcal{E}}^{(\infty)}$. In fact, both $\mathcal{E}^{(\infty)}$ and $\check{\mathcal{E}}^{(\infty)}$ may be thought as limit sets of the sequences of domains $\{\mathcal{E}^{(\nu)}\}_{\nu\in\mathbb{Z}}$, $\{\check{\mathcal{E}}^{(\nu)}\}_{\nu\in\mathbb{Z}}$ holding real and complex frequencies respectively and it can be shown that, in the limit $\mathcal{E}^{(\infty)} \equiv \check{\mathcal{E}}^{(\infty)}$ (see remark 3.25 below). Firstly though, one should estimate, given a R-neighborhood of $\boldsymbol{\Lambda}^* = (0, \omega_2)$, the measure of its subset of (real) basic frequencies corresponding to values of (real) parameters $\boldsymbol{\zeta}^* = (\xi, \eta)$ with ξ negative (complex tori) or *too* small (in the sense of the restrictions (3.2.39)). These frequencies must be discarded from the initial step of the iterative scheme in the KAM process.

Consider the sets $\mathcal{U}$, $\mathcal{W}$, and $\mathcal{V}$ defined by (3.2.27), (3.2.28) and (3.2.29) respectively, and let us denote by ψ the map taking the frequencies $\boldsymbol{\Lambda}$ onto the space of parameters $\boldsymbol{\zeta}$, so we can write $\boldsymbol{\zeta} = \psi(\boldsymbol{\Lambda})$, and conversely, $\boldsymbol{\Lambda} = \psi^{-1}(\boldsymbol{\zeta})$. In fact, ψ^{-1}, is the map used to define the set $\mathcal{W}$ (and it can be checked, that, for R small enough, ψ^{-1} is nondegenerate under our early assumptions $\xi > 0$ and $a > 0$). Hence we introduce the set,

$$\mathcal{E}^{(0)}(R) = \mathcal{W}\left(\tfrac{1}{8}\check{c}R\right) \setminus \psi^{-1}\left(\mathcal{V}\left(64(M(R))^{\alpha}\right)\right), \tag{3.5.1}$$

i. e., we must eliminate from $\mathcal{W}(\frac{1}{8}\check{c}R)$, the measure filled by the pre-image of $\widetilde{\mathcal{V}}(R) = \mathcal{V}(64M^\alpha) \cap \psi(\mathcal{W}(\frac{1}{8}\check{c}R))$. Let us denote $\widetilde{\mathcal{W}}(R) = \psi^{-1}(\mathcal{V}(64M^\alpha)) \cap \mathcal{W}(\frac{1}{8}\check{c}R)$. Therefore,

$$\text{meas}\left(\widetilde{\mathcal{W}}(R)\right) = \int\limits_{\widetilde{\mathcal{W}}(R)} d\boldsymbol{\Lambda} = \int\limits_{\widetilde{\mathcal{V}}(R)} |\det D\psi^{-1}(\boldsymbol{\zeta})| d\boldsymbol{\zeta},$$

but direct computation (applying the bounds for $\mathcal{Z}_3$ in (3.2.9) and (3.2.10)) shows that, $|\det D\psi^{-1}(\boldsymbol{\zeta})| \leq \text{constant} \cdot \mu^{-1}$ and using the first of (3.2.42) –in remark 3.12–, we have $\mu^{-1} \leq \xi^{-1/2}$, so

$$\begin{aligned} \text{meas}\left(\widetilde{\mathcal{W}}(R)\right) &\leq \int\limits_{\widetilde{\mathcal{V}}(R)} constant \cdot \xi^{-1/2} d\xi d\eta \\ &\leq constant \cdot R \int\limits_0^{64M^\alpha} \xi^{-1/2} d\xi \leq constant' \cdot M^{\alpha/2}. \end{aligned} \tag{3.5.2}$$

This justifies that meas $(\widetilde{\mathcal{W}}(R))$ is of order $(M(R))^{\alpha/2}$.

3.5.1 Measure of the complementary of $\mathcal{E}^{(\infty)}$

As it has been already pointed out at the beginning of section 3.5, the set we are going to measure is the complementary of $\mathcal{E}^{(\infty)}$, which is given as the limit set of the sequence $\{\mathcal{E}^{(\nu)}\}_{\nu\in\mathbb{Z}_+}$, (with $\mathcal{E}^{(\nu)} \subset \mathbb{R}^2$, for all $\nu \in \mathbb{Z}_+$). Below, we specify, recursively from $\mathcal{E}^{(0)}$, the elements of this sequence and use them to define the sets $\check{\mathcal{E}}^{(\nu)}$ –i. e., those holding the *complex* frequencies taken at the ν-th step of the iterative scheme, see (3.3.20) in section 3.3–.

Thus, if $\mathcal{E}^{(0)}$ is the domain introduced in the previous section –equation (3.5.1)–, for $\nu = 1, 2, \ldots$; $\mathcal{E}^{(\nu)}$ will be the set

$$\begin{aligned} \mathcal{E}^{(\nu)} \equiv \Big\{ &\boldsymbol{\Lambda} \in \mathcal{E}^{(\nu-1)} : |k_1 \Omega_1^{(\nu-1)}(\boldsymbol{\Lambda}) + k_2\Omega_2 + \ell\mu| \geq \frac{3\gamma^{(\nu-1)}(R)}{|\boldsymbol{k}|_1^\tau}, \\ &\text{for } \boldsymbol{k} \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\}, \text{ with } |\boldsymbol{k}|_1 < 2N^{(\nu-1)}; \text{ and } \ell \in \mathbb{Z}, \text{ with } 0 \leq |\ell| \leq 2 \Big\} \end{aligned} \tag{3.5.3}$$

Definition 3.23 (Complex σ-widening). *Given a set $\mathcal{D} \subseteq \mathbb{R}^2$ and a real positive $\sigma > 0$, one defines the complex σ-widening of $\mathcal{D}$, $\mathcal{D} + \sigma$, as*

$$\mathcal{D} + \sigma = \bigcup_{\boldsymbol{z}\in\mathcal{D}} \{\boldsymbol{z}' \in \mathbb{C}^2 : |\boldsymbol{z} - \boldsymbol{z}'| \leq \sigma\},$$

i. e., $\mathcal{D} + \sigma$ is the union of all (complex) balls of radius σ centered at points of $\mathcal{D}$.

With this last definition we introduce $\check{\mathcal{E}}^{(\nu)}$ as a $2r^{(\nu)}$-widening of $\mathcal{E}^{(\nu)}$, i. e.:

$$\check{\mathcal{E}}^{(\nu)} = \mathcal{E}^{(\nu)} + 2r^{(\nu)}, \qquad \nu = 1, 2, \ldots \tag{3.5.4}$$

Let $N^{(\nu-1)}$ be the order up to which the harmonics (of the "obstructing term", $\widehat{H}^{(\nu-1)}$) are removed at the ν-th step of the iterative process. It was shown, at the end of section 3.3 that it suffices to take:

$$N^{(\nu-1)} = \lfloor \mathcal{N}^{(\nu-1)} \rfloor + 1, \tag{3.5.5}$$

being

$$\mathcal{N}^{(\nu-1)} = \frac{\ln \widehat{N} + (1 - 42\alpha - 8\alpha\tau) \ln M^{(\nu-1)}}{-K\delta^{(\nu-1)}}, \tag{3.5.6}$$

with $K \geq 9$. See (3.3.31).

With these previous items we can state lemma 3.26, which gives the convenient choice (in the sense of remark 3.25) for the width $r^{(\nu)}$. It is left without proof because this is included in that of lemma 3.26 (see below).

Lemma 3.24. *If the thickness $r^{(\nu)}$ in (3.5.4) is given by,*

$$r^{(\nu)} = \frac{\gamma^{(\nu-1)}}{c_2}(2N^{(\nu-1)})^{-\tau-1}, \qquad \nu = 1, 2, \ldots \tag{3.5.7}$$

(being c_2 a constant independent of R), then for $\nu = 1, 2, \ldots$; $m = 1, 2$:

$$\boldsymbol{\Lambda} \in \mathcal{E}_m^{(\nu)} + mr^{(\nu)} \Rightarrow |k_1 \Omega_1^{(\nu-1)}(\boldsymbol{\Lambda}) + k_2\Omega_2 + \ell\mu| \geq (3-m)\frac{\gamma^{(\nu-1)}(R)}{|\boldsymbol{k}|_1^\tau},$$
$$\textit{for } \boldsymbol{k} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\} \textit{ with } |\boldsymbol{k}|_1 < 2N^{(\nu-1)} \textit{ and } \ell \in \mathbb{Z} \textit{ with } 0 \leq |\ell| \leq 2. \tag{3.5.8}$$

Remark 3.25. Thus, with such choice of $r^{(\nu)}$, the sets $\mathcal{E}^{(\nu)} + 2r^{(\nu)}$ satisfy the Diophantine conditions (3.3.20) as required. This justifies the definition, in (3.5.4), of $\breve{\mathcal{E}}^{(\nu)}$ as a complex $2r^{(\nu)}$-widening of $\mathcal{E}^{(\nu)}$, for $\nu = 1, 2 \ldots$; however, it is necessary to ask $r^{(1)}$ to be small enough to make $\mathcal{E}^{(1)} + 2r^{(1)} \subseteq \breve{\mathcal{E}}^{(0)}$. One may always achieve this taking R sufficiently small, since form (3.5.7), (3.5.5), (3.5.6) –we recall that $\gamma^{(\nu-1)} = (M^{(\nu-1)})^{\alpha}$–, the limit

$$r^{(\nu)}(R) \to 0^+ \quad \text{ when } R \to 0^+,$$

follows immediately for every $\nu = 1, 2, \ldots$ fixed. Furthermore, we have[8]: $r^{(\nu)} \leq r^{(\nu-1)}$, $\nu = 1, 2, \ldots$ and $\lim_{\nu\to\infty} r^{(\nu)} = 0^+$ (for any R fixed and sufficiently small). Therefore, since by their construction (see (3.5.3)) is $\mathcal{E}^{(\nu)} \subseteq \mathcal{E}^{(\nu-1)}$, then: $\mathcal{E}^{(\nu)} + mr^{(\nu)} \subseteq \mathcal{E}^{(\nu-1)} + mr^{(\nu-1)}$ for $\nu = 2, 3, \ldots$; $m = 1, 2$. This last point allow us to consider $\breve{\mathcal{E}}^{(\infty)}$ also as the limit set of the sequence $\{\breve{\mathcal{E}}^{(\nu)}\}_{\nu\in\mathbb{Z}_+}$; but $r^{(\nu)} \to 0^+$ when $\nu \to +\infty$ so $\breve{\mathcal{E}}^{(\infty)} \equiv \mathcal{E}^{(\infty)}$, as pointed out at the beginning of section 3.5. ♠

To give estimates of the measure of $\mathcal{E}^{(0)} \setminus \mathcal{E}^{(\infty)}$, we may consider this set as the limit of the collection of sets: $\mathcal{E}^{(0)} \setminus \mathcal{E}^{(s)}$, which, in turn, can be split as the disjoint union:

$$\mathcal{E}^{(0)} \setminus \mathcal{E}^{(s)} = \bigcup_{j=0}^{s-1} \mathcal{E}^{(j)} \setminus \mathcal{E}^{(j+1)}, \tag{3.5.9}$$

and hence, one should control the measure of $\mathcal{E}^{(j)} \setminus \mathcal{E}^{(j+1)}$ for $j \geq 0$. To do this, we consider the decomposition

$$\mathcal{E}^{(j)} \setminus \mathcal{E}^{(j+1)} = \bigcup_{0\leq|\ell|\leq 2} \bigcup_{\substack{\boldsymbol{k}\in\mathbb{Z}^2\setminus\{\mathbf{0}\} \\ |\boldsymbol{k}|_1<2N^{(j)}}} \mathcal{R}_{\ell,\boldsymbol{k}}^{(j)}, \tag{3.5.10}$$

[8] If R is kept fixed –and small enough–, $N^{(\nu)}$ monotonically increases to infinity and $\gamma^{(\nu)}$ monotonically decreases to zero with $\nu \in \mathbb{N}$.

being

$$\mathcal{R}^{(j)}_{\ell,\boldsymbol{k}} = \left\{ \boldsymbol{\Lambda} \in \mathcal{E}^{(j)} : |k_1 \Omega^{(j)}_1(\boldsymbol{\Lambda}) + k_2\Omega_2 + \ell\mu| < \frac{3\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}} \right\}, \tag{3.5.11}$$

note, that the set $\mathcal{R}^{(j)}_{l,\boldsymbol{k}}$ is a slice of the set $\mathcal{U}(\frac{1}{8}\check{c}R) = \{\boldsymbol{\Lambda} \in \mathbb{R}^2 : |\boldsymbol{\Lambda} - \boldsymbol{\Lambda}_0| \leq \frac{1}{8}\check{c}R\}$, with $\boldsymbol{\Lambda}_0^* = (0, \omega_2)$ –see (3.2.27)–. Therefore, we take $\boldsymbol{\Lambda}^{(1)}, \boldsymbol{\Lambda}^{(2)} \in \mathcal{R}^{(j)}_{\ell,\boldsymbol{k}}$ in this set and such that $\boldsymbol{\Lambda}^{(1)} - \boldsymbol{\Lambda}^{(2)}$ is (approximately) parallel to the vector (ℓ, k_2), with $(\ell, k_2) \neq (0,0)$ (see remark 3.27 below). Thus, the measure of $\mathcal{R}^{(j)}_{\ell,\boldsymbol{k}}$ can be estimated by the product of a bound of $|\boldsymbol{\Lambda}^{(1)} - \boldsymbol{\Lambda}^{(2)}|$ by a bound of the measure of the worst (i. e., the widest) section of a hyperplane (of codimension 1) with the set $\mathcal{U}(\frac{1}{8}\check{c}R)$.

Thus, if $\boldsymbol{\Lambda}^{(1)}$, $\boldsymbol{\Lambda}^{(2)} \in \mathcal{R}^{(j)}_{\ell,\boldsymbol{k}}$, the definition (of $\mathcal{R}^{(j)}_{\ell,\boldsymbol{k}}$) above and the triangle inequality lead to

$$|\langle(\ell, k_2), \boldsymbol{\Lambda}^{(1)} - \boldsymbol{\Lambda}^{(2)}\rangle| \leq \frac{6\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}} + |k_1| \times |\Omega^{(j)}_1(\boldsymbol{\Lambda}^{(1)}) - \Omega^{(j)}_1(\boldsymbol{\Lambda}^{(2)})| \tag{3.5.12}$$

and also from (3.5.11), if $\boldsymbol{\Lambda} \in \mathcal{R}^{(j)}_{\ell,\boldsymbol{k}}$:

$$|k_1| \leq \frac{1}{|\Omega^{(j)}_1(\boldsymbol{\Lambda})|} \left(\frac{3\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}} + |(\ell, k_2)|_2 |\boldsymbol{\Lambda}|_2 \right),$$

where $|\cdot|_2$ denotes the Euclidean norm. Now, we use that, if $\|(\Omega^{(0)}_1)^{-1}\|_{\check{\mathcal{E}}^{(0)}} \leq \widetilde{\mathfrak{g}}^{(0)}$ (being $\widetilde{\mathfrak{g}}^{(0)}$ a positive constant) then –for R small enough–,

$$\|(\Omega^{(j)}_1)^{-1}\|_{\check{\mathcal{E}}^{(j)}} \leq \widetilde{\mathfrak{g}}, \qquad j = 1, 2, \ldots$$

holds with, for instance, $\widetilde{\mathfrak{g}} = 2\widetilde{\mathfrak{g}}^{(0)}$. Here, we shall not check this last assertion, but point out that it can be shown, from the second inequality in (3.3.33), with the same recursive tricks used of section 3.4 to derive the step-independent bounds (3.4.5) for $\langle \mathcal{A}^{(j)} \rangle_{\boldsymbol{\theta}}^{-1}$. Hence k_1 will, for any $\boldsymbol{\Lambda} \in \widetilde{\mathcal{E}}^{(n)}$, satisfy the following inequality:

$$|k_1| \leq c_4 |(\ell, k_2)|_2 + c_5 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}}, \tag{3.5.13}$$

where c_4 and c_5 are two constants independent of R and on the step.

On the other hand:

$$\begin{aligned} |\boldsymbol{k}|_1 = |k_1| + |k_2| &\overset{\{\text{by (3.5.13)}\}}{\leq} c_4|(\ell, k_2)|_2 + c_5 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}} + |k_2| \\ &\leq (c_4 + 1) \times |(\ell, k_2)|_2 + c_5 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}}. \end{aligned}$$

But then,

$$(c_4 + 1) \times |(\ell, k_2)|_2 \geq |\boldsymbol{k}|_1 - c_5 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau}}$$

and if $0 < R < 1$ is taken small enough to make, for instance, $c_5 \frac{\gamma^{(0)}(R)}{|\boldsymbol{k}|_1^{\tau}} \leq \frac{1}{2}$:

$$
\begin{aligned}
(c_4 + 1) \times |(\ell, k_2)|_2 &\overset{(9)}{\geq} |\boldsymbol{k}|_1 - c_5 \frac{\gamma^{(0)}(R)}{|\boldsymbol{k}|_1^{\tau}} \\
&\geq |\boldsymbol{k}|_1 - \frac{1}{2} \geq |\boldsymbol{k}|_1 - \frac{1}{2}|\boldsymbol{k}|_1 = \frac{1}{2}|\boldsymbol{k}|_1,
\end{aligned}
$$

(for, obviously, $\frac{1}{2} \leq \frac{1}{2}|\boldsymbol{k}|_1$, if $|\boldsymbol{k}|_1 \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\}$). Therefore,

$$
|(\ell, k_2)|_2 \leq c_6 |\boldsymbol{k}|_1 \tag{3.5.14}
$$

holds, for $0 < R < 1$ sufficiently small, with c_6 a constant independent on R, $\boldsymbol{k}$, ℓ and on the step.

Thus, assuming now that $\boldsymbol{\Lambda}^{(1)} - \boldsymbol{\Lambda}^{(2)}$ is parallel to the vector (ℓ, k_2), the inequality (3.5.12) yields, if both (3.5.13) and (3.5.14) are taken into account:

$$
|\boldsymbol{\Lambda}^{(1)} - \boldsymbol{\Lambda}^{(2)}| \leq \frac{6\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau+1}} + \left(c_4 + c_7 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau+1}} \right) |\Omega_1^{(j)}(\boldsymbol{\Lambda}^{(1)}) - \Omega_1^{(j)}(\boldsymbol{\Lambda}^{(2)})|. \tag{3.5.15}
$$

The next "technical" lemma provides bounds for the difference $|\Omega_1^{(j)}(\boldsymbol{\Lambda}^{(1)}) - \Omega_1^{(j)}(\boldsymbol{\Lambda}^{(2)})|$. We delay its proof for the end of this section.

Lemma 3.26. *Given $\boldsymbol{\Lambda}, \boldsymbol{\Lambda}' \in \mathcal{E}^{(j)}$ the following inequality,*

$$
|\Omega_1^{(j)}(\boldsymbol{\Lambda}) - \Omega_1^{(j)}(\boldsymbol{\Lambda}')| \leq 3c_3 R |\boldsymbol{\Lambda} - \boldsymbol{\Lambda}'|, \qquad j = 0, 1, 2, \ldots, \tag{3.5.16}
$$

holds provided $R \leq 1$ is sufficiently small.

Applying the result of the lemma in (3.5.15), we can state, for R small enough:

$$
|\boldsymbol{\Lambda}^{(1)} - \boldsymbol{\Lambda}^{(2)}|_2 \leq \Pi_1 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau+1}}
$$

and since the diameter of $\mathcal{E}^{(0)}(R)$ can be bounded by $c_8 R$ –with c_8 an appropriate constant independent of R–, from the previous estimates one gets

$$
\operatorname{meas}(\mathcal{R}_{\ell,\boldsymbol{k}}^{(j)}) \leq \Pi_2 \frac{\gamma^{(j)}(R)}{|\boldsymbol{k}|_1^{\tau+1}} R;
$$

so, in view of (3.5.10):

$$
\operatorname{meas}(\mathcal{E}^{(j)} \setminus \mathcal{E}^{(j+1)}) \leq \Pi_3 R \gamma^{(j)}(R) \sum_{\boldsymbol{k} \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\}} \frac{1}{|\boldsymbol{k}|_1^{\tau+1}}, \tag{3.5.17}
$$

where Π_3 does not depend on R, and (to simplify) the sum is taken over all $\boldsymbol{k} \in \mathbb{Z} \setminus \{\boldsymbol{0}\}$, not only on those such that $|\boldsymbol{k}|_1 \leq 2N^{(j)}$. Moreover, we recall that for r, ν nonnegative integers, $r > 1$, is

$$
\#\{\boldsymbol{k} \in \mathbb{Z}^r : |\boldsymbol{k}|_1 = \nu\} \leq 2r\nu^{r-1}
$$

(9) Recall that $\gamma^{(j)} \equiv (M^{(j)})^{\alpha}$ and $M^{(j)}$ –for $0 < R < 1$ kept fixed and sufficiently small–, decreases monotonically with j. See footnote on page 128.

and therefore, the sum in the r. h. s. of the inequality (3.5.17) may be bounded by the sum $4\sum_{\nu\geq 1}\nu^{-\tau}$, which converges, since $\tau > 1$. Resuming (3.5.17),

$$\operatorname{meas}(\mathcal{E}^{(j)} \setminus \mathcal{E}^{(j+1)}) \leq 4\Pi_3 R\gamma^{(j)}(R)\sum_{\nu\geq 1}\nu^{-\tau} \leq \Pi_4 R\gamma^{(j)}(R),$$

with $\gamma^{(j)}(R) = (M^{(j)})^{\alpha}$. Finally, $\mathcal{E}^{(0)} \setminus \mathcal{E}^{(\infty)}$ was defined as the limit of the disjoint union (3.5.9), consequently:

$$\begin{aligned}\operatorname{meas}(\mathcal{E}^{(0)} \setminus \mathcal{E}^{(\infty)}) &= \sum_{j\geq 0}\operatorname{meas}(\mathcal{E}^{(j)} \setminus \mathcal{E}^{(j+1)})\\ &\leq \Pi_4 R\left((M^{(0)})^{\alpha} + \sum_{j\geq 1}(\Pi M^{(0)})^{s^j\alpha}\right) \leq 2\Pi_4(M^{(0)})^{\alpha},\end{aligned}$$

where we have used again the fact that, for $R \leq 1$ small enough, the sum

$$\sum_{n\geq 0}(\Pi M^{(0)})^{\alpha s^{j+1}}, \tag{3.5.18}$$

is convergent and tends to 0 with R.

Remark 3.27. In the derivation of the inequality (3.5.15) from (3.5.12), it has been assumed implicitly $(\ell, k_2) \neq (0,0)$. Actually this works for R small enough. Indeed, for if $(\ell, k_2) \neq (0,0)$ and R is sufficiently small, by (3.5.13):

$$|k_1| \leq c_5\frac{2\gamma^{(j)}(R)}{|k_1|^{\tau}},$$

but, this cannot occur even for $j = 0$, because k_1 must be a nonzero integer (since $\boldsymbol{k} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$ and $k_1 = 0$). ♠

Finally, taking into account that measure of $\mathcal{W}(\frac{1}{8}\check{c}R)\backslash\mathcal{E}^{(0)}(R)$ is, according with (3.5.2), of order $(M^{(0)})^{\alpha/2}$, we conclude that the measure of the set of frequencies $\boldsymbol{\Lambda} = (\mu, \Omega_2)$, corresponding to the destroyed tori is of order $(M^{(0)})^{\alpha/2}$. However, such a result is tied to lemma 3.26, which proof is still owed.

Proof of lemma 3.26. Let $\mathcal{N}^*$ be the set introduced by the second of (3.2.23). We suppose, that for $0 < R < 1$ small enough,

$$\left\|\frac{\partial \Omega_1^{(0)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{N}^*} \leq c_3 R \tag{3.5.19}$$

holds with c_3 a constant independent of R –see expansion (3.2.22)–. Actually, this estimate can be derived from (3.2.13), the bounds in (3.2.9), (3.2.10) and lemmas 3.2, 3.7, 3.8. However, it could be necessary to shrink slightly $\mathcal{N}^*$ (we do not give the details here).

From their construction –see (3.5.3)– it is obvious that always $\mathcal{E}^{(j)} \subseteq \mathcal{E}^{(j)} + mr^{(j)}$, $m = 1, 2$, $j = 1, 2, \ldots$ Then, using the Cauchy's inequalities:

$$\left\|\frac{\partial \Omega_1^{(j)}}{\partial \boldsymbol{\Lambda}} - \frac{\partial \Omega_1^{(j-1)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{E}^{(j)}+r^{(j)}} \leq \frac{\|\Omega_1^{(j)} - \Omega_1^{(j-1)}\|_{\check{\mathcal{E}}^{(j)}}}{r^{(j)}},$$

which implies,

$$\left\|\frac{\partial \Omega_1^{(j)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{E}^{(j)}+r^{(j)}} \leq \left\|\frac{\partial \Omega_1^{(j-1)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{E}^{(j-1)}+r^{(j-1)}} + \frac{\|\Omega_1^{(j)} - \Omega_1^{(j-1)}\|_{\breve{\mathcal{E}}^{(j)}}}{r^{(j)}}$$

($\mathcal{E}^{(j)} + r^{(j)} \subseteq \mathcal{E}^{(j-1)} + r^{(j-1)}$, see remark 3.25) and applying this recursively,

$$\left\|\frac{\partial \Omega_1^{(j)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{E}^{(j)}+r^{(j)}} \leq \left\|\frac{\partial \Omega_1^{(0)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{N}^*} + \sum_{\nu=1}^{j} \frac{\|\Omega_1^{(\nu)} - \Omega_1^{(\nu-1)}\|_{\breve{\mathcal{E}}^{(\nu)}}}{r^{(\nu)}}, \tag{3.5.20}$$

since it is assumed (remark 3.25) that $r^{(1)}$ is small enough to let $\mathcal{E}^{(1)} + 2r^{(1)} \subseteq \breve{\mathcal{E}}^{(0)}$ then, as $\breve{\mathcal{E}}^{(0)} \subseteq \mathcal{N}^*$, it should be $\mathcal{E}^{(1)} + r^{(1)} \subseteq \mathcal{N}^*$. Let $\boldsymbol{\Lambda}, \boldsymbol{\Lambda}' \in \mathcal{E}^{(j)}$. We shall distinguish two cases:

First case: if $|\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \leq r^{(j)}$, one may apply the mean value theorem (because the segment joining $\boldsymbol{\Lambda}'$ and $\boldsymbol{\Lambda}$ will be included in $\mathcal{E}^{(j)} + r^{(j)}$, so

$$\begin{aligned} |\Omega_1^{(j)}(\boldsymbol{\Lambda}') - \Omega_1^{(j)}(\boldsymbol{\Lambda})| &\leq \left\|\frac{\partial \Omega^{(j)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{E}^{(j)}+r^{(j)}} |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \\ &\leq \left(\left\|\frac{\partial \Omega^{(0)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{N}^*} + \sum_{\nu=1}^{j} \frac{\|\Omega_1^{(\nu)} - \Omega_1^{(\nu-1)}\|_{\breve{\mathcal{E}}^{(\nu)}}}{r^{(\nu)}}\right) \times |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}|. \end{aligned} \tag{3.5.21}$$

Second case: if $|\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| > r^{(j)}$. Then, using the triangle inequality,

$$\begin{aligned} &|\Omega_1^{(j)}(\boldsymbol{\Lambda}') - \Omega_1^{(j)}(\boldsymbol{\Lambda})| \leq \\ &\leq |\Omega_1^{(j)}(\boldsymbol{\Lambda}') - \Omega_1^{(j-1)}(\boldsymbol{\Lambda}')| + |\Omega_1^{(j)}(\boldsymbol{\Lambda}) - \Omega_1^{(j-1)}(\boldsymbol{\Lambda})| + |\Omega_1^{(j-1)}(\boldsymbol{\Lambda}') - \Omega_1^{(j-1)}(\boldsymbol{\Lambda})| \\ &\qquad \leq 2\|\Omega_1^{(j)} - \Omega_1^{(j-1)}\|_{\breve{\mathcal{E}}^{(j)}} + |\Omega_1^{(j-1)}(\boldsymbol{\Lambda}') - \Omega_1^{(j-1)}(\boldsymbol{\Lambda})|. \end{aligned}$$

This last inequality can be also applied recursively to give,

$$|\Omega_1^{(j)}(\boldsymbol{\Lambda}') - \Omega_1^{(j)}(\boldsymbol{\Lambda})| \leq \left(\left\|\frac{\partial \Omega_1^{(0)}}{\partial \boldsymbol{\Lambda}}\right\|_{\mathcal{N}^*} + \sum_{\nu=1}^{j} \frac{2\|\Omega_1^{(\nu)} - \Omega_1^{(\nu-1)}\|_{\breve{\mathcal{E}}^{(\nu)}}}{r^{(\nu)}}\right) \times |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}|. \tag{3.5.22}$$

So, by comparison with (3.5.21) of the first case, it can be seen that the difference, $|\Omega_1^{(j)}(\boldsymbol{\Lambda}') - \Omega_1^{(j)}(\boldsymbol{\Lambda})|$ is bounded according with (3.5.22) for all $\boldsymbol{\Lambda}', \boldsymbol{\Lambda} \in \mathcal{E}^{(j)}$.

Next, we take for the radii $r^{(\nu)}$ the value guessed in lemma 3.24, explicitly:

$$r^{(\nu)} = \frac{\gamma^{(\nu-1)}}{c_2}(2N^{(\nu-1)})^{-\tau-1}, \qquad \nu = 1, 2, \ldots,$$

where $N^{(\nu)}$ is the number of harmonics fixed by (3.5.5) and (3.5.6), whereas c_2 is a constant to be suitable fixed (see below). It turns out that such a choice of $r^{(\nu)}$ works for $\nu = 1$. To check this out, we take $\boldsymbol{\Lambda} \in \mathcal{E}^{(1)}$ and $\boldsymbol{\Lambda}'$ such that $|\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \leq mr^{(1)}$. Therefore, applying the triangle inequality,

$$\begin{aligned} |k_1\Omega_1^{(0)}(\boldsymbol{\Lambda}') + k_2\Omega_2' + \ell\mu'| \geq |k_1\Omega_1^{(0)}(\boldsymbol{\Lambda}) + k_2\Omega_2 + \ell\mu| - |k_1(\Omega_1^{(0)}(\boldsymbol{\Lambda}') - \Omega_1^{(0)}(\boldsymbol{\Lambda}))| \\ - |k_2(\Omega_2' - \Omega_2) + \ell(\mu' - \mu)| \end{aligned}$$

but, by the definition of $\mathcal{E}^{(1)}$ the the first term on the r. h. s. is bounded from below by $3\gamma^{(0)}(R)/|\boldsymbol{k}|_1$, and hence:

$$|k_1\Omega_1^{(0)}(\boldsymbol{\Lambda}')+k_2\Omega_2'+\ell\mu'| \geq \frac{3\gamma^{(0)}(R)}{|\boldsymbol{k}|_1^{\tau}} - 2(|k_2|+|\ell|)|\boldsymbol{\Lambda}'-\boldsymbol{\Lambda}| - |k_1| \times |\Omega_1^{(0)}(\boldsymbol{\Lambda}')-\Omega_1^{(0)}(\boldsymbol{\Lambda})|$$

$$\geq \frac{3\gamma^{(0)}(R)}{|\boldsymbol{k}|_1^{\tau}} - 2c_1|\boldsymbol{k}|_1 \times |\boldsymbol{\Lambda}'-\boldsymbol{\Lambda}| - |\boldsymbol{k}|_1 \times |\Omega_1^{(0)}(\boldsymbol{\Lambda}')-\Omega_1^{(0)}(\boldsymbol{\Lambda})|, \quad (3.5.23)$$

with $|\boldsymbol{k}|_1 \leq 2N^{(0)}$ and c_1 is such that $|k_2|+|\ell| \leq c_1|\boldsymbol{k}|_1$ for all $\boldsymbol{k} \in \mathbb{Z}^2 \setminus \{\boldsymbol{0}\}$ and $\ell \in \mathbb{Z}$, $0 \leq |\ell| \leq 2$. Then, using, the mean value theorem and (3.5.19):

$$|\Omega_1^{(0)}(\boldsymbol{\Lambda}')-\Omega_1^{(0)}(\boldsymbol{\Lambda})| \leq c_3R|\boldsymbol{\Lambda}'-\boldsymbol{\Lambda}|$$

so if we take, for example, $c_2 = 2c_1+2c_3$, the r. h. s. of (3.5.23) becomes greater than:

$$\frac{3\gamma^{(0)}}{|\boldsymbol{k}|_1^{\tau}} - 2mc_2N^{(0)}r^{(1)} \geq \frac{3\gamma^{(0)}}{|\boldsymbol{k}|_1^{\tau}} - 2mN^{(0)}\gamma^{(0)}(2N^{(0)})^{-\tau-1} \geq \frac{(3-m)\gamma^{(0)}}{|\boldsymbol{k}|_1^{\tau}}.$$

(since $|\boldsymbol{k}|_1 \leq 2N^{(0)} \Rightarrow (2N^{(0)})^{-\tau} \leq |\boldsymbol{k}|_1^{-\tau}$). Now assume (3.5.8) holds for $\nu = 2,\dots,j-1$ and $m = 1,2$ with the corresponding $r^{(\nu)}$ given by (3.5.7). Let us take $\boldsymbol{\Lambda} \in \mathcal{E}^{(j)}$ and, $\boldsymbol{\Lambda}'$ such that $|\boldsymbol{\Lambda}'-\boldsymbol{\Lambda}| \leq mr^{(j)}$. Then, as before:

$$\begin{aligned}|k_1\Omega_1^{(j-1)}(\boldsymbol{\Lambda}')+k_2\Omega_2'+\ell\mu'| \geq{}& \frac{3\gamma^{(j-1)}(R)}{|\boldsymbol{k}|_1^{\tau}} - 2c_1|\boldsymbol{k}|_1 \times |\boldsymbol{\Lambda}'-\boldsymbol{\Lambda}| \\ &- |\boldsymbol{k}|_1 \times |\Omega_1^{(j-1)}(\boldsymbol{\Lambda}')-\Omega_1^{(j-1)}(\boldsymbol{\Lambda})|, \quad (3.5.24)\end{aligned}$$

and one may use (3.5.22) to bound the difference $|\Omega_1^{(j-1)}(\boldsymbol{\Lambda}')-\Omega_1^{(j-1)}(\boldsymbol{\Lambda})|$ taking into account:

(a) the second of (3.3.33). i. e.,

$$\|\Omega_1^{(\nu)}-\Omega_1^{(\nu-1)}\|_{\mathcal{E}^{(\nu)}} \leq \widehat{N}(M^{(\nu-1)})^{1-22\alpha-3\alpha\tau},$$

(b) the inequality,

$$(M^{(\nu-1)})^{-2\alpha} \geq 2c_2^{\frac{1}{\tau+1}}N^{(\nu-1)}, \quad (3.5.25)$$

which is fulfilled for R small enough[10].

---

[10] In fact, it is easier to prove, the inequality:

$$\frac{\mathcal{N}^{(\nu-1)}+1}{(M^{(\nu-1)})^{-2\alpha}} < 2c_2^{-\frac{1}{\tau+1}}, \quad (*)$$

(with $\mathcal{N}^{(\nu-1)}$ given by (3.5.6)), which implies (3.5.25) –since $\mathcal{N}^{(\nu-1)}+1 > N^{(\nu-1)}$, as follows from (3.5.5) and (3.5.6)–. To check $(*)$, it suffices to write the quotient on the l. h. s. on the form:

$$\frac{\mathcal{N}^{(\nu-1)}+1}{(M^{(\nu-1)})^{-2\alpha}} = \left(x^2 - d_3x + d_4x\ln\frac{1}{x}\right) \circ (M^{(\nu-1)})^{\alpha},$$

with $d_3 = \frac{1}{9}\ln\widehat{N}$, $d_4 = \frac{1-40\alpha-8\alpha\tau}{9\alpha}$. Thus $(*)$ follows if one realizes that $x^2 + x(d_4\ln\frac{1}{x} - d_3) \to 0^+$, when $x \to 0^+$.

In particular, from this last one, it follows:

$$r_\nu = \frac{\gamma^{(\nu-1)}}{c_2}(2N^{(\nu-1)})^{-\tau-1} \geq \gamma^{(\nu-1)}(M^{(\nu-1)})^{2\alpha(\tau+1)}.$$

(being $\gamma^{(\nu-1)} = (M^{(\nu-1)})^\alpha$). With these items (and also with the estimate on the derivative of $\Omega_1^{(0)}$ in (3.5.19)) on mind,

$$\begin{aligned} |\Omega_1^{(j-1)}(\boldsymbol{\Lambda}') - \Omega_1^{(j-1)}(\boldsymbol{\Lambda})| &\leq \left(c_3 R + 2\sum_{\nu=1}^{j-1} \frac{\widehat{N}(M^{(\nu-1)})^{1-23\alpha-4\alpha\tau}}{(M^{(\nu-1)})^\alpha (M^{(\nu-1)})^{2\alpha(\tau+1)}} \right) |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \\ &\leq \left(c_3 R + 2\sum_{\nu\geq 0} (\Pi M^{(0)})^{\frac{s^{\nu+1}}{2}} \right) |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}|, \end{aligned} \tag{3.5.26}$$

where the inequality $\widehat{N}(M^{(\nu-1)})^{\frac{s}{2}} \leq (\Pi M^{(0)})^{\frac{s^\nu}{2}}$ has been applied and we have re-defined $s = 2(1 - 26\alpha - 6\alpha\tau)$ asking α to be sufficiently small to let $s > 1$. Furthermore,

$$2\sum_{\nu\geq 0} (\Pi M^{(0)})^{\frac{s^{\nu+1}}{2}} \leq c_3 R, \tag{3.5.27}$$

will hold for $0 < R < 1$ sufficiently small, and then, resuming (3.5.26),

$$|\Omega_1^{(j-1)}(\boldsymbol{\Lambda}') - \Omega_1^{(j-1)}(\boldsymbol{\Lambda})| \leq (c_3 R + c_3 R)|\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \leq 2c_3 |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}|.$$

Now, this last bound can be substituted in the last term of the r. h. s. of (3.5.24) to give:

$$\begin{aligned} |k_1 \Omega_1^{(j-1)}(\boldsymbol{\Lambda}') + k_2 \Omega_2' + \ell\mu'| &\geq \frac{3\gamma^{(j-1)}}{|\boldsymbol{k}|_1^\tau} - (2c_1 + 2c_3)|\boldsymbol{k}|_1 |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \\ &\geq \frac{3\gamma^{(j-1)}}{|\boldsymbol{k}|_1^\tau} - mc_2 |\boldsymbol{k}|_1 \frac{\gamma^{(j-1)}}{c_2}(2N^{(j-1)})^{-\tau-1} \\ &\geq \frac{3\gamma^{(j-1)}}{|\boldsymbol{k}|_1^\tau} - m|\boldsymbol{k}|_1 \gamma^{(j-1)} |\boldsymbol{k}|_1^{-\tau-1} \\ &= \frac{(3-m)\gamma^{(j-1)}}{|\boldsymbol{k}|_1^\tau}. \end{aligned}$$

This completes the induction, so (3.5.8) holds for all $\nu \in \mathbb{N}$. Finally, from the inequality (3.5.22), using (3.5.7) for $r^{(\nu)}$, $\nu = 1, 2, \ldots$, and the inequalities (3.5.25), (3.5.27), one derives:

$$\begin{aligned} |\Omega_1^{(j)}(\boldsymbol{\Lambda}') - \Omega_1^{(j)}(\boldsymbol{\Lambda})| &\leq \left(c_3 R + 2\sum_{\nu\geq 0} (\Pi M^{(0)})^{\frac{\nu^{\nu+1}}{2}} \right) |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}| \\ &\leq 3c_3 R |\boldsymbol{\Lambda}' - \boldsymbol{\Lambda}|, \qquad \text{for } j = 1, 2, \ldots \end{aligned}$$

And we are done with the proof of the lemma. $\square$

So far, in what concerns the preservation of the invariant tori, the proof theorem 3.9 is almost done. To finish, we define the set

$$\mathfrak{A} = \bigcup_{0<R<R_*} \mathcal{E}^{(\infty)}(R),$$

being R_* the largest value of R for which the iterative scheme converges. Hence, most of those bifurcating elliptic two-dimensional invariant tori found from the normal form in chapter 1, will survive in the complete system. To be more precise, we have seen that, if the formal (and so, the unperturbed) elliptic family of invariant tori is re-parameterized by the normal frequency and one of its internal frequencies –the vector $\boldsymbol{\Lambda}^* = (\mu, \Omega_2)$ in our notation–, then, the control on the small divisors needed by the KAM-iterative scheme, give rise to gaps in the initial set of such "mixed" frequencies. These gaps are spread at each iterate, rarefying the successive frequency sets which degenerate, in the limit, to a Cantorian set with "plenty" measure, in the sense of small measure of its complementary (in fact, of the order of the remainder of the normal form raised to a positive constant).

3.5.2 Whitney-smoothness of the surviving tori

The third item of theorem 3.9 follows from the application of the *inverse approximation lemma* (see Zehnder, 1975) to the sequence of compositions of time-one flow transformations, $\{\breve{\psi}^{(j)}\}_{j\in\mathbb{Z}_+}$, defined in section 3.4.1. This will prove that $\Psi \equiv \lim_{j\to\infty} \breve{\psi}^{(j)}$ is $C^{\infty}_{Wh}(\mathfrak{A})$ (i. e., regular in the sense of Whitney) and also give a bound for the norm: $|||\Psi|||_{\mathfrak{A}}$, where $|||\cdot|||_{\mathfrak{A}}$, denotes the norm in $C^{\infty}_{Wh}(\mathfrak{A})$ (for a definition of this norm and an account on Whitney-smoothness, see Broer, Huitema and Sevryuk, 1996, chapter 6).

At this point, we quote the Whitney extension theorem (see Whitney, 1934), as appears in Wiggings and Rudnev (1997).

Theorem 3.28 (The extension theorem). *For any closed set $A \subset \mathbb{R}^n$ there exists a (non-unique) linear extension operator $\mathcal{F} : C^{\infty}_{Wh}(A) \to C^{\infty}(R^2)$; $u \mapsto U \equiv \mathcal{F}u$, such that for all the derivatives of u in the sense of Whitney, $D^kU|_A \leq u_k$, $k = 0, 1, \ldots$ and $|||U|||_{\mathbb{R}^n} \leq c|||u|||_A$, with c a constant depending only on n, but not on the set A.*

For definitions of Whitney differentiation we refer Pöschel (1989). If the function u depends on other variables, and if this dependence is analytic, smooth or periodic, then (quoting from the book of Broer et al.): such variables should be treated as parameters and an extension operator $\mathcal{F}$ can be chosen preserving analyticity (respectively smoothness) as well as periodicity with respect to these parameters.

Thus, selecting an appropriate operator, $\mathcal{F}$ one can extend Ψ, defined in $\mathbb{T}^2 \times \mathfrak{A}$ to $\widetilde{\Psi} \equiv \mathcal{F}\Psi$, defined in $\mathbb{T}^2 \times \mathbb{R}^2$. Alternatively, since Ψ parametrizes a Cantorian manifold of invariant tori, the extended function $\widetilde{\Psi}$ can be thought of a parametrization of the completed manifold. Let us now define $\widetilde{\mathfrak{A}}(R) \equiv \mathbb{T}^2 \times (\mathcal{W}(\frac{1}{8}\check{c}R) \setminus \mathfrak{A})$, $\mathcal{G}(R) \equiv \widetilde{\Psi}(\widetilde{\mathfrak{A}}(R))$. Therefore, as $\int_{\mathcal{G}(R)} \omega = \int_{\widetilde{\mathfrak{A}}(R)} \widetilde{\Psi}^*\omega$ (here ω is the volume form on the manifold and $\widetilde{\Psi}^*$ is the pullback of $\widetilde{\Psi}$). Applying the estimates for the derivatives of $\widetilde{\Psi}$ furnished by the extension theorem and the inverse approximation lemma, it can be seen that the measure of the "gaps" on the completed manifold is bounded by: *constant* $\cdot \int_{\widetilde{\mathfrak{A}}(R)} d\boldsymbol{\theta} \wedge d\boldsymbol{\Lambda} \leq$ *constant* $\cdot M^{\alpha/2}(R)$, i. e., they are of the same order than the gaps generated by the KAM process in the domain of basic frequencies.

3.5.3 Some final remarks

To conclude the chapter we give some hints addressed to fill those aspects not mentioned explicitly in the previous sections.

Remark 3.29. We may wonder about the real character of the constructed invariant tori. Actually, one can set up the same iterative scheme described in section 3.3 using the real Hamiltonian (developed around a selected unperturbed invariant torus), establish the corresponding homological equations, complexify, by means of the change (3.2.35) but *only to solve them*, and apply the inverse transformation to obtain a real generating function (and so a new real transformed Hamiltonian). In this way one construct a sequence of real Hamiltonians, $\{\mathcal{H}^{(j)}\}_{j\in\mathbb{N}\cup\{0\}}$ which, in virtue of the properties of the symplectic transformations satisfy $H^{(j)} = \mathcal{H}^{(j)}\circ\mathcal{F}$, $j = 0, 1, \dots$ (where $\mathcal{F}$ stands for the complexifying change) so, by the unicity of the limit, it must hold also for the limit Hamiltonian. Thus, equivalently, complexification could be applied to the initial Hamiltonian, perform all the iterative steps and transform back the (complex) limit Hamiltonian to recover the real invariant tori. However, what is worth here, if one looks for real quasi-periodic solutions is to force $\xi > 0$ in the change (3.2.2). This was done throwing away those $\boldsymbol{\Lambda}$ which maps to some $\boldsymbol{\zeta}^* = (\xi, \eta)$ with $\xi \le 0$ (see (3.2.28)). Otherwise, letting $\xi < 0$, complex solutions of a real Hamiltonian are obtained. ♠

Remark 3.30. In theorem 3.9 it was assumed coefficient a in (3.2.7) is positive (we recall, that this case corresponds to the *direct* bifurcation in proposition 1.29 of chapter 1). When this is not the case, though, form the expression (3.2.19) for the characteristic exponents of the unperturbed tori, it can be seen –as already mentioned in section 3.2.2–, that these quasi-periodic solutions could correspond to whether elliptic or hyperbolic tori, on the sign of

$$\mu^2 = -4\eta^2 - 2a\xi - 2\xi\partial_{1,1}\mathcal{Z}_3(\xi, \mathcal{I}(\boldsymbol{\zeta}), 2\xi\eta)$$

is negative or positive respectively. Then it is clear that the curve $\mu(\xi, \eta) = 0$ –which, by the implicit function theorem can be put, locally at the origin of the (ξ, η) plane, as the graph of a function $\xi(\eta)$–, plays the rôle of the separatrix between both types of normal behavior and corresponds to a one parameter family of (unperturbed) *parabolic* tori.

The persistence of the elliptic tori for this case, can be stated using essentially the same tricks described along the chapter. Nevertheless those initial domains $\mathcal{U}$, $\mathcal{W}$ and $\mathcal{V}$ should be more accurately delimited to ensure that the values of the parameters $\boldsymbol{\zeta}^* = (\xi, \eta)$ range far enough from the separatrix. ♠

Appendix A

Basic lemmas

The lemmas A.1 to A.4 of this appendix state some basic properties of the norms (2.2.3) and (2.2.4), introduced in chapter 2 and now generalized to analytic functions depending on a higher number of angles and normal coordinates (see section below). These lemmas can be also found in Jorba and Villanueva (1997a, 1998).

A.1 Notation

Previous to the statement of the lemmas, we shall generalize the norm introduced at the beginning of chapter 2. For some $\rho > 0$, and $R > 0$, we define the set

$$\mathcal{D}_{r,m}(\rho, R) = \left\{(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) \in \mathbb{C}^r \times \mathbb{C}^m \times \mathbb{C}^r \times \mathbb{C}^m : |\mathrm{Im}\ \boldsymbol{\theta}| \leq \rho, |\boldsymbol{z}| \leq R, |\boldsymbol{I}| \leq R^2\right\}. \tag{A.1.1}$$

We put $\boldsymbol{z}^* = (\boldsymbol{x}^*, \boldsymbol{y}^*)$ and $|\cdot|$ denotes the supremum norm of a complex vector (the same notation will be used for the matrix norm induced). We shall consider analytic functions $f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})$ defined on $\mathcal{D}_{r,m}(\rho, R)$, and 2π-periodic with respect to $\boldsymbol{\theta}$. We write the Taylor expansion of f w. r. t. $\boldsymbol{z}$ and $\boldsymbol{I}$ as

$$f = \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} f_{\boldsymbol{l},\boldsymbol{s}}(\boldsymbol{\theta})\boldsymbol{z}^{\boldsymbol{l}}\boldsymbol{I}^{\boldsymbol{s}}, \tag{A.1.2}$$

(with $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$). In turn, the coefficients $f_{\boldsymbol{l},\boldsymbol{s}}(\boldsymbol{\theta})$ can be expanded in Fourier series,

$$f_{\boldsymbol{l},\boldsymbol{s}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^r} f_{\boldsymbol{l},\boldsymbol{s},\boldsymbol{k}} \exp(\mathrm{i}\langle\boldsymbol{k}, \boldsymbol{\theta}\rangle). \tag{A.1.3}$$

We use the expansions (A.1.2) and (A.1.3), to introduce the norms,

$$|f_{\boldsymbol{l},\boldsymbol{s}}|_\rho = \sum_{\boldsymbol{k}\in\mathbb{Z}^r} |f_{\boldsymbol{l},\boldsymbol{s},\boldsymbol{k}}| \exp(|\boldsymbol{k}|_1\rho), \tag{A.1.4}$$

$$|f|_{\rho,R} = \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} |f_{\boldsymbol{l},\boldsymbol{s}}|_\rho R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1}, \tag{A.1.5}$$

As in chapter 2, we remark here that, when the sums defining them are convergent, these norms bound the supremum norms of $f_{\boldsymbol{l},\boldsymbol{s}}$ (on the complex strip of width $\rho > 0$) and of f (on $\mathcal{D}_{r,m}(\rho, R)$).

Moreover, from its definition, the Poisson bracket of two functions depending on $(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})$ is given by,

$$\{f, g\} = \frac{\partial f}{\partial \boldsymbol{\theta}} \left(\frac{\partial g}{\partial \boldsymbol{I}}\right)^* - \frac{\partial f}{\partial \boldsymbol{I}} \left(\frac{\partial g}{\partial \boldsymbol{\theta}}\right)^* + \frac{\partial f}{\partial \boldsymbol{z}} J_m \left(\frac{\partial g}{\partial \boldsymbol{z}}\right)^* .$$

A.2 Lemmata

Lemma A.1. *Let $f(\boldsymbol{\theta})$ and $g(\boldsymbol{\theta})$ be analytic functions of r complex arguments, 2π-periodic*[1] *in $\boldsymbol{\theta}$, with $|\mathrm{Im}\,\boldsymbol{\theta}| \leq \rho$, and taking values in $\mathbb{C}$. If $\{f_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^r}$ denote the Fourier coefficients of f, $f = \sum_{\boldsymbol{k}\in\mathbb{Z}^r} f_{\boldsymbol{k}} \exp(i\langle \boldsymbol{k}, \boldsymbol{\theta}\rangle)$. Then we have:*

(i) $|f_k| \leq |f|_\rho \exp(-|\boldsymbol{k}|_1 \rho)$.

(ii) $|fg|_\rho \leq |f|_\rho |g|_\rho$.

(iii) *For every $0 < \delta < \rho$,*

$$\left|\frac{\partial f}{\partial \theta_j}\right|_{\rho-\delta} = \frac{|f|_\rho}{\delta \exp(1)}, \qquad j = 1, \ldots, r.$$

(iv) Let $\{d_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^r\setminus\{\boldsymbol{0}\}} \subset \mathbb{C}$, with $|d_{\boldsymbol{k}}| \geq \mu/|\boldsymbol{k}|_1^\gamma$, for some $\mu > 0$ and $\gamma \geq 0$. If we assume that the average of f over the angles $\boldsymbol{\theta}$ is zero, i. e., $\langle f\rangle_{\boldsymbol{\theta}} = 0$, then, for any $0 < \delta < \rho$, we have that the function g defined as

$$g(\boldsymbol{\theta}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^r\setminus\{0\}} \frac{f_k}{d_k} \exp(i\langle \boldsymbol{k}, \boldsymbol{\theta}\rangle),$$

satisfies the bound

$$|g|_{\rho-\delta} \leq \left(\frac{\gamma}{\delta \exp(1)}\right)^\gamma \frac{|f|_\rho}{\mu}.$$

All these bounds can be extended to the case in which f and g can take values in $\mathbb{C}^{n_1}$ or $\mathbb{M}_{n_1,n_2}(\mathbb{C})$ (the space of the $n_1 \times n_2$ matrices with coefficients in $\mathbb{C}$).

Proof. (i) It follows directly from the definition of the norm (2.2.2):

$$|f|_\rho = \sum_{\boldsymbol{k}\in\mathbb{Z}^r} |f_k| \exp(|k|_1\rho) \geq |f_k| \exp(|\boldsymbol{k}|_1 \rho),$$

for all $\boldsymbol{k} \in \mathbb{Z}^r$, and hence $|f_k| \leq |f|_\rho \exp(-|\boldsymbol{k}|_1\rho)$. The second item is also straightforward. Indeed,

$$\begin{aligned} |fg|_\rho &= \left| \sum_{\boldsymbol{k},\boldsymbol{k}'\in\mathbb{Z}^r} f_{\boldsymbol{k}} g_{\boldsymbol{k}'} \exp\left(i\langle \boldsymbol{k}+\boldsymbol{k}', \boldsymbol{\theta}\rangle\right)\right|_\rho \\ &= \sum_{\boldsymbol{k},\boldsymbol{k}'\in\mathbb{Z}^r} |f_{\boldsymbol{k}}||g_{\boldsymbol{k}'}| \exp\left(|\boldsymbol{k}+\boldsymbol{k}'|_1\rho\right) \\ &\leq \sum_{\boldsymbol{k},\boldsymbol{k}'\in\mathbb{Z}^r} |f_{\boldsymbol{k}}||g_{\boldsymbol{k}'}| \exp(|\boldsymbol{k}|_1\rho)\exp(|\boldsymbol{k}'|_1\rho) \\ &= |f|_\rho |g|_\rho, \end{aligned}$$

[1] In this context it means $f(\boldsymbol{\theta} + 2\boldsymbol{\nu}\pi) = f(\boldsymbol{\theta})$ for all $\nu \in \mathbb{Z}^r$, and the same for g.

since, evidently $\exp(|\boldsymbol{k}+\boldsymbol{k}'|_1\rho) \leq \exp(|\boldsymbol{k}|_1\rho)\exp(|\boldsymbol{k}'|_1\rho)$, for all $\boldsymbol{k},\boldsymbol{k}' \in \mathbb{Z}^r$. To prove (iii) and (iv), we shall use that for any $\gamma > 0$ and $\delta > 0$,

$$\sup_{x\geq 1}\{x^\gamma \exp(-\delta x)\} \leq \left(\frac{\gamma}{\delta\exp(1)}\right)^\gamma, \tag{A.2.1}$$

for it can be seen, directly, that the function $h(x) = x^\gamma\exp(-\delta x)$, with $\gamma,\delta > 0$ and restricted to $x \geq 0$, reach its absolute maximum in this interval at $x = \gamma/\delta$, but $h(\frac{\gamma}{\delta}) = \left(\frac{\gamma}{\delta\exp(1)}\right)^\gamma$. Therefore,

$$\begin{aligned}\left|\frac{\partial f}{\partial\theta_j}\right|_{\rho-\delta} &= \sum_{\boldsymbol{k}\in\mathbb{Z}^r\setminus\{\boldsymbol{0}\}} |k_j||f_{\boldsymbol{k}}|\exp(|\boldsymbol{k}|_1(\rho-\delta)) \\ &\leq \sum_{\boldsymbol{k}\in\mathbb{Z}^r\setminus\{\boldsymbol{0}\}} |\boldsymbol{k}|_1\exp(-|\boldsymbol{k}|_1\delta)|f_{\boldsymbol{k}}|\exp(|\boldsymbol{k}|_1\rho) \\ &\leq \frac{1}{\delta\exp(1)}\sum_{\boldsymbol{k}\in\mathbb{Z}^r}|f_{\boldsymbol{k}}|\exp(|\boldsymbol{k}|_1\rho).\end{aligned}$$

This proves (iii); here, k_j, $j = 1,\ldots,r$ is the j-th component of the integer vector $\boldsymbol{k} \in \mathbb{Z}^r$, then $|k_j| \leq |\boldsymbol{k}|_1$. Moreover, inequality (A.2.1) with $\gamma = 1$ has been applied. Similarly, for the last item,

$$\begin{aligned}|g|_{\rho-\delta} &= \sum_{\boldsymbol{k}\in\mathbb{Z}^r\setminus\{\boldsymbol{0}\}}\left|\frac{f_{\boldsymbol{k}}}{d_{\boldsymbol{k}}}\right|\exp(|\boldsymbol{k}|_1(\rho-\delta)) \\ &\leq \sum_{\boldsymbol{k}\in\mathbb{Z}^r\setminus\{\boldsymbol{0}\}}\frac{|\boldsymbol{k}|_1^\gamma}{\mu}|f_{\boldsymbol{k}}|\exp(|\boldsymbol{k}|_1\rho)\exp(-|\boldsymbol{k}|_1\delta),\end{aligned}$$

but, as $|\boldsymbol{k}|_1 \geq 1$ for all $\boldsymbol{k} \in \mathbb{Z}^r \setminus \{\boldsymbol{0}\}$, it is $|\boldsymbol{k}|_1^\gamma\exp(-|\boldsymbol{k}|_1\delta) \leq \left(\frac{\gamma}{\delta\exp(1)}\right)^\gamma$, by (A.2.1). Then (iv) immediately follows. □

Lemma A.2. *Let $f(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})$ and $g(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})$ be analytic functions defined on $\mathcal{D}_{r,m}(\rho,R)$ and 2π-periodic on $\boldsymbol{\theta}$. Then,*

(i) If $f = \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} f_{\boldsymbol{l},\boldsymbol{s}}(\boldsymbol{\theta})\boldsymbol{z}^{\boldsymbol{l}}\boldsymbol{I}^{\boldsymbol{s}}$, we have $|f_{\boldsymbol{l},\boldsymbol{s}}|_\rho \leq |f|_{\rho,R}/R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1}$.

(ii) $|fg|_{\rho,R} \leq |f|_{\rho,R}|g|_{\rho,R}$.

(iii) For every $0 < \delta < \rho$ and $0 < \chi < 1$, we have for $j = 1,\ldots,r$ and $s = 1,\ldots,2m$:

$$\left|\frac{\partial f}{\partial\theta_j}\right|_{\rho-\delta,R} \leq \frac{|f|_{\rho,R}}{\delta\exp(1)}, \quad \left|\frac{\partial f}{\partial I_j}\right|_{\rho,R\chi} \leq \frac{|f|_{\rho,R}}{(1-\chi^2)R^2},$$
$$\left|\frac{\partial f}{\partial z_s}\right|_{\rho,R\chi} \leq \frac{|f|_{\rho,R}}{(1-\chi)R}.$$

As in lemma A.1, all the bounds work if f and g take values in $\mathbb{C}^{n_1}$ or $\mathbb{M}_{n_1,n_2}(\mathbb{C})$.

Proof. (i) As in the preceding lemma, from the definition of the norm in (A.1.5): $|f|_{\rho,R} = \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} |f_{\boldsymbol{l},\boldsymbol{s}}|_\rho R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1} \geq |f_{\boldsymbol{l},\boldsymbol{s}}|_\rho R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1}$, for all $(\boldsymbol{l},\boldsymbol{s}) \in \mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r$. To prove (ii), we take the norm of the product fg, so

$$
\begin{aligned}
|fg|_{\rho,R} &= \left| \sum_{\boldsymbol{l},\boldsymbol{s}} \sum_{\boldsymbol{l}',\boldsymbol{s}'} f_{\boldsymbol{l},\boldsymbol{s}}(\boldsymbol{\theta})\, g_{\boldsymbol{l}',\boldsymbol{s}'}(\boldsymbol{\theta}) \boldsymbol{I}^{\boldsymbol{s}+\boldsymbol{s}'} \boldsymbol{z}^{\boldsymbol{l}+\boldsymbol{l}'} \right|_{\rho,R} \\
&\leq \sum_{\boldsymbol{l},\boldsymbol{s}} \sum_{\boldsymbol{l}',\boldsymbol{s}'} |f_{\boldsymbol{l},\boldsymbol{s}}\, g_{\boldsymbol{l}',\boldsymbol{s}'}|_\rho R^{|\boldsymbol{l}+\boldsymbol{l}'|_1+2|\boldsymbol{s}+\boldsymbol{s}'|_1},
\end{aligned}
$$

where the summation indices $(\boldsymbol{l},\boldsymbol{s}),(\boldsymbol{l}',\boldsymbol{s}') \in \mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r$. Now, by the second statement of lemma A.1, $|f_{\boldsymbol{l},\boldsymbol{s}}\, g_{\boldsymbol{l}',\boldsymbol{s}'}|_\rho \leq |f_{\boldsymbol{l},\boldsymbol{s}}|_\rho |g_{\boldsymbol{l}',\boldsymbol{s}'}|_\rho$ and then,

$$
\begin{aligned}
|fg|_{\rho,R} &\leq \left(\sum_{\boldsymbol{l},\boldsymbol{s}} |f_{\boldsymbol{l},\boldsymbol{s}}|_\rho R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1} \right) \left(\sum_{\boldsymbol{l}',\boldsymbol{s}'} |g_{\boldsymbol{l}',\boldsymbol{s}'}|_\rho R^{|\boldsymbol{l}'|_1+2|\boldsymbol{s}'|_1} \right) \\
&= |f|_{\rho,R} |g|_{\rho,R},
\end{aligned}
$$

which is the desired result. The first inequality in (iii) is proved using item (iii) of lemma A.1,

$$
\begin{aligned}
\left| \frac{\partial f}{\partial \theta_j} \right|_{\rho-\delta,R} &= \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} \left| \frac{\partial f_{\boldsymbol{l},\boldsymbol{s}}}{\partial \theta_j} \right| R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1} \\
&\leq \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} \frac{|f_{\boldsymbol{l},\boldsymbol{s}}|_\rho}{\delta \exp(1)} R^{|\boldsymbol{l}|_1+2|\boldsymbol{s}|_1} \\
&= \frac{|f|_{\rho,R}}{\delta \exp(1)}.
\end{aligned}
$$

To derive the other two, we consider the function $\widehat{f}(\boldsymbol{x},I,\boldsymbol{y}) = \sum_{(\boldsymbol{l},\boldsymbol{s})\in\mathbb{Z}_+^{2m}\times\mathbb{Z}_+^r} |f_{\boldsymbol{l},\boldsymbol{s}}|_\rho \boldsymbol{z}^{\boldsymbol{l}} \boldsymbol{I}^{\boldsymbol{s}}$, defined for $|x_j|,|y_j| \leq R$, $j=1,\dots,m$; $|I_\nu| \leq R^2$, $\nu = 1,\dots,r$. Then, applying Cauchy estimates to $\widehat{f}$, we have, for instance:

$$
\begin{aligned}
\left| \frac{\partial f}{\partial z_k} \right|_{\rho,R\chi} &= \left| \frac{\partial \widehat{f}}{\partial z_k} \right|_{R\chi} = \sup_{\substack{|x_j|,|y_j|\leq R\chi,\, j=1,\dots,m \\ |I_\nu|\leq R^2\chi^2,\, \nu=1,\dots,r}} \left| \frac{\partial \widehat{f}}{\partial z_k} \right| \\
&\leq (1-\chi)^{-1} R^{-1} \left(\sup_{\substack{|x_j|,|y_j|\leq R,\, j=1,\dots,m \\ |I_\nu|\leq R^2,\, \nu=1,\dots,r}} |\widehat{f}| \right) = \frac{|\widehat{f}|_R}{(1-\chi)R} = \frac{|f|_{\rho,R}}{(1-\chi)R},
\end{aligned}
$$

with $k = 1,\dots,2m$, and one proceeds similarly for $\left| \frac{\partial f}{\partial I_\sigma} \right|_{\rho,R\chi}$, $\sigma = 1,\dots,r$. □

Lemma A.3. *Let us consider $f(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})$ and $g(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})$ complex valued functions, such that f and* grad g *are analytic functions defined on $\mathcal{D}_{r,m}(\rho,R)$ and 2π-periodic on $\boldsymbol{\theta}$.*

Then, for every $0 < \delta < \rho$ *and* $0 < \chi < 1$*, we have:*

$$
\begin{aligned}
|\{f,g\}|_{\rho-\delta,R\chi} \leq \frac{r|f|_{\rho,R}}{\delta\exp(1)}\left|\frac{\partial g}{\partial \boldsymbol{I}}\right|_{\rho-\delta,R\chi} &+ \frac{r|f|_{\rho,R}}{R^2(1-\chi^2)}\left|\frac{\partial g}{\partial \boldsymbol{\theta}}\right|_{\rho-\delta,R\chi} \\
&+ \frac{2m|f|_{\rho,R}}{R(1-\chi)}\left|\frac{\partial g}{\partial \boldsymbol{z}}\right|_{\rho-\delta,R\chi}.
\end{aligned} \tag{A.2.2}
$$

Proof. It follows directly from lemma A.2. □

Lemma A.4. *Let us take* $0 < \rho_0 < \rho$ *and* $0 < R_0 < R$*, and consider the analytic functions* $\boldsymbol{\Theta}$*,* $\boldsymbol{\mathcal{I}}$ *with values in* $\mathbb{C}^r$*,* $\mathbb{C}^{r'}$ *respectively, and* $\boldsymbol{\mathcal{X}}$*,* $\boldsymbol{\mathcal{Y}}$ *with values in* $\mathbb{C}^{m'}$*, all defined for* $(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y}) \in \mathcal{D}_{r,m}(\rho_0,R_0)$*, and* 2π*-periodic on* $\boldsymbol{\theta}$*. We assume that* $|\boldsymbol{\Theta}|_{\rho_0,R_0} \leq \rho-\rho_0$*,* $|\boldsymbol{\mathcal{I}}|_{\rho_0,R_0} \leq R^2$*, and that* $\max\{|\boldsymbol{\mathcal{X}}|_{\rho_0,R_0}, |\boldsymbol{\mathcal{Y}}|_{\rho_0,R_0}\} \leq R$*. Let* $f(\boldsymbol{\phi},\boldsymbol{u},\boldsymbol{J},\boldsymbol{v})$ *be a given (*2π*-periodic on* $\boldsymbol{\phi}$*) analytic function, defined on* $\mathcal{D}_{r',m'}(\rho,R)$*. If we introduce:*

$$
F(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y}) = f(\boldsymbol{\theta}+\boldsymbol{\Theta},\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{I}},\boldsymbol{\mathcal{Y}})
$$

then, $|F|_{\rho_0,R_0} \leq |f|_{\rho,R}$*.*

Proof. We may expand $f(\boldsymbol{\theta}+\boldsymbol{\Theta},\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{I}},\boldsymbol{\mathcal{Y}})$ in Taylor series,

$$
F(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y}) = \sum_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}} f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{\theta}+\boldsymbol{\Theta})\,\boldsymbol{\mathcal{I}}^{\boldsymbol{s}}\boldsymbol{\mathcal{X}}^{\boldsymbol{m}}\boldsymbol{\mathcal{Y}}^{\boldsymbol{n}},
$$

where the indices $\boldsymbol{s} \in \mathbb{Z}_+^{r'}$, $\boldsymbol{m},\boldsymbol{n} \in \mathbb{Z}_+^{m'}$ (and the dependence of $\boldsymbol{\Theta}$, $\boldsymbol{\mathcal{I}}$, $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{Y}}$ on $\boldsymbol{\theta}$, $\boldsymbol{x}$, $\boldsymbol{I}$ and $\boldsymbol{y}$ has not been written explicitly). Similarly, the coefficients $f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{\theta}+\boldsymbol{\Theta})$ can be expanded in Fourier series,

$$
f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{\theta}+\boldsymbol{\Theta}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^r} f_{\boldsymbol{k},\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\theta}+\boldsymbol{\Theta}\rangle),
$$

so their norm $|\cdot|_{\rho_0,R_0}$ is bounded according to,

$$
|f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{\theta}+\boldsymbol{\Theta})|_{\rho_0,R_0} \leq \sum_{\boldsymbol{k}\in\mathbb{Z}^r} |f_{\boldsymbol{k},\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}||\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\theta}+\boldsymbol{\Theta}\rangle)|_{\rho_0,R_0} \tag{A.2.3}
$$

and, by item (ii) of lemma A.2,

$$
|\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\theta}+\boldsymbol{\Theta}\rangle)|_{\rho_0,R_0} \leq \exp\left(|\boldsymbol{k}|_1\rho_0\right)|\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\Theta}\rangle)|_{\rho_0,R_0}.
$$

Next, expanding $\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\Theta}\rangle)$, applying the properties of the $|\cdot|_{\rho_0,R_0}$ norm and taking into account that $|\boldsymbol{\Theta}|_{\rho_0,R_0} \leq \rho-\rho_0$, one has $|\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\Theta}\rangle)|_{\rho_0,R_0} \leq \exp(|\boldsymbol{k}|_1(\rho-\rho_0))$. Hence $|\exp(\mathrm{i}\langle\boldsymbol{k},\boldsymbol{\theta}+\boldsymbol{\Theta}\rangle)|_{\rho_0,R_0} \leq \exp\left(|\boldsymbol{k}|_1\rho\right)$ which, in view of (A.2.3) leads to the inequality $|f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{\theta}+\boldsymbol{\Theta})|_{\rho_0,R_0} \leq |f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}|_\rho$. Finally, from the assumptions of the lemma on $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{I}}$:

$$
\begin{aligned}
|F|_{\rho_0,R_0} &\leq \sum_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}} |f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}(\boldsymbol{\theta}+\boldsymbol{\Theta})|_{\rho_0,R_0}|\boldsymbol{\mathcal{I}}|_{\rho_0,R_0}^{|\boldsymbol{s}|_1}|\boldsymbol{\mathcal{X}}|_{\rho_0,R_0}^{|\boldsymbol{m}|_1}|\boldsymbol{\mathcal{Y}}|_{\rho_0,R_0}^{|\boldsymbol{n}|_1} \\
&\leq \sum_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}} |f_{\boldsymbol{s},\boldsymbol{m},\boldsymbol{n}}|_\rho R^{2|\boldsymbol{s}|_1+|\boldsymbol{m}|_1+|\boldsymbol{n}|_1} \\
&= |f|_{\rho,R},
\end{aligned}
$$

which is the desired result. □

Lemma A.5. *Let us consider* $\boldsymbol{\Theta}^{(j)}$, $\boldsymbol{\mathcal{I}}^{(j)}$, $\boldsymbol{\mathcal{X}}^{(j)}$ *and* $\boldsymbol{\mathcal{Y}}^{(j)}$, $j = 1, 2$, *in the same conditions of the ones of lemma A.4, but with the following bounds:* $|\boldsymbol{\Theta}^{(j)}|_{\rho_0, R_0} \le \rho - \rho_0 - \delta$, $|\boldsymbol{\mathcal{I}}^{(j)}|_{\rho_0, R_0} \le R^2 - \sigma$, *and* $\max\{|\boldsymbol{\mathcal{X}}^{(j)}|_{\rho_0,R_0}, |\boldsymbol{\mathcal{Y}}^{(j)}|_{\rho_0,R_0}\} \le R - \chi$, *with* $0 < \delta < \rho - \rho_0$, $0 < \sigma < R^2$ *and* $0 < \chi < R$. *Then, if one takes the function* f *of lemma A.4 to define*

$$F^{(j)}(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) = f(\boldsymbol{\theta} + \boldsymbol{\Theta}^{(j)}, \boldsymbol{\mathcal{X}}^{(j)}, \boldsymbol{\mathcal{I}}^{(j)}, \boldsymbol{\mathcal{Y}}^{(j)}), \qquad \text{with } j = 1, 2$$

one has $|F^{(1)} - F^{(2)}|_{\rho_0, R_0} \le \mathcal{K}|f|_{\rho, R}$, *where if we put* $\boldsymbol{\mathcal{Z}}^* = (\boldsymbol{\mathcal{X}}^*, \boldsymbol{\mathcal{Y}}^*)$ *then*

$$\mathcal{K} \equiv \frac{|\boldsymbol{\Theta}^{(1)} - \boldsymbol{\Theta}^{(2)}|_{\rho_0, R_0}}{\delta \exp(1)} + r' \frac{|\boldsymbol{\mathcal{I}}^{(1)} - \boldsymbol{\mathcal{I}}^{(2)}|_{\rho_0, R_0}}{\sigma} + \frac{1}{\chi} \sum_{j=1}^{2m'} |\boldsymbol{\mathcal{Z}}_j^{(1)} - \boldsymbol{\mathcal{Z}}_j^{(2)}|_{\rho_0, R_0}.$$

Proof. It is based upon the same ideas used to prove lemma A.4. □

Lemma A.6. *Let* f *and* g *be complex valued functions,* f *analytic in the domain* $\mathcal{D}_{r,m}(\rho + \delta, R\exp(\delta))$ *and* $\operatorname{grad} g$ *analytic in* $\mathcal{D}_{r,m}(\rho + \delta', R\exp(\delta'))$, *with* $0 < \delta, \delta' < 1/2$. *Then*

$$|\{f, g\}|_{\rho, R} \le A(R, \delta, \delta', r, m)\, |f|_{\rho+\delta, R\mathrm{e}^{\delta}}\, |g|_{\rho+\delta', R\mathrm{e}^{\delta'}} \tag{A.2.4}$$

with,

$$A(R, \delta, \delta', r, m) := \frac{r + 8m}{(R\exp(\delta))(R\exp(\delta'))\delta\delta'}$$

Proof. Using lemma A.3 we obtain,

$$|\{f, g\}|_{\rho, R} \le \frac{r|f|_{\rho+\delta, R\mathrm{e}^{\delta}}}{\delta \exp(1)} \left|\frac{\partial g}{\partial \boldsymbol{I}}\right|_{\rho, R} + \frac{r|f|_{\rho+\delta, R\mathrm{e}^{\delta}}}{R^2(\exp(2\delta) - 1)} \left|\frac{\partial g}{\partial \boldsymbol{\theta}}\right|_{\rho, R} + \frac{2m|f|_{\rho+\delta, R\mathrm{e}^{\delta}}}{R(\exp(\delta) - 1)} \left|\frac{\partial g}{\partial \boldsymbol{z}}\right|_{\rho, R},$$

and the norms of the derivatives may be estimated with (iii) of lemma A.2,

$$\left|\frac{\partial g}{\partial \boldsymbol{I}}\right|_{\rho, R} \le \frac{1}{R^2(\exp(2\delta') - 1)} |g|_{\rho+\delta', R\mathrm{e}^{\delta'}}, \qquad \left|\frac{\partial g}{\partial \boldsymbol{\theta}}\right|_{\rho, R} \le \frac{1}{\delta' \exp(1)} |g|_{\rho+\delta', R\mathrm{e}^{\delta'}},$$
$$\left|\frac{\partial g}{\partial \boldsymbol{z}}\right|_{\rho, R} \le \frac{1}{R(\exp(\delta') - 1)} |g|_{\rho+\delta', R\mathrm{e}^{\delta'}},$$

This allows us to bound the norm of the Poisson parenthesis of f and g according to,

$$\begin{aligned}
|\{f, g\}|_{\rho, R} &\le \left[\frac{r/\mathrm{e}}{\delta(\mathrm{e}^{2\delta'} - 1)} + \frac{r/\mathrm{e}}{\delta'(\mathrm{e}^{2\delta} - 1)} + \frac{2m}{(\mathrm{e}^{\delta} - 1)(\mathrm{e}^{\delta'} - 1)}\right] \frac{|f|_{\rho+\delta, R\mathrm{e}^{\delta}} |g|_{\rho+\delta', R\mathrm{e}^{\delta'}}}{R^2} \\
&\le \left[\frac{r/\mathrm{e}}{\delta\delta'\mathrm{e}^{2\delta'}} + \frac{r/\mathrm{e}}{\delta\delta'\mathrm{e}^{2\delta}} + \frac{8m}{\delta\delta'\mathrm{e}^{\delta}\mathrm{e}^{\delta'}}\right] \frac{|f|_{\rho+\delta, R\mathrm{e}^{\delta}} |g|_{\rho+\delta', R\mathrm{e}^{\delta'}}}{R^2} \qquad \text{(A.2.5)} \\
&= \frac{1}{(R\mathrm{e}^{\delta})(R\mathrm{e}^{\delta'})\delta\delta'} \left[8m + \frac{2r}{\mathrm{e}} \cosh(\delta - \delta')\right] |f|_{\rho+\delta, R\mathrm{e}^{\delta}} |g|_{\rho+\delta', R\mathrm{e}^{\delta'}}
\end{aligned}$$

where, to state the second inequality we have used that, whenever $0 < x \le 1$,

$$\frac{x}{2} \le 1 - \exp(-x). \tag{A.2.6}$$

So, for example, $\mathrm{e}^{2\delta'} - 1 = \mathrm{e}^{2\delta'}(1 - \mathrm{e}^{-2\delta'}) \geq \mathrm{e}^{2\delta'}\delta'$ ($2\delta' \leq 1$, since $\delta' \leq \frac{1}{2}$), and similarly for the other factors of the same kind appearing at the denominators[(2)]. Moreover, the argument of the hyperbolic cosine appearing inside the square brackets in (A.2.5), $\delta - \delta'$, must be $-\frac{1}{2} < \delta - \delta' < \frac{1}{2}$ (since $0 < \delta, \delta' \leq \frac{1}{2}$), so $\cosh(\delta - \delta') \leq \frac{\mathrm{e}}{2}$. Hence, $8m + \frac{2r}{\mathrm{e}} \cosh(\delta - \delta') \leq r + 8m$. But $r + 8m$ is just the numerator of $A(r, \delta, \delta', r, m)$, so the last term in the expression (A.2.5) –the stuff on the right of the equal sign–, cannot be greater than the quotient $\frac{r+8m}{(Re^{\delta})(Re^{\delta'})\delta\delta'}$. This shows the estimate (A.2.4) is fulfilled and concludes the proof of the lemma. □

The following corollary is gleaned directly from this last lemma, and we use it in chapter 2 to bound the Poisson brackets $\{f_{l,s-j}, G_{2+j}\}$ appearing in Giorgilli-Galgani algorithm (see remark 2.15).

Corollary A.7. *Consider $\rho, R > 0$ and a finite sequence of positive numbers, $\{\delta_\nu\}_{3 \leq \nu \leq s+2}$, such that their sum satisfies the condition*

$$\sum_{\nu=1}^{s} \delta_{\nu+2} \leq \min\{1/4, \rho/4\}. \tag{A.2.7}$$

Let ρ_ν, R_ν for $\nu = 2, \dots, s+2$, be the quantities defined by,

$$\rho_2 := \rho, \quad \rho_\nu := \rho_2 - 2\sum_{\sigma=3}^{\nu} \delta_\sigma, \tag{A.2.8}$$

$$R_2 := R, \quad R_\nu := R_2 \exp\left(-2\sum_{\sigma=3}^{\nu} \delta_\sigma\right). \tag{A.2.9}$$

Therefore, if f, g are two complex valued functions, such that f and g are analytic in $\mathcal{D}_{r,m}(\rho_{s-j+2}, R_{s-j+2})$ and $\mathcal{D}_{r,m}(\rho_{2+j} + \delta_{2+j}, R_{2+j}\exp\delta_{2+j})$ respectively, it turns out that

$$|\{f,g\}|_{\rho_{s+2},R_{s+2}} \leq \frac{(r+8m)\mathrm{e}}{2R^2\delta_{s+2}\delta_{j+2}} |f|_{\rho_{s-j+2},R_{s-j+2}} |g|_{\rho_{2+j}+\delta_{2+j},R_{2+j}\exp(\delta_{2+j})} \tag{A.2.10}$$

for all $j = 1, \dots, s$.

Proof. Given a fixed j, with $1 \leq j \leq s$, take

$$\delta = 2\sum_{\sigma=s-j+3}^{s+2} \delta_\sigma,$$

and,

$$\delta' = \begin{cases} \delta_{j+2} + 2\sum_{\sigma=j+3}^{s+2} \delta_\sigma, & \text{if } 1 \leq j \leq s-1 \ , \\ \delta_{s+2}, & \text{if } j = s \ , \end{cases}$$

(note that, in particular $\delta, \delta' \leq 1/2$). With this choice of δ and δ' it is true that $\rho_{s+2} + \delta = \rho_{s-j+2}$, $\rho_{s+2} + \delta' = \rho_{2+j} + \delta_{2+j}$ (and also $R_{s+2}\exp(\delta) = R_{s-j+2}$, $R_{s+2}\exp(\delta') = R_{2+j}\exp(\delta_{2+j})$). Thus, application of lemma A.6 leads to the inequality (A.2.10) if, in addition, we use that,

[(2)]In the referred expressions –and to make them shorter–, $\mathrm{e} = \exp(1)$ and $\mathrm{e}^x = \exp(x)$, with letter "e" in roman font, not *italicized*, as usual when displaying mathematical formulas. This convention is kept throughout and applied when becomes convenient.

(A.7-i) $\exp(\delta), \exp(\delta') > 1$.

(A.7-ii) $\delta \geq 2\delta_{s+2}$ and $\delta' \geq \delta_{j+2}$, for all $1 \leq j \leq s$.

(A.7-iii) $R_{s+2} \geq R\exp(-1/2)$, as follows directly from $\sum_{j=1}^{s} \delta_{j+2} \leq 1/4$ and (A.2.9) for $\nu = s+2$.

And we are done with the proof. $\square$

Lemma A.8. *Let $S(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})$ be a function defined on $\mathcal{D}_{r,m}(\rho, R)$ with $\rho > 0$ and $R > 0$, being* grad S *analytic on $\mathcal{D}_{r,m}(\rho, R)$ and 2π-periodic on $\boldsymbol{\theta}$. If we assume that,*

$$\left|\frac{\partial S}{\partial \boldsymbol{\theta}}\right|_{\rho,R} \leq R^2(1-\chi^2), \qquad \left|\frac{\partial S}{\partial \boldsymbol{I}}\right|_{\rho,R} \leq \delta, \quad \left|\frac{\partial S}{\partial \boldsymbol{z}}\right|_{\rho,R} \leq R(1-\chi), \tag{A.2.11}$$

for certain $0 < \chi < 1$ and $0 < \delta < \rho$, then one has

(i) $\phi_t^S : \mathcal{D}_{r,m}(\rho-\delta, R\chi) \longrightarrow \mathcal{D}_{r,m}(\rho, R)$, for every $-1 \leq t \leq 1$, where ϕ_t^S is the flow time t of the Hamiltonian system given by S.

(ii) If one writes $\phi_t^S - \mathrm{Id} = (\boldsymbol{\Theta}_t^{S*}, \boldsymbol{\mathcal{X}}_t^{S*}, \boldsymbol{\mathcal{I}}_t^{S*}, \boldsymbol{\mathcal{Y}}_t^{S*})$, then, for every $-1 \leq t \leq 1$, we have that $\boldsymbol{\Theta}_t^S$, $\boldsymbol{\mathcal{I}}_t^S$ and $\boldsymbol{\mathcal{Z}}_t^{S*} = (\boldsymbol{\mathcal{X}}_t^{S*}, \boldsymbol{\mathcal{Y}}_t^{S*})$ are analytic functions on $\mathcal{D}_{r,m}(\rho-\delta, R\chi)$, 2π-periodic on $\boldsymbol{\theta}$. Moreover, the following bounds hold:

$$|\boldsymbol{\Theta}_t^S|_{\rho-\delta,R\chi} \leq \left|\frac{\partial S}{\partial \boldsymbol{\theta}}\right|_{\rho,R}, \qquad |\boldsymbol{\mathcal{I}}_t^S|_{\rho-\delta,R\chi} \leq \left|\frac{\partial S}{\partial \boldsymbol{\mathcal{I}}}\right|_{\rho,R}, \qquad |\boldsymbol{\mathcal{Z}}_t^S|_{\rho-\delta,R\chi} \leq \left|\frac{\partial S}{\partial \boldsymbol{\theta}}\right|_{\rho,R}. \tag{A.2.12}$$

Proof. Consider $(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) \in \mathcal{D}_{r,m}(\rho-\delta, R\chi)$. Letting $\psi_t^S - \mathrm{Id} = (\boldsymbol{\Theta}_t^{S*}, \boldsymbol{\mathcal{X}}_t^{S*}, \boldsymbol{\mathcal{I}}_{\boldsymbol{t}}^{\boldsymbol{S}}, \boldsymbol{\mathcal{Y}}_t^{S*})$, we define,

$$\begin{aligned}\tau_{(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})} = \sup\{|t| : |\boldsymbol{\Theta}_t^S|_{\rho-\delta,R\chi} \leq \delta,& \\ |\boldsymbol{\mathcal{I}}_t^S|_{\rho-\delta,R\chi} \leq R^2(1-\chi^2), &\quad |\boldsymbol{\mathcal{Z}}_t^S|_{\rho-\delta,R\chi} \leq R(1-\chi)\}.\end{aligned}$$

Clearly, $\tau_{(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})}$ and since ϕ_t^S is the flow time t of the Hamiltonian S,

$$|\boldsymbol{\Theta}_t^S|_{\rho-\delta,R\chi} \leq \left|\int_0^t \frac{\partial S}{\partial \boldsymbol{I}} \circ \phi_s^S \, ds\right|_{\rho-\delta,R\chi} \leq |t| \cdot \left|\frac{\partial S}{\partial \boldsymbol{I}}\right|_{\rho,R} \leq \delta\,|t|, \tag{A.2.13a}$$

$$|\boldsymbol{\mathcal{I}}_t^S|_{\rho-\delta,R\chi} \leq \left|\int_0^t \frac{\partial S}{\partial \boldsymbol{\theta}} \circ \phi_s^S \, ds\right|_{\rho-\delta,R\chi} \leq |t| \cdot \left|\frac{\partial S}{\partial \boldsymbol{\theta}}\right|_{\rho,R} \leq R^2(1-\chi^2)\,|t|, \tag{A.2.13b}$$

$$|\boldsymbol{\mathcal{Z}}_t^S|_{\rho-\delta,R\chi} \leq \left|\int_0^t \frac{\partial S}{\partial \boldsymbol{z}} \circ \phi_s^S \, ds\right|_{\rho-\delta,R\chi} \leq |t| \cdot \left|\frac{\partial S}{\partial \boldsymbol{z}}\right|_{\rho,R} \leq R(1-\chi)\,|t|, \tag{A.2.13c}$$

where we have used lemma A.5 and the assumed inequalities (A.2.11). Then, one has, for $-1 \leq t \leq 1$, $|\boldsymbol{\Theta}_t^S|_{\rho-\delta,R\chi} \leq \delta$, $|\boldsymbol{\mathcal{I}}_t^S|_{\rho-\delta,R\chi} \leq R^2(1-\chi^2)$, $|\boldsymbol{\mathcal{Z}}_t^S|_{\rho-\delta,R\chi} \leq R(1-\chi)$ so $\tau_{(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})} \geq 1$ and hence the result of the first item follows. The inequalities (A.2.12) are derived straightforward from (A.2.13a), (A.2.13b) and (A.2.13c), whereas the analyticity and $2Pi$-periodicity in $\boldsymbol{\theta}$, for $-1 \leq t \leq 1$, of $\boldsymbol{\Theta}_t^S$, $\boldsymbol{\mathcal{I}}_t^S$, $\boldsymbol{\mathcal{Z}}_t^S$ follow from the analyticity and 2π periodicity in $\boldsymbol{\theta}$ of grad S. $\square$

Lemma A.9. *Consider $\xi \in \mathbb{C}$ such that $|\xi| \geq L > 0$ and let $f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})$ be a function defined in $\mathcal{D}_{r,m}(\rho, R)$ such that:*

$$|f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})|_{\rho,R} \leq M, \qquad \text{with } 0 < M < L.$$

Let $g(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})$, $h(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})$ be given respectively by

$$g(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) = \sqrt{\xi + f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})}, \qquad h(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) = \frac{1}{\sqrt{\xi + f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})}}.$$

Then, one has:

$$|g(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})|_{\rho,R} \leq |\sqrt{\xi}| \left(2 - \sqrt{2 - \frac{M}{L}}\right), \qquad |h(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})|_{\rho,R} \leq \frac{1/|\sqrt{\xi}|}{\sqrt{1 - \frac{M}{L}}}.$$

Proof. To prove the first inequality, we use that $|\frac{f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})}{\xi}| \leq \frac{M}{L} < 1$. Therefore, we can develop the square root using binomial expansion:

$$g(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) = \sqrt{\xi} \sum_{j \geq 0} \binom{1/2}{j} \left(\frac{f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})}{\xi}\right)^j,$$

and taking the norm $|\cdot|_{\rho,R}$ at both sides:

$$\begin{aligned}
|g(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})|_{\rho,R} &\leq |\sqrt{\xi}| \sum_{j \geq 0} \left|\binom{1/2}{j}\right| \frac{|g(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})|_{\rho,R}^j}{|\xi|^j} \\
&\leq |\sqrt{\xi}| \left(\binom{1/2}{0} + \sum_{j \geq 1} \binom{1/2}{j} (-1)^{j+1} \left(\frac{M}{L}\right)^j\right) \\
&= |\sqrt{\xi}| \left(1 + \binom{1/2}{0} - 1 - \sum_{j \geq 1} \binom{1/2}{j} (-1)^j \left(\frac{M}{L}\right)^j\right) \\
&= |\sqrt{\xi}| \left(2 - \sum_{j \geq 0} \binom{1/2}{j} (-1)^j \left(\frac{M}{L}\right)^j\right) \\
&= |\sqrt{\xi}| \left(2 - \sqrt{1 - \frac{M}{L}}\right),
\end{aligned}$$

where it has been used that $\left|\binom{1/2}{j}\right| = (-1)^{j+1} \binom{1/2}{j}$, for $j \in \mathbb{N}$ (as one may check easily) and that $\binom{1/2}{0} \equiv 1$, by definition. The second inequality of the lemma follows similarly using the corresponding binomial expansion. Explicitly,

$$h(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y}) = \frac{1}{\sqrt{\xi}} \sum_{j \geq 0} \binom{-1/2}{j} \left(\frac{f(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{I}, \boldsymbol{y})}{\xi}\right)^j,$$

and, as before, taking norms $|\cdot|_{\rho,R}$ at both sides one obtains:

$$
\begin{aligned}
|h(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})|_{\rho,R} &\leq \frac{1}{|\sqrt{\xi}|}\sum_{j\geq 0}\left|\binom{-1/2}{j}\right|\frac{|f(\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{I},\boldsymbol{y})|^j_{\rho,R}}{|\xi|^j} \\
&\leq \frac{1}{|\sqrt{\xi}|}\sum_{j\geq 0}\binom{-1/2}{j}(-1)^j\left(\frac{M}{L}\right)^j \\
&= \frac{1/|\sqrt{\xi}|}{\sqrt{1-\frac{M}{L}}},
\end{aligned}
$$

since it is straightforward to check that $\left|\binom{-1/2}{j}\right| = (-1)^j\binom{-1/2}{j}$, for all $j \in \mathbb{Z}_+$. Thus the lemma is proved. □

Remark A.10 *(Notation).* To avoid any possible confusion with the square brackets, $[\cdot]$, in what follows, $\lfloor\cdot\rfloor$ will denote the integral part or the greatest integer function. This symbol was already settled at the introduction. ♠

The next results become useful when we look for estimates of those sums $f_{l,s} = \sum_{j=1}^{s}\frac{j}{s}L_{G_{2+j}}f_{l,s-j}$ in the transformation algorithm (chapter 2).

Lemma A.11. *Consider a non decreasing sequence of positive numbers $\{\beta_s\}_{2\leq s\leq r}$, with $\beta_2 = 1$. Given two positive integers, s and $r > 2$, define the product*

$$
\mathcal{W}(r,s) = \left(\prod_{\nu=1}^{r-2}\beta_{\nu+2}\right)^{\left\lfloor\frac{s}{r-2}\right\rfloor}\left(\prod_{\nu=0}^{[s,r-2]}\beta_{\nu+2}\right), \tag{A.2.14}
$$

(definition 2.12 of chapter 2) where $[m,n]$ denote the remainder of the integer division of m by n (definition 2.5.9 in the same chapter). It turns out that the products above satisfy

$$
\mathcal{W}(r,s-j)\,\beta_3\beta_4\cdots\beta_{j+2} \leq \mathcal{W}(r,s), \tag{A.2.15}
$$

for all the integers $s\geq 1$, $r>2$ and $j = 1,2,\ldots,\min(s,r-2)$.

Proof. Both possible cases are thought of separately. For a given and fixed $r\geq 3$, we shall find out if the inequality (A.2.14) of the lemma is satisfied whether $s > r-2$ or $s\leq r-2$.

Case 1. $s > r-2$; let q and p be the quotient and the remainder of the integer division of s by $r-2$, i. e., $q = \left\lfloor\frac{s}{r-2}\right\rfloor$ and $p = [s,r-2]$. Hence, $p < r-2$; $q\geq 1$; $j = 1,\ldots,r-2$ and according to the formula (A.2.14)

$$
\mathcal{W}(r,s) = (\beta_3\cdots\beta_r)^q\beta_3\cdots\beta_{p+2}, \tag{A.2.16}
$$

so depending on the values of j, three possibilities (sub-cases) arise.

(i) $j < p$. We may write $s-j = q(r-2)+p-j$, so $\left\lfloor\frac{s-j}{r-2}\right\rfloor = q$ and $[s-j,r-2] = p-j$. Therefore,

$$
\begin{aligned}
\mathcal{W}(r,s-j)\beta_3\cdots\beta_{j+2} &= (\beta_3\cdots\beta_r)^q\beta_3\cdots\beta_{p-j+2}\beta_3\cdots\beta_{j+2} \\
&\leq (\beta_3\cdots\beta_r)^q\beta_3\cdots\beta_{p-j+2}\beta_{p-j+3}\cdots\beta_{p+2},
\end{aligned}
$$

by the monotonic character of the quantities $\{\beta_\nu\}_{1\leq\nu\leq r}$, it is $\beta_3\leq\beta_{p-j+3}, \beta_4\leq\beta_{p-j+4},\ldots,\beta_{j+2}\leq\beta_{p+2}$.

(ii) $j = p$. Then, $s - j = q(r-2)$ and now $\left\lfloor \frac{s-j}{r-2} \right\rfloor = q$, but $[s-j, r-2] = 0$. Hence,

$$\mathcal{W}(r, s-j)\beta_3 \cdots \beta_{j+2} = (\beta_3 \cdots \beta_r)^q \beta_3 \cdots \beta_{p+2},$$

so (A.2.15) becomes an equality when $j = [r, s-r]$.

(iii) $j > p$. We put, $s - j = (q-1)(r-2) + p + r - j - 2$. Consequently, it must be $\left\lfloor \frac{s-j}{r-2} \right\rfloor = q - 1$ and $[s-j, r-2] = p + r - j - 2$. Note that $p + r - j \geq p + 2$ (because $j \leq r-2$). Assume first, however, that this last inequality is strict. Hence, by (A.2.14),

$$\begin{aligned}\mathcal{W}(r, s-j)\beta_3 \cdots \beta_{j+2} &= (\beta_3 \cdots \beta_r)^{q-1} \beta_3 \cdots \beta_{p+2}\beta_{p+3} \cdots \beta_{p+r-j}\beta_3 \cdots \beta_{j+2} \\ &\leq (\beta_3 \cdots \beta_r)^{q-1} \beta_3 \cdots \beta_{p+2}\beta_3 \cdots \beta_{j+2}\beta_{j+3} \cdots \beta_r \\ &= (\beta_3 \cdots \beta_r)^q \beta_3 \cdots \beta_{p+2},\end{aligned}$$

where we make use again of the non-decreasing character of the sequence. In particular –as $p < j$–, apply $\beta_{p+3} \leq \beta_{j+3}, \beta_4 \leq \beta_{p+4}, \ldots, \beta_{p+r-j} \leq \beta_r$. The discussion of this sub-case is not complete, though, without seeing what happens when $p+r-j = p+2$; then $j = r-2$ and,

$$\begin{aligned}\mathcal{W}(r, s-j)\beta_3 \cdots \beta_{j+2} &= (\beta_3 \cdots \beta_r)^{q-1} \beta_3 \cdots \beta_{p+2}\beta_3 \cdots \beta_r \\ &= (\beta_3 \cdots \beta_r)^q \beta_3 \cdots \beta_{p+2},\end{aligned}$$

so, as in the previous item (A.2.14) is an equality also for the maximum allowed value of j, (i. e., $j = r-2$).

Therefore, all the sub-cases of this first case verify the inequality (A.2.15), which shows the validity of the lemma when $s > r-2$.

Case 2. If $s \leq r-2$, it is readily checked out from its definition through (A.2.14) that,

$$\mathcal{W}(r, m) = \beta_2\beta_3 \cdots \beta_{m+2}$$

for all $0 \leq m \leq r-2$. Accordingly,

$$\begin{aligned}\mathcal{W}(r, s-j)\beta_3 \cdots \beta_{j+2} &= \beta_3 \cdots \beta_{s-j+2}\beta_3 \cdots \beta_{j+2} \\ &\leq \beta_3 \cdots \beta_{s-j+2}\beta_{s-j+3} \cdots \beta_{s+2},\end{aligned}$$

since $\beta_3 \leq \beta_{s-j+3}, \beta_4 \leq \beta_{s-j+4}, \ldots, \beta_{j+2} \leq \beta_{s+2}$.

Thus, we have exhausted by far all the possibilities, and seen that in everyone of the different cases and sub-cases, the r. h. s. of the $\leq$ symbol (=, for a couple of them) matches the r. h. s. of (A.2.14). This proves the assertion of the lemma. □

Lemma A.12. *Consider a, d, δ real positive numbers; s, r two integers with $s \geq 1$, $r \geq 3$ and a sequence $\{\delta_\nu\}_{3 \leq \nu \leq s+2}$ given by,*

$$\delta_\nu = \begin{cases} \delta, & \text{if } 3 \leq \nu \leq r\ , \\ \dfrac{\delta}{s}, & \text{if } r < \nu \leq s+2\ . \end{cases} \tag{A.2.17}$$

Let $\{\vartheta_\nu\}_{0\leq\nu\leq s}$ be a sequence defined recursively through,

$$\begin{aligned} &\vartheta_0 = 1, \\ &\vartheta_\nu = \frac{\delta_3}{\delta_{\nu+2}} \sum_{j=1}^{\min\{\nu,r-2\}} \tfrac{j}{\nu}\, a^{j-1} d\, \vartheta_{\nu-j}, \qquad 1 \leq \nu \leq s. \end{aligned} \tag{A.2.18}$$

Then,

$$\vartheta_s \leq d(de + 2ae)^{s-1}. \tag{A.2.19}$$

Proof. All along this proof, we shall suppose $r \geq 4$. This does not mean any loss in generality, for it may be easily checked out that the assertion of the lemma works also for $r = 3$ (see remark at the end). By direct computation, we have:

$$\begin{aligned} \vartheta_1 &= d \\ \vartheta_2 &= \frac{\delta_3}{\delta_4}\frac{1}{2}d^2 + \frac{\delta_3}{\delta_4}ad = \frac{\delta_3}{\delta_4}\left(\frac{d}{2} + 2a\right) d \\ &= \left(\frac{d}{2} + 2a\right) d \leq (d + 2a)\vartheta_1. \end{aligned}$$

So the result (A.2.19) is satisfied for $s = 1, 2$. Suppose now that $3 \leq s \leq r - 2$, and take $3 \leq \nu \leq s$ (hence $\delta_\nu = \delta$). Then, from the definition of ϑ_ν,

$$\begin{aligned} \vartheta_\nu &= \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + \frac{\delta_3}{\delta_{\nu+2}}\sum_{j=2}^{\nu-1}\frac{j}{\nu}a^{j-1}d\vartheta_{\nu-j} + \frac{\delta_3}{\delta_{\nu+2}}a^{\nu-1}d \\ &= \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + \frac{\delta_{\nu+1}}{\delta_{\nu+2}}\left(\frac{\delta_3}{\delta_{\nu+1}}\sum_{i=1}^{\nu-2}\frac{i+1}{\nu}a^i d\vartheta_{\nu-1-i} + \frac{\delta_3}{\delta_{\nu+1}}a^{\nu-1}d\right), \end{aligned} \tag{A.2.20}$$

where the indices in the sum have been shifted taking $j = i+1$. Now taking into account that $\frac{i+1}{\nu} \leq \frac{2i}{\nu-1}$, for all the integers i, ν with $i \geq 1$, $\nu \geq 2$. Moreover, and with the aid of the naive inequality $a^{\nu-1}d \leq 2a^{\nu-1}d$, we should derive,

$$\begin{aligned} \vartheta_\nu &\leq \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + 2a\frac{\delta_{\nu+1}}{\delta_{\nu+2}}\left(\frac{\delta_3}{\delta_{\nu+1}}\sum_{i=1}^{\nu-2}\frac{i}{\nu-1}a^{i-1}d\vartheta_{\nu-1-i} + \frac{\delta_3}{\delta_{\nu+1}}a^{\nu-2}d\right) \\ &= \left(\frac{d\delta_3}{\nu\delta_{\nu+2}} + 2a\frac{\delta_{\nu+1}}{\delta_{\nu+2}}\right)\vartheta_{\nu-1} \\ &\leq \left(\frac{d}{\nu} + 2a\right)\vartheta_{\nu-1} \\ &\leq (d+2a)\vartheta_{\nu-1}, \end{aligned} \tag{A.2.21}$$

(it has been taken into account that $\delta_\nu = \delta_3$ for all $3 \leq \nu \leq r$ and $\nu \leq s \leq r-2$). So we have for every ν, $1 \leq \nu \leq s \leq r-2$,

$$\vartheta_\nu \leq (d + 2a)\vartheta_{\nu-1}, \tag{A.2.22}$$

and recursive backward application from $\nu = s$ down to $\nu = 2$ of this last relation, yields the estimate (A.2.19):

$$\vartheta_s \le (d+2a)\vartheta_{s-1} \le (d+2a)^2\vartheta_{s-2} \le \cdots \le (d+2a)^{s-1}\vartheta_1.$$

Let us now study the case $s > r-2$. Assume, first, that $s > r-1$, and for $r-1 \le \nu \le s$, we proceed in the same way than in (A.2.20) and (A.2.21), i. e.,

$$\begin{aligned}
\vartheta_\nu &= \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + \frac{\delta_3}{\delta_{\nu+2}}\sum_{j=2}^{r-2}\frac{j}{\nu}a^{j-1}d\vartheta_{\nu-j} \\
&= \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + \frac{\delta_3}{\delta_{\nu+2}}\sum_{i=1}^{r-3}\frac{i+1}{\nu}a^{i}d\vartheta_{\nu-1-i} \\
&\le \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + 2a\frac{\delta_{\nu+1}}{\delta_{\nu+2}}\left(\frac{\delta_3}{\delta_{\nu+1}}\sum_{i=1}^{r-2}\frac{i}{\nu-1}a^{i-1}d\vartheta_{\nu-1-i}\right) \\
&= \frac{d\delta_3}{\nu\delta_{\nu+2}}\vartheta_{\nu-1} + 2a\frac{\delta_{\nu+1}}{\delta_{\nu+2}}\vartheta_{\nu-1} \\
&= \frac{\delta_3}{\nu\delta_{\nu+2}}\left(d + 2a\nu\frac{\delta_{\nu+1}}{\delta_3}\right)\vartheta_{\nu-1},
\end{aligned}$$

where identical shift of the summation indices as in (A.2.20) has been made. Also, we have used again the inequality $\frac{i+1}{\nu} \le \frac{2i}{\nu-1}$, for all $i \ge 1$, $\nu \ge 2$ and the fact that the sum from 1 to $r-3$ should be less than the corresponding one from 1 to $r-2$ (since the latter contains one positive term more). Furthermore, if $\nu > r-1$, then by definition (A.2.17) it must be $\nu\frac{\delta_{\nu+1}}{\delta_3} \le 1$, so

$$\vartheta_\nu \le \frac{\delta_3}{\nu\delta_{\nu+2}}(d+2a)\vartheta_{\nu-1}, \tag{A.2.23}$$

whereas for $\nu = r-1$,

$$\begin{aligned}
\vartheta_{r-1} &\le \frac{1}{(r-1)\delta_{r+1}}\left(\delta_3 d + 2a(r-1)\delta_r\right)\vartheta_{r-2} \\
&= \frac{\delta_3}{\delta_{r+1}}\left(\frac{d}{r-1} + 2a\frac{\delta_r}{\delta_3}\right)\vartheta_{r-2} \le \frac{\delta_3}{\delta_{r+1}}(d+2a)\vartheta_{r-2}.
\end{aligned} \tag{A.2.24}$$

If recursion is now applied backwards from $\nu = s$ down to $\nu = r-1$,

$$\begin{aligned}
\vartheta_s \le \frac{\delta_3}{s\delta_{s+2}}(d+2a)\vartheta_{s-1} \le \frac{\delta_3^2}{s(s-1)\delta_{s+2}\delta_{s+1}}(d+2a)^2\vartheta_{s-2} \le \dots \\
\dots \le \frac{\delta_3^{s-r+1}}{s(s-1)\cdots r\delta_{s+2}\delta_{s+1}\cdots\delta_{r+2}}(d+2a)^{s-r+1}\vartheta_{r-1}.
\end{aligned}$$

Now using the bound (A.2.24) for ϑ_{r-1} and that $\vartheta_{r-2} \le d(d+2a)^{r-3}$, so we obtain

$$\vartheta_s \le \frac{(r-1)!\delta_3^{s-r+2}}{s!\delta_{r+1}\delta_{r+2}\cdots\delta_{s+2}}(d+2a)^{s-1}d;$$

but, as $\delta_{r+1} = \cdots = \delta_s = \delta_3/s$, the fraction of the r. h. s. in this expression can be arranged and bounded according to,

$$\begin{aligned}\frac{(r-1)!\delta_3^{s-r+2}}{s!\delta_{r+1}\cdots\delta_{s+2}} &= \frac{(r-1)!\delta_3^{s-r+2}s^{s-r+2}}{s!\delta_3^{s-r+2}} \times \frac{s^{r-2}}{s^{r-2}} \\ &= \frac{(r-1)!}{s^{r-2}} \times \frac{s^s}{s!} \underset{(s>r-1)}{\leq} \frac{(r-1)!}{(r-1)^{r-2}} \times \frac{s^s}{s!}.\end{aligned}$$

And the factors in the last term are,

$$\begin{aligned}\frac{(r-1)!}{(r-1)^{r-2}} &\leq 1, \qquad \text{for } r \geq 2\ , \\ \frac{s^s}{s!} &\leq \mathrm{e}^{s-1}, \qquad \text{for } s \geq 1\ ,\end{aligned}$$

As induction easily shows. For example, the latter is verified for $s = 1$ and, if assumed to work for $s > 1$, then

$$\frac{(s+1)^{s+1}}{(s+1)!} = \frac{(s+1)(s+1)^s}{(s+1)s!} = \frac{s^s}{s!}\left(1+\frac{1}{s}\right)^s \leq \mathrm{e}^s,$$

(by the hypothesis, and the fact that $(1+1/s)^s \leq \mathrm{e}$, for all $s \in \mathbb{N}$). Hence, we conclude that ϑ_s, when $s > r-1$, may be bounded by,

$$\vartheta_s \leq d(ed + 2\mathrm{e}a)^{s-1},$$

whilst, if $s = r-1$, then $\delta_{r-1} = \delta_3$, and directly

$$\vartheta_{r-1} \leq \frac{\delta_3}{\delta_{r+1}}(d+2a)\vartheta_{r-2} = (r-1)d(d+2a)^{r-2} \leq d(d\mathrm{e} + 2\mathrm{e}a)^{r-2}.$$

using that $x \leq \mathrm{e}^{x-1}$, for $x \geq 1$. Therefore, the inequality (A.2.19) works for $s \geq r-1$.

Nevertheless, it was supposed, up to now, that $r \geq 4$. To complete the proof, we discuss apart the case $r = 3$, for which direct computation from the definition (A.2.18) shows,

$$\vartheta_0 = 1, \qquad \vartheta_\nu = \frac{\delta_3 d}{\nu\delta_{\nu+2}}\vartheta_{\nu-1},$$

and so:

$$\vartheta_1 = d, \quad \vartheta_2 = \frac{d\delta_3}{2\delta_4}d, \quad \vartheta_3 = \frac{d\delta_3^2}{2\cdot 3\delta_4\delta_5}d^2, \ldots, \vartheta_s = \frac{d\delta_3^{s-1}}{s!\delta_4\delta_5\cdots\delta_{s+2}}d^{s-1},$$

and making the same computations as above, it is seen that,

$$\vartheta_s = \frac{d\delta_3^{s-1}}{s!\delta_4\delta_5\cdots\delta_{s+2}}d^{s-1} \leq \frac{d\delta_3^{s-1}}{\delta_3^{s-1}} \times \frac{s^{s-1}}{s!}d^{s-1} \leq \frac{1}{s}d(\mathrm{e}d)^{s-1}.$$

Thus, estimate (A.2.19) also works when $r = 3$. $\square$

Appendix B

Background

In this appendix, we introduce the essential facts on Hamiltonian dynamical systems which will be necessary along the text. First, classical Hamiltonian systems are introduced, then extended through a geometric point of view, using symplectic geometry. Special emphasis is put on the transformation theory, and on the subsequent normal form reduction, because this will be the main tool we shall use along chapter 1.

Most of the theorems –in particular, those concerning with the geometric approach to the mechanics–, are stated without proof. The interested reader may access to a wide amount of literature related with this subject, specially significant for us are the book of Arnol'd (1974) and the one of Abraham and Marsden (1978).

B.1 Hamiltonian systems

Consider first the following system of ordinary first order differential equations on $\mathbb{R}^{2n}$,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{B.1.1}$$

with $i = 1, \ldots, n$. The variables $\boldsymbol{q}^* = (q_1, \ldots, q_n) \in \mathbb{R}^n$, are the *positions*, while $\boldsymbol{p}^* = (p_1, \ldots, p_n) \in \mathbb{R}^n$ are referred as the *momenta*. Both of them, taken together, are often called *coordinates*. The system (B.1.1) is said to be a *Hamiltonian system* of differential equations, while the function H in (B.1.1) is the *Hamiltonian function* or many times, simply, the *Hamiltonian*. The number n (half times the dimension of the space) is the number of *degrees of freedom* of the system.

If now we introduce the notation $\boldsymbol{\zeta}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*)$, identifying $\zeta_i = q_i$ and $\zeta_{i+n} = p_i$ for $i = 1, \ldots, n$, the equations (B.1.1) can be written in vectorial form as,

$$\dot{\boldsymbol{\zeta}} = J_n \cdot \mathrm{grad} H(\boldsymbol{\zeta}), \tag{B.1.2}$$

where J_n is the matrix of the standard symplectic form (see definition B.2 and theorem B.12 below):

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \tag{B.1.3}$$

being I_n the $n \times n$ identity matrix. The vector field of the Hamiltonian equations (B.1.2), which we shall note by $X_H = J_n \cdot \mathrm{grad} H$, is the *Hamiltonian vector field* associated to the Hamiltonian H.

Now suppose that we make a change of coordinates given by $\boldsymbol{z} = \Phi(\boldsymbol{\zeta})$, with $\Phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ an smooth map. Then, if $\boldsymbol{\zeta}(t)$ is a solution (B.1.2), $\boldsymbol{z}(t) = \Phi(\zeta(t))$ must satisfy, $\dot{\boldsymbol{z}} = S\dot{\boldsymbol{\zeta}} = SJ\mathrm{grad}_{\boldsymbol{\zeta}} H(\boldsymbol{\zeta}) = SJS^*\mathrm{grad}_{\boldsymbol{z}} H(\Phi^{-1}(\boldsymbol{z}))$, where $S^i_j = D_j\Phi_i(\boldsymbol{\zeta})$ is the Jacobian matrix of Φ, S^* the transpose matrix of $\mathbf{S}$ and Φ^{-1} the inverse of Φ. Therefore, the equations for $\boldsymbol{z}$ are of Hamiltonian type, with a new Hamiltonian function given by $K = H \circ \Phi$, if $SJS^* = J$(1). A transformation satisfying this condition is called *canonical* or *symplectic*. If this is the case, the transformed equations can be written explicitly,

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i},$$

for $i = 1, \ldots, n$ and $\boldsymbol{z} = (\boldsymbol{Q}, \boldsymbol{P})$. It is said that "canonical transformations take Hamiltonian systems into Hamiltonian systems".

The above description corresponds to the classical definition of Hamiltonian systems and symplectic transformations. There, the space of the positions and momenta (the *phase space*) is $\mathbb{R}^n \times \mathbb{R}^n$. For many mechanical systems, though, their natural phase space is not an Euclidean space, but a manifold. For example, the phase space for the motion of a rigid body about a fixed point is $T^*SO(3)$, the cotangent bundle of the group of rotations $SO(3)$. See Marsden and Ratiu (1999).

To define a Hamiltonian system on a manifold we need to introduce first several concepts from the symplectic geometry, but before proceeding we fix some useful notation.

Let M be a manifold, then $\mathfrak{F}(M)$ denotes the set of smooth mappings from M into $\mathbb{R}$ and $\mathfrak{X}(M)$, $\Omega^k(M)$ the smooth vector fields and $k-$forms on M respectively. Also, if N is another manifold and $\varphi : M \to N$ is an smooth map, $d\varphi$ stands for the corresponding differential map between the tangent spaces i. e., $d\varphi_m : T_mM \to T_{\varphi(m)}N$, with $m \in M$.

When we have: maps, vector fields, forms, defined on a manifold N, and a regular map $\varphi : M \to N$, from another manifold M to N, there is a standard mechanism to extend these objects into M. This is known as the *pullback* by φ and is usually denoted by φ^*.

First, if $\alpha \in \Omega^k(N)$, the pullback of α by φ, $\varphi^*\alpha$, is a k-form on M (so $\varphi^*\alpha \in \Omega^k(M)$), given by

$$(\varphi^*\alpha)_m(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) = \alpha_{\varphi(m)}(d\varphi_m \cdot \boldsymbol{v}_1, \ldots, d\varphi_m \cdot \boldsymbol{v}_k),$$

with $m \in M$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in T_mM$. In the same way we define the pull back of a map $g \in \mathfrak{F}(N)$ by $\varphi^*g = g \circ \varphi$ and if φ is a diffeomorphism, the pullback of a vector field $Y \in \mathfrak{X}(N)$ as

$$(\varphi^*Y)(m) = (d\varphi_m)^{-1} \cdot Y(\varphi(m)),$$

for any $m \in M$. Note that $\varphi^*Y \in \mathfrak{X}(M)$.

If X is a continuous vector field on M, a finite dimensional differentiable manifold, by the theorem of Peano (see Sotomotor, 1979), there exists, for each $m \in M$, an $\varepsilon = \varepsilon(m) > 0$ and a C^1-map $c : (-\varepsilon, \varepsilon) \to M$ such that $c(0) = m$ and $\dot{c}(t) = \frac{d}{dt}c(t) = X(c(t))$, for all $t \in (-\varepsilon, \varepsilon)$. In addition, if the vector field is smooth, i. e., $X \in \mathfrak{X}^r(M)$ (the set

(1) This is equivalent to the more usual condition $S^*JS = S$

of C^r-fields defined on M) with $r \geq 1$, given two C^1 curves on M, $\alpha_i : I_i \to$ such that $c_i(0) = m$, $\dot{c}_i(t) = X(c_i(t))$ for $i = 1, 2$, then $c_1 = c_2$ on $I_1 \cap I_2$.

The maps c with the properties described in this last paragraph, are called *integral curves* of the field X at the point m. The image of an integral curve at m is known as the *orbit* or *trajectory* of the vector field X through the point m.

Now let $\mathfrak{I}_m$ be the set

$$\mathfrak{I}_m = (\omega_-(m), \omega_+(m)) = \bigcup_{I \in \mathfrak{I}} I, \tag{B.1.4}$$

where $\mathfrak{I}$ is the class of open intervals, $I \subset \mathbb{R}$, such that $0 \in I$ and there exists an integral curve $c : I \to M$ of the vector field X at m (so $c(0) = m$). $\mathfrak{I}_m$ is the *maximal interval* of definition of the integral curves through m at $t = 0$. Then we can define $c : \mathfrak{I}_m \to M$ of class C^1 and such that $c(0) = m$, $\dot{c}(t) = X(c(t))$ for all $t \in \mathfrak{I}_m$. This is the *maximal integral curve* through m at $t = 0$. It is denoted usually by $t \mapsto \phi(t; m) = \phi_t(m) \in M$.

Consider the set $\mathcal{D}_X = \{(t, m) \in \mathbb{R} \times M : t \in \mathfrak{I}_m\}$. The *flow* of X on the manifold M (we denote it again with ϕ), is the map

$$\begin{aligned} \phi : \mathcal{D} \subset \mathbb{R} \times M &\to M \\ (t, m) &\mapsto \phi(t; m) = \phi_t(m), \end{aligned}$$

defined by the properties $\dot{\phi}(t; m) = X(\phi(t; m))$ and $\phi(0; m) = m$. Here, as it has been introduced before, the dot symbol denotes the derivative with respect to the parameter (the time) t.

It can be proved (see the book of Sotomayor referenced above), that $\mathcal{D}_X$ is an open set of $\mathbb{R} \times M$ holding $\{0\} \times M$, and also that the flow ϕ is a C^r map if X is a field of class C^r on M.

When $\mathfrak{I}_m = \mathbb{R}$ for all $m \in M$, the field X is said to be *complete*. This takes place, for example when M is a compact manifold (see the book of Palis and Melo, 1982, for a proof). Then, $\phi_t : M \to M$, $t \in \mathbb{R}$, defined by $m \in M \mapsto \phi_t(m) = \phi(t; m) \in M$ is a diffeomorphism and $\phi_t \circ \phi_s = \phi_{t+s}$ for all $t, s \in \mathbb{R}$. Moreover, the set $\{\phi_t\}_{t \in R}$ defines the *one parameter group of diffeomorphisms*.

If X is not complete, the relation $\phi_t(\phi_s(m)) = \phi_{t+s}(m)$ or equivalently, $\phi(t; \phi(s; m)) = \phi(t + s; m)$ with $m \in M$ holds only whenever both members are defined (that is only if $(s, m), (t, \phi(s, m)) \in \mathcal{D}$). In this case, we say that the flow is *local*.

Remark B.1. In many books, ϕ_t is used also to denote the flow. When t is fixed, and X is complete then it denotes a diffeomorphism $\phi_t : M \to M$. We shall use this convention because both meanings can be usually distinguished from the context. ♠

Let $\alpha \in \Omega^k(M)$, be a k-form on M, $X \in \mathfrak{X}(M)$ and ϕ_t the (local) flow of X. The dynamic definition of the *Lie derivative* of α along X is given by

$$\mathcal{L}_X \alpha = \left.\frac{d}{dt}\right|_{t=0} \phi_t^* \alpha$$

This definition, together with the properties of pullbacks, leads to the *Lie derivative theorem*,

$$\frac{d}{dt} \phi_t^* \alpha = \phi_t^* \mathcal{L}_X \alpha. \tag{B.1.5}$$

(see Abraham, Marsden and Ratiu, 1983, chapter 5).

If $f \in \mathfrak{F}(M)$, the *Lie derivative* of f along X is the *directional derivative*

$$\mathcal{L}_X f = df \cdot X,$$

and the formula (B.1.5) of the Lie derivative theorem has the same expression for functions, i. e.

$$\frac{d}{dt}\phi_t^* f = \phi_t^* \mathcal{L}_X f, \tag{B.1.6}$$

where, as above, ϕ_t is the flow of the vector field X and $f \in \mathfrak{F}(M)$.

For a k-form α on a manifold M and a vector field X, the *interior product* or the *contraction* of X and α, denoted $i_X\alpha$, is defined by

$$(i_X\alpha)_m = \alpha_m\left(X(m), \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}\right),$$

with $m \in M$, and $\boldsymbol{v}_1, \ldots, v_{k-1} \in T_mM$. From this last definition, it is possible to prove that "the pullback of a contraction, $\varphi^* i_X\alpha$, is equal to the contraction of the pullback", i. e.

$$\varphi^* i_X\alpha = i_{\varphi^* X}\varphi^*\alpha. \tag{B.1.7}$$

The Lie derivative and the interior product are related by the *Cartan's magic formula*

$$\mathcal{L}_X\alpha = di_X\alpha + i_X d\alpha, \tag{B.1.8}$$

where d stands for the *exterior derivative* of the corresponding forms. The reader can find a proof of (B.1.8) in the chapter 6 of the same referred book of Abraham, Marsden and Ratiu (1983).

From the commutation of the pullback and the exterior derivative, i. e., $d\varphi^* = \varphi^* d$ and from the Cartan's magic formula (B.1.8), it is easy to obtain the following relation for the pullback of a Lie derivative

$$\varphi^*\mathcal{L}_X\alpha = \mathcal{L}_{\varphi^* X}\varphi^*\alpha. \tag{B.1.9}$$

($\alpha \in \Omega^k(M)$, $X \in \mathfrak{X}(M)$ and $\varphi : M \to N$, a diffeomorphism). So "the pullback of a Lie derivative is the Lie derivative of the pullback".

For a function $f \in \mathfrak{F}(M)$, and directly from the definitions of pullback and Lie derivative: $(\varphi^*\mathcal{L}_X f)(m) = (\mathcal{L}_X f)(\varphi(m)) = df_{\varphi(m)} \cdot X(\varphi(m))$, with $m \in M$, but this can be expressed as

$$\begin{aligned} df_{\varphi(m)} \cdot d\varphi_m \cdot d\varphi^{-1}_{\varphi(m)} \cdot X(\varphi(m)) &= d(f \circ \varphi)_m \cdot (\varphi^* X)(m) \\ &= d(\varphi^* f)_m \cdot (\varphi^* X)(m) = (\mathcal{L}_{\varphi^* X}\varphi^* f)(m), \end{aligned}$$

for all $m \in M$. So the same formula (B.1.9) works also for the pullback of a Lie derivative of a function; i. e.

$$\varphi^*\mathcal{L}_X f = \mathcal{L}_{\varphi^* X}\varphi^* f. \tag{B.1.10}$$

A Hamiltonian system is born from a *symplectic structure* defined on a manifold. Next we introduce this and other related concepts.

Definition B.2. *Let M be a regular or smooth manifold. A symplectic form or a symplectic structure, is a two-form ω^2 on M, such that*

(i) ω^2 is closed: $dw^2 = 0$, and

(ii) for each $m \in M$, $\omega^2_m : T_mM \times T_mM \to \mathbb{R}$ is nondegenerate, i. e.: if $\omega^2_m(\boldsymbol{u}, \boldsymbol{v}) = 0$ for all $\boldsymbol{v} \in T_mM$ then $\boldsymbol{u} = 0$.

The pair (M, ω^2), of a manifold M together with a symplectic form ω^2 on M is a *symplectic manifold*.

Definition B.3. *Let (M, ω^2) be a symplectic manifold, and $H \in \mathfrak{F}(M)$. The vector field X_H determined by the condition*

$$i_{X_H}\omega^2 = dH, \tag{B.1.11}$$

is called the Hamiltonian vector field of the Hamiltonian function H, and (M, ω^2, X_H) is a Hamiltonian system.

Remark B.4. The condition (B.1.11) in the definition above, is equivalent to

$$\omega^2_m\left(X_H(m), \boldsymbol{v}\right) = dH_m \cdot \boldsymbol{v}$$

for all $m \in M$ and $\boldsymbol{v} \in T_mM$. Thus, Nondegeneracy of ω^2 guarantees that X_H exists. ♠

From (B.1.11) it also follows that, on a connected symplectic manifold, any two Hamiltonians for the same X_H have the same differential, so they must differ by a constant.

To each Hamiltonian vector field X_H we associate its Hamilton's equations $\dot{x} = X_H$, whose solutions are integral curves of the field X_H. This is the "natural" extension of the Hamilton's equations (B.1.1) on manifolds.

Example B.5. For $M = \mathbb{R}^{2n}$ (or $\mathbb{C}^{2n}$) and if ω^2 is the *standard canonical two form* $\omega^2 = \sum_{i=1}^n dq_i \wedge dp_i$ (so $\omega^2(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^* J \boldsymbol{v}$, with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{2n}$ (or $\mathbb{C}^{2n}$)), it is readily checked that in this case $X_H = J_n \cdot \mathrm{grad} H$, and the Hamiltonian equations are the ones given by (B.1.2). ◇

Next we introduce the notion of symplectic map. It just generalizes our early definition of canonical transformation given on page 160 (see proposition B.9 below).

Definition B.6. *A map $\varphi : M \to N$ between symplectic manifolds (M, ω^2) and (N, α^2) is called symplectic if $\varphi^*\alpha^2 = \omega^2$. Note: when $\varphi : M \to M$, then φ is symplectic if $\varphi^*\omega^2 = \omega^2$.*

Example B.7. For $M = \mathbb{R}^{2n}$ and ω^2 the standard symplectic two form given in example B.5, and a diffeomorphism $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$; the just given definition of symplectic map reduces to the condition $S^*JS = J$, for in this case

$$\left(\varphi^*\omega^2\right)_x(\boldsymbol{u}, \boldsymbol{v}) = \omega^2\left(d\varphi_x \cdot u, d\varphi_x \cdot v\right) = \boldsymbol{u}^* S^* J S \boldsymbol{v},$$

where S is the Jacobian matrix of φ i. e., $S^i_j = D_j\varphi_i(x)$, as before. But the last term should be equal to $\omega^2(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^* J \boldsymbol{v}$, for all $x \in \mathbb{R}^{2n}$ and for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{2n}$, so it must be $S^*JS = J$, which is equivalent to our first definition of canonical transformation, $SJS^* = J$,

$$SJS^* = J \Leftrightarrow \left(JS^{-1}\right) SJS^* \left(JS\right) = \left(JS^{-1}\right) J \left(JS\right) \Leftrightarrow S^*JS = J,$$

and the property $J^2 = -I_{2n}$ is used. ◇

The following theorem provides an important result: the flow of a Hamiltonian vector field is, for each fixed time t, a symplectic map which leaves the Hamiltonian function invariant.

Theorem B.8. *Let X_H be a Hamiltonian vector field on the symplectic manifold (M, ω^2), and let ϕ_t be the flow of X_H. Then*

(i) ϕ_t is symplectic; i. e.: $\phi_t^* \omega^2 = \omega^2$,

(ii) H is constant along the the flow, i. e. $H \circ \phi_t = H$.

Proof. The proof of (i) follows immediately from the Lie derivative theorem (B.1.5), and the application of the Cartan's formula

$$\frac{d}{dt}\left(\phi_t^* \omega^2\right) = \phi_t^* \mathcal{L}_{X_H} \omega^2 = \phi_t^* \left(d i_{X_H} \omega^2 + i_{X_H} d\omega^2\right),$$

but $d\omega^2 = 0$, because ω^2 is a closed form and $d i_{X_H} \omega^2 = d(dH) = 0$. Thus, $\phi_t^* \omega^2 = \phi_{t=0}^* \omega^2 = \omega^2$.

To prove (ii), consider $c(t)$ to be a integral curve of X_H. Then applying the chain rule and the formula (B.1.11) of the definition B.3,

$$\begin{aligned}\frac{d}{dt}(H \circ c)(t) &= dH_{c(t)} \cdot X_H\left(c(t)\right) \\ &= \omega^2_{c(t)}\left(X_H\left(c(t)\right), X_H\left(c(t)\right)\right) = 0.\end{aligned}$$

by the skew-symmetry of ω^2. So H is constant along any integral curve of X_H. In particular along those ones given by the flow. □

The Hamiltonian H is often referred as the *energy* of the Hamiltonian system (M, ω^2, X_H). So (ii) states that the energy is conserved.

Now, we define for any two functions $f, g \in \mathfrak{F}(M)$, their *Poisson bracket* by

$$\{f, g\} = \omega^2(X_f, X_g), \tag{B.1.12}$$

where X_f and X_g are the Hamiltonian vector fields associated to the functions f and g as given by (B.1.11).

The Poisson brackets can also be expressed in terms of the Lie derivative, since

$$\mathcal{L}_{X_f} g = i_{x_f} dg = i_{x_f} i_{x_g} \omega^2 = \omega^2(X_f, X_g) = -\omega^2(X_g, X_f) = -\mathcal{L}_{X_g} f,$$

and therefore

$$\{f, g\} = -\mathcal{L}_{X_f} g = \mathcal{L}_{X_g} f.$$

Note that, for $M = \mathbb{R}^{2n}$, and for the standard symplectic 2-form, the Poisson bracket of two functions f and g can be expressed as

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}, \tag{B.1.13}$$

or, also

$$\{f, g\} = (\operatorname{grad} f)^* J \cdot \operatorname{grad} g. \tag{B.1.14}$$

in vectorial notation. Here $(\text{grad} f)^*$ is the transpose of the gradient of f.

Now, let (M, ω^2) and (N, α^2) be symplectic manifolds. A diffeomorphism $\varphi : M \to N$ preserves the Poisson bracket if $\varphi^* \{f, g\} = \{\varphi^* f, \varphi^* g\}$, for all $f, g \in \mathfrak{F}(N)$.

The Poisson bracket gives an straight characterization of the symplectic maps as those which leave it invariant.

Proposition B.9. *With the notation introduced in the paragraph above, the following are equivalent:*

(i) φ *is symplectic.*

(ii) φ *preserves the Poisson bracket of any functions* $f, g \in \mathfrak{F}(N)$.

(iii) $\varphi^* X_f = X_{\varphi^* f}$*, for any* $f \in \mathfrak{F}(N)$.

Proof. First we see that (ii) is equivalent to (iii), since

$$\varphi^* \{f, g\} = \varphi^* \mathcal{L}_{X_g} f = \mathcal{L}_{\varphi^* X_g} \varphi^* f = \{\varphi^* f, \varphi^* g\} .$$

Now consider

$$i_{X_{\varphi^* f}} \omega^2 = d(\varphi^* f) = \varphi^*(df) = \varphi^* i_{X_f} \alpha^2 = i_{\varphi^* X_f} \varphi^* \alpha^2$$

then, by the nondegeneracy of the 2-forms ω^2 and α^2, and by the fact that any $\boldsymbol{v} \in T_m M$ equals to some $X_h(z)$ for a function $h \in \mathfrak{F}(N)$, it follows that φ is symplectic if and only if $\varphi^* X_f = X_{\varphi^* f}$, for all $f \in \mathfrak{F}(N)$. So, it is proved that (i) and (iii) are equivalent. This ends the proof of the proposition. □

Remark B.10. (iii) implies that Hamilton's equations are preserved under canonical transformations. Recall that from our initial definition of symplectic transformation (page 160 and see also example B.7), we saw directly the conservation of the Hamiltonian equations. (iii) adds that the converse is also true. ♠

Proposition B.9 allows to generalize the conservation of the energy established in theorem B.8. We do this through the following corollary.

Corollary B.11 (of proposition B.9). *Consider a Hamiltonian vector field* X_H *on the symplectic manifold* (M, ω^2)*, and let* ϕ_t *be its corresponding flow; then, for any* $f \in \mathfrak{F}(M)$*, we have*

$$\frac{d}{dt}(f \circ \phi_t) = \{f, H\} \circ \phi_t = \{f \circ \phi_t, H\} \tag{B.1.15}$$

Proof. Using formula (B.1.6) of the Lie derivative theorem and the definition of the Poisson bracket in terms of the Lie derivative, $\{g, f\} = \mathcal{L}_{X_f} g$, we obtain

$$\begin{aligned} \frac{d}{dt}(f \circ \phi_t) = \frac{d}{dt} \phi_t^* f &= \phi_t^* \mathcal{L}_{X_H} f \\ &= \phi_t^* \{f, H\} = \{\phi_t^* f, \phi_t^* H\} = \{f \circ \phi_t, H\}, \end{aligned}$$

where the preservation of the Poisson bracket and the Hamiltonian under the flow ϕ_t has been applied. □

A function $g \in \mathfrak{F}(M)$ is in *involution* or *Poisson commute* if $\{g, H\} = 0$. Then, by (B.1.15), g is constant along the flow of the Hamiltonian vector field X_H. This functions are called *integrals* or *constants of the motion*. A classical theorem of Liouville states that, if a Hamiltonian n-degrees of freedom system has k functionally independent integrals in involution, then the number of degrees of freedom can be reduced to $n-k$. When $k=n$ it is said that the system is *integrable*. Then, and under certain additional hypotheses, the trajectories of the system in the phase space are straight lines on high-dimensional cylinders or tori and the Hamiltonian equations can be integrated by quadratures. When the motion takes place on tori, it is possible (though not always trivial) to introduce the so called *action-angle variables*.

For a more precise formulation and for the proof of the Liouville theorem, beyond the outline given here, the reader is aimed to consult the book of Abraham and Marsden (1978) and also the one of Arnol'd (1974). In Goldstein (1980), chapter 10, there are several examples and exercises on integration of mechanical systems by changing to action-angle variables.

The next theorem states that, locally, it is possible to write the Hamilton's equations in the form (B.1.1) of page 159.

Theorem B.12 (Darboux). *Consider the symplectic manifold (M, ω^2) and $m \in M$. Then there exists a chart (U, ψ), with $\psi(m) = 0$; such that*

$$\omega^2\big|_U = \sum_{i=1}^{n} dq_i \wedge dp_i,$$

being $(q_1, \ldots, q_n; p_1, \ldots, p_n)$ the coordinate functions of the chart map ψ. This coordinates are often referred as symplectic *or* canonical coordinates.

In particular, it follows that the manifold M is even dimensional. The Darboux's theorem allows us to extend, over the whole manifold M, any local result proved for the standard symplectic manifold $(\mathbb{R}^{2n}, \omega^2 = \sum_{i=1}^{n} dq_i \wedge dp_i)$, whenever it is invariant under canonical transformations. For a constructive proof see Arnol'd (1974).

Henceforth we shall consider that the Darboux's theorem has been applied so in the rest of the present monograph we shall work by default with the standard symplectic manifold; so if nothing is stated in the contrary sense, $M = \mathbb{R}^{2n}$ and $\omega^2 = \sum_{i=1}^{n} dq_i \wedge dp_i$.

We also want to stress that all the definitions and theorems presented on these section are valid on *finite dimensional* manifolds. In the different books quoted along the text, it is possible to find generalizations of the corresponding notions and results on infinite dimensional manifolds.

B.2 Poincaré maps

Poincaré maps are a useful trick in the study of dynamical systems. They transform continuous dynamical systems (flows) into discrete (mappings) at the same time that reduce the dimension. The material we include here is taken from Delshams (1994). Also good references are the books of Sotomotor (1979) and Palis and Melo (1982).

Consider $p_0 \in \mathbb{R}^m$ a nonsingular point of a smooth vector field $X \in \mathfrak{X}(R^m)$ (not necessarily Hamiltonian); this means that $X(p_0) \neq 0$. Let $p_1 = \phi(T; p_0)$ with $T \neq 0$ be

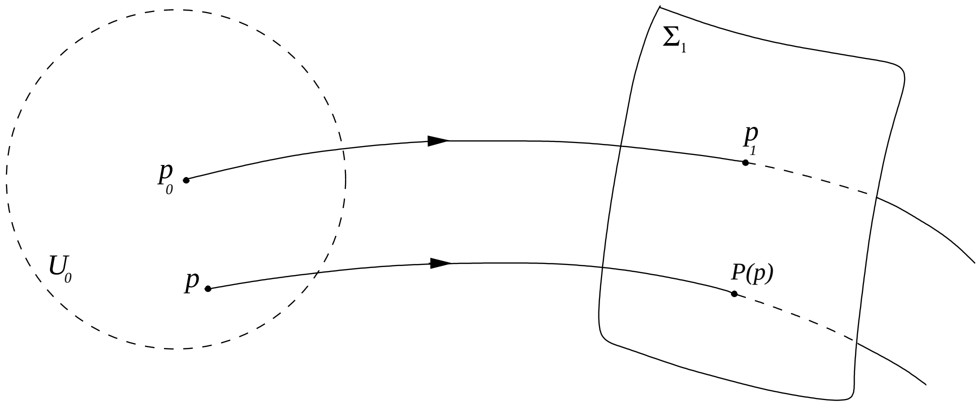


Figure B.1: Poincaré map

another point on the integral curve of X at p_0 and Σ_1 a *transversal section* of the field at p_1, i. e., Σ_1 is a smooth $m-1$ hypersurface and $X(p_1) \notin T_{p_1}\Sigma_1$.

We shall suppose that, there exists a regular function $h : U_1 \to \mathbb{R}$, with U_1 a neighborhood of p_1, such that $\Sigma_1 \cap U_1 = \{x \in \mathbb{R}^m : h(x) = 0\}$ and with $Dh(x) \neq 0$ for all $x \in U_1$. Now consider the map $(x,t) \mapsto h(\phi(t;x))$, defined in a neighborhood of p_0. By the implicit function theorem, we have neighborhoods U_0 of p_0, I of T and an unique *time map* $\tau : U_0 \to I$ such that

$$\phi(t;x) \in \Sigma_1 \text{ with } (t,x) \in I \times U_0 \Leftrightarrow t = \tau(x).$$

(in particular, $\tau(p_0) = T$).

The map $x \in U_0 \mapsto P(x) = \phi\left(\tau(x);x\right) \in \Sigma_1$ is called the *Poincaré map*. It is a smooth map, though *degenerate* in the following sense: given $p \in U_0$, on the piece of the orbit at p

$$\mathcal{O}_{U_0}(p) = \{x = \phi(t;p) : \phi(s;p) \in U_0 \text{ for all } s \in [0,t_p]\}$$

contained in U_0, the map P is constant, i. e., $P(p) = P(x)$, for all $x \in \mathcal{O}_{U_0}(p)$ (see figure B.2). To avoid this degeneration, we select another transversal section, Σ_0, at p_0 and restrict $P|_{\Sigma_0} : \Sigma_0 \cap U_0 \to P\left(\Sigma_0 \cap U_0\right)$ (see figure B.2). In fact, the *restricted* map is the so called Poincaré map. Furthermore, it is not difficult to prove that P is a diffeomorphism.

An interesting case is for periodic orbits, $p_1 = \phi\left(T;p_0\right) = p_0$. Then we can take $\Sigma_0 = \Sigma_1$, because $P(\Sigma_0 \cap U_0) \cap (\Sigma_0 \cap U_0) \neq \emptyset$. For this cases we could iterate P and consider P^n, when possible. Note that we have lower in one unit the dimension of the dynamical system, which now is discrete for it is described by a diffeomorphism $P : \Sigma' = \Sigma_0 \cap U_0 \to P(\Sigma') \subset \Sigma_0$.

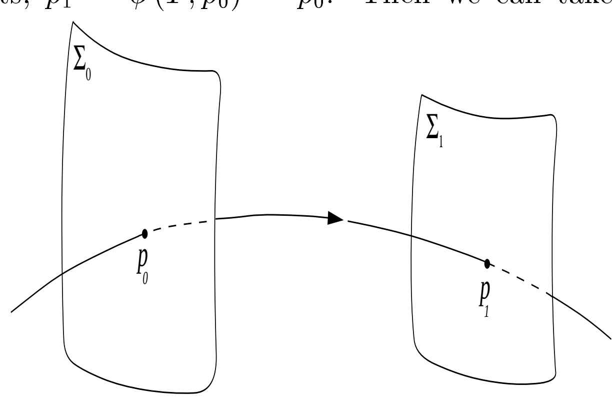


Figure B.2: Transversal section Σ_0 at p_0.

Moreover, the dynamics properties of the flow X are translated to the map P. Thus, periodic points on the map correspond to periodic orbits of X of the same hyperbolic type (see below) and if $\mathcal{A}$ is an invariant set under the flow of X, then the intersection $\Sigma' \cap \mathcal{A}$ is invariant under P, and so on.

For example, it can be seen the relation between the differential matrices DP and $D\phi = \frac{\partial \phi}{\partial x}$ (for the notation, see remark B.14 at the next section).

For a given orbit (not necessarily periodic), consider the Poincaré map $P : \Sigma_0 \to \Sigma_1$ of the field X and the transversal sections Σ_0, Σ_1 defined above. We have $p_1 = \phi(\tau(p_0); p_0)$ and $X(p_0) \neq 0$. Σ_i are transversal sections at p_i so $X(p_i) \notin T_{p_i}\Sigma_i$ with $i = 0, 1$.

On U_0 we can define $P(x) = \phi(\tau(x); x)$ with $\tau(p_0) = p_1$. Consider now the direct sum $\mathbb{R}^m = \operatorname{Span}\{X(p_i)\} \oplus T_{p_i}\Sigma_i$; $i = 0, 1$. To fix ideas, suppose we choose two bases for $\mathbb{R}^m$, say $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$; given by $\boldsymbol{u}_1 = X(p_0)$, $\operatorname{Span}\{\boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\} = T_{p_0}\Sigma_0$ and in the same way, $\boldsymbol{v}_1 = X(p_1)$, $\operatorname{Span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\} = T_{p_1}\Sigma_1$.

Next, we apply the relations (see remark B.14 below):

$$D\phi(\tau(p_0); p_0) \cdot X(p_0) = X(p_1), \tag{B.2.1}$$

and

$$DP(p_0) = X(p_1) D\tau(p_0) + D\phi(\tau(p_0); p_0), \tag{B.2.2}$$

to the vectors of the basis $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$, to obtain

$$\begin{aligned} D\phi(\tau(p_0); p_0) \cdot \boldsymbol{u}_1 &= \boldsymbol{v}_1, \\ D\phi(\tau(p_0); p_0) \cdot \boldsymbol{u}_i &= DP(p_0) \cdot \boldsymbol{u}_i - (D\tau(p_0) \cdot \boldsymbol{u}_i) X(p_1), \end{aligned}$$

for $i = 2, \ldots, m$ in the second equation. But note that $(D\tau(p_0) \cdot \boldsymbol{u}_i) X(p_1) \in \operatorname{Span}\{X(p_1)\}$ and $DP(p_0) \cdot \boldsymbol{u}_i \in T_{p_1}\Sigma_1$. So the following proposition is proved.

Proposition B.13. *Let $P : \Sigma_0 \cap U_0 \to \Sigma_1$ be the Poincaré map of the flow of $X \in \mathfrak{X}(\mathbb{R}^m)$, with the transversal sections Σ_0 and Σ_1. Consider any point $p_0 \in \Sigma_0 \cap U_0$, $p_1 = P(p_0) = \phi(T; p_0)$. Then in the bases $\{\boldsymbol{u}_i\}_{i=1,\ldots,m}$, $\{\boldsymbol{v}_i\}_{i=1,\ldots,m}$ described above, the matrices $D\phi(T; p_0)$ and $DP(p_0)$ are related by*

$$D\phi(T; p_0) = \left(\begin{array}{c|c} 1 & \alpha \\ \hline 0 & DP(p_0) \end{array} \right),$$

with $\alpha = -D\tau(p_0) : T_{p_0}\Sigma \to \mathbb{R}$.

Remark B.14. $D\phi(t; p_0) = \frac{\partial \phi}{\partial x}(t; p_0)$, so it is the derivative with respect to the initial conditions of the flow, and satisfies the following initial value problem:

$$\frac{d}{dt} D\phi(t; p_0) = DX(\phi(t; p_0))\, D\phi(t; p_0), \quad D\phi(0; p_0) = I_m$$

(and I_m is the $m \times m$ identity matrix). Thus, $D\phi(t; p_0)$ is a *fundamental* (in fact the *principal* at $t = 0$) matrix of the linear system

$$\dot{x} = A(t)\, x, \tag{B.2.3}$$

with $A(t) = X(\phi(t; p_0))$. We say that the equations (B.2.3) are –for the vector field X–, the *first variational equations* of the orbit at p_0. ♣

B.2.1 Poincaré maps of periodic orbits

For a Poincaré map associated to a periodic orbit i. e., when $p_1 = \phi(T; p_0) = p_0$, with $X(p_0) \neq 0$ and taking $\Sigma_0 = \Sigma_1$, the matrix $A(t)$ in (B.2.3) is T-periodic: $A(t) = A(t+T)$. Then $D\phi(T; p_0)$, is a *monodromy matrix* (since $D\phi(t+T; p_0) = D\phi(t; p_0)\, D\phi(T; p_0)$) and its eigenvalues are called the *characteristic multipliers* of the orbit. Strictly speaking, the characteristic multipliers of a system (B.2.3) with $t \mapsto A(t)$ continuous and T-periodic are defined to be the eigenvalues of *any* monodromy matrix.

In fact, the characteristic multipliers do not depend on the particular monodromy matrix chosen; that is, the particular fundamental solution used to define the monodromy matrix: if $\Psi(t)$ is a fundamental matrix solution with monodromy matrix C (so $\Psi(t+T) = \Psi(t)\, C$) and $\widetilde{\Psi}(t)$ is another fundamental matrix; then there exists a nonsingular matrix, D, such that $\widetilde{\Psi}(t) = \Psi(t)D$. Hence $\widetilde{\Psi}(t+T) = \widetilde{\Psi}(t)D^{-1}CD$ so the monodromy matrix for $\widetilde{\Psi}(t)$ is $D^{-1}CD$; and similar matrices have the same eigenvalues.

Applying the Floquet theorem, any fundamental matrix $\Psi(t)$ of (B.2.3) can be written as the product of two $m \times m$ matrices

$$\Psi(t) = P(t) \exp(tB), \tag{B.2.4}$$

with $P(t)$ T-periodic and B a constant matrix given by $M = \exp(TB)$, where M is the monodromy matrix of ϕ (so $\Psi(t+T) = \Psi(t)\, M$). As M is not singular, the matrix B exists though it will be complex in general. For a more complete account on Floquet's theorem see Smale (1974), or practically any text book on differential equations.

If λ is a characteristic multiplier (i. e. an eigenvalue of the monodromy matrix M), each complex number μ such that

$$\lambda = e^{T\mu},$$

is called *characteristic* or *Floquet* exponents. Note that the imaginary part of μ is not determined uniquely, since $\mu + i2k\pi/T$ also verifies the above condition for $k \in \mathbb{Z}$. As the characteristic multipliers are determined uniquely, it is usual to choose the characteristic exponents such that they coincide with the eigenvalues of B in (B.2.4).

With this choice, μ_i is a characteristic exponent if and only if $\lambda_i = e^{T\mu_i}$ is a characteristic multiplier, so the solution is asymptotically stable if and only if:

$$\operatorname{Re}\mu_i < 0. \text{ for all } i = 1, \ldots, m \ (\Leftrightarrow |\lambda_i| < 1, \text{ for all } i = 1, \ldots, m)$$

It turns out that, if γ is a periodic orbit, its associated Poincaré map $P : \Sigma_0 \cap U_0 \to \Sigma_0$ does not depend upon the selected point p_0, $p_0 \in \gamma$ "up to C^r-conjugations"; more precisely, in the sense of the following.

Proposition B.15. *Let γ be a T-periodic orbit of $X \in \mathfrak{X}^r(M)$ and Σ_0 a transversal section of X at a point $p_0 \in \gamma$. We define the Poincaré map $P_0 : \Sigma_0 \cap U_0 \to \Sigma_0 \cap U_0'$ by $x \mapsto \phi(\tau(x); x)$, $\tau(p_0) = T$. Then, if P_1 is another Poincaré map associated to γ at another point $p_1 \in \gamma$, there exists a C^r-map $h : W_0 \to W_1$, with W_i a neighborhood of p_i such that $W_i \subset U_i \cap U_i'$, $i = 0, 1$; satisfying $P_1 \circ h = h \circ P_0$ on W_0. Furthermore h is a C^r-diffeomorphism; i. e., there exists a C^r-map $g : W_1 \to W_0$ such that $g = h^{-1}$ on $W_1 \cap \Sigma_1$.*

So, if $A(t) = DX(\phi(t;p_0))$ with $\phi(t+T;p_0) = \phi(t;p_0)$, the set

$$\{\lambda \in \mathbb{C},\ \text{characteristic multiplier of } \dot{x} = A(t)\,x\} = \{1\} \cup \operatorname{Spec}(DP(p_0)),$$

because 1 is always an eigenvalue of $D\phi(T;p_0)$ with eigenvector $X(p_0)$, since

$$D\phi(T;p_0)X(p_0) = X(p_0).$$

Therefore the characteristic multipliers which determine the behavior of the periodic orbit γ are those ones of $DP(p_0)$; and the study of the linear stability of the periodic orbit γ is reduced to the stability of the fixed point p_0 of the Poincaré map P. For example, p_0 (and thus γ) is hyperbolic when $\operatorname{Spec}(DP(p_0)) \cap S^1 = \emptyset$, i. e., if all the characteristic multipliers, but one lie outside the unit circle in the complex plane.

As through the present work we shall deal with Hamiltonian systems, and they always have at least the integral corresponding to the energy (see theorem B.8), the following proposition becomes of interest.

Proposition B.16. *Let X be a smooth vector field on $\mathbb{R}^m$, $X \in \mathfrak{X}^r(\mathbb{R}^m)$ with $r \geq 1$ and $F : M \to \mathbb{R}$ of class C^1, a first integral of X. Consider the flow $\phi : \mathcal{D} \subset \mathbb{R} \times M \to M$ of the vector field X, a T-periodic orbit $\gamma = \{\phi(t,p_0), t \in [0,T]\}$ and the Poincaré map associated to $p_0 \in \gamma$. Then 1 is an eigenvalue of $DP(p_0)$, and an eigenvalue of $D\phi(T;p_0)$ with multiplicity at least two.*

To prove this proposition we need a previous result

Lemma B.17. *The map*

$$\begin{array}{rcl} W : \mathfrak{J}_{p_0} \times \mathbb{R}^m & \to & \mathbb{R} \\ (t, \boldsymbol{v}) & \mapsto & DF(\phi(t;p_0))\,\boldsymbol{v}, \end{array}$$

is a time-dependent first integral of the variational equation $\dot{x} = A(t)\,x$ associated to the periodic orbit of the vector field X at p_0.

Proof Of the lemma. $\boldsymbol{v}(t)$ is a solution of the variational equation if and only if $\boldsymbol{v}(t) = D\phi(t;p_0)\boldsymbol{v}(0)$. Therefore,

$$W(t, \boldsymbol{v}(t)) = DF(\phi(t;p_0))\,D\phi(t;p_0)\boldsymbol{v}(0) = DF(p_0)\boldsymbol{v}(0) = W(0, \boldsymbol{v}(0)),$$

as follows by deriving both sides of $F(\phi(t;x)) = F(x)$ with respect to x and then substitute $x = p_0$. $\square$

Proof of proposition B.16. By the lemma we have that

$$DF(\phi(t;p_0))\,D\phi(t;p_0) = DF(\phi(0;p_0))\,D\phi(0;p_0) = DF(p_0).$$

Taking $t = T$, (and therefore $\varphi(T;p_0) = p_0$), in the expression above, it reduces to $DF(p_0)D\phi(T;p_0) = DF(p_0)$, or equivalently,

$$(D\phi(T;p_0))^* \operatorname{grad}F(p_0) = \operatorname{grad}F(p_0).$$

So grad$F(p_0)$ is an eigenvector of $(D\phi(T;p_0))^*$ associated to an eigenvalue equal to 1. By the other hand, we already know that $D\phi(T;p_0)X(p_0) = X(p_0)$. Also, the relation: $\langle \mathrm{grad}F(p_0), X(p_0)\rangle = DF(p_0)X(p_0) = 0$ must be satisfied since F is a first integral. From here, it follows that the generalized eigenspace $E_1 = \{\boldsymbol{v} \in \mathbb{R}^m : (D\phi(T;p_0) - I_m)^m = 0\}$, has, at least, dimension 2, i. e., 1 is an eigenvalue of the monodromy matrix with multiplicity at least two. If this is not so, and $\dim E_1 = 1$, then the rest of the eigenvalues should be different from one and therefore, in a suitable basis the monodromy could be expressed as

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A_{m-1} \end{array}\right);$$

thus, in this basis $DF(p_0) = e_1^*$ and $X(p_0) = e_1$ so $DF(p_0)\cdot X(p_0) \neq 0$; which cannot be possible because F is a first integral of the vector field X. □

Without loss of generality, we can suppose grad$F(p_0) \in T_{p_0}\Sigma_0$; e. g., we can take a surface of section Σ_0 such that, $T_{p_0}\Sigma_0 = X(p_0)^{\perp}$ (see proposition B.15). In this case, taking $\boldsymbol{v}_1 = X(p_0)$, $\boldsymbol{v}_2 = \mathrm{grad}F(p_0)$ and $\boldsymbol{v}_3, \ldots, \boldsymbol{v}_m$ such that Span$\{\boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_m\} = T_{p_0}\Sigma_0$. Then in the basis $\{\boldsymbol{v}_i\}_{1\le i\le m}$, the matrix of monodromy reads

$$D\phi(T;p_0) = \left(\begin{array}{c|cccc} 1 & * & * & \cdots & * \\ \hline 0 & 1 & 0 & \cdots & 0 \\ \vdots & * & & & \\ \vdots & \vdots & & DP_{F_0}(p_0) & \\ 0 & * & & & \end{array}\right).$$

The box $(m-2)\times(m-2)$ noted as $DP_{F_0}(p_0)$ in the matrix above, corresponds to the Jacobian matrix at the point p_0 of the *reduced Poincaré map* $P_{F_0} : \Sigma_{F_0} \cap U_0 \to \Sigma_{F_0}$, $x \mapsto \phi(\tau(x); x)$, where

$$\Sigma_{F_0} = \{x \in \Sigma : F(x) = F(p_0) = F_0\}, \tag{B.2.5}$$

and it is a $(m-2)$-dimensional manifold of $\mathbb{R}^m$.

If we have $r \ge 1$ first integrals $F_1, \ldots, F_r$ with grad$F_1(p_0), \ldots$ grad$F_r(p_0)$ linear independent, it is often possible to reduce the dynamical system in $r+1$ dimensions and work on

$$\Sigma_{F_1^0,\ldots,F_r^0} = \left\{x \in \Sigma : F_j(x) = F_j(p_0) = F_j^0, j = 1, \ldots, r\right\}.$$

When this is possible, 1 is an eigenvalue of multiplicity at least $r+1$ of $D\phi(T;p_0)$ (multiplicity r of $DP(p_0)$) and in such cases, periodic orbits are not isolate, but usually we have *families* of periodic orbits.

B.2.2 Stability for three degrees of freedom Hamiltonian systems

In this section we introduce some notations and conventions for the study of the (linear) stability of Hamiltonian systems with three degrees of freedom. The approach we present here is classical and it can be found mainly in Broucke (1969).

In particular we are interested in the study of (linear) stability of periodic orbits. Since autonomous (i. e., not time dependent) Hamiltonian systems have always a first integral corresponding to the energy, we can fix the energy level $H = h_0$, and work with the reduced Poincaré map $P_{h_0} : \Sigma_{h_0} \to \Sigma_{h_0}$, at some convenient point p_0 on the periodic orbit. Here, the *isoenergetic* surface of section Σ_{h_0} is given by (B.2.5) with $F_0 = h_0 = H(p_0)$. By definition $P_{h_0}(p_0) = p_0$, and the study of the stability of the periodic orbit reduces to the study of the stability of the fixed point p_0 of the map P_{h_0}. Thus, the linear normal behavior of the periodic orbit is determined by the eigenvalues of the differential of the Poincaré map at p_0, $DP_{h_0}(p_0)$.

But P_{h_0} is a symplectic map, so $DP_{h_0}(p_0)$ is a linear symplectic map. The following is a well known result (see Arnol'd, 1974).

Proposition B.18. *The characteristic polynomial of a real symplectic transformation* $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$,

$$p(\lambda) = \det(A - \lambda I_{2n})$$

is cyclic, i. e., $p(\lambda) = \lambda^{2n} p\left(\frac{1}{\lambda}\right)$.

Then, if λ is a eigenvalue of A, $1/\lambda$ must also be an eigenvalue of A. On the other hand, the characteristic polynomial is real, so if λ is a complex eigenvalue, its complex conjugate $\overline{\lambda}$ must also be an eigenvalue. Thus, in a real symplectic map, the eigenvalues appear;

(i) in 4-tuples: λ, $\overline{\lambda}$, $1/\lambda$, $1/\overline{\lambda}$, $(|\lambda| \neq 1, \operatorname{Im}\lambda \neq 0)$,

(ii) pairs on the real axis: $\lambda = \overline{\lambda}$, $1/\lambda = 1/\overline{\lambda}$,

(iii) pairs on the unit circle: $\lambda = 1/\overline{\lambda}$, $\overline{\lambda} = 1/\lambda$.

A linear map A is *stable* if, for any given $\varepsilon > 0$, there exists $\delta > 0$ such that $|x| < \delta \Rightarrow |A^q x| < \varepsilon$, for all $q \in \mathbb{N}$. With this definition, it is straightforward to deduce that if all the $2n$ eigenvalues of a linear symplectic map A, are distinct and lie on the unit circle of the complex plane, then the map is stable. Even more, then, A is proved to be *strongly stable*, which means that any other symplectic map, A_1, "close enough" to A is also stable.

In our particular case, $A = DP_{h_0}(p_0)$ and $n = 2$. Proposition B.18 implies thus that all the possible distributions of the four eigenvalues λ_1, $1/\lambda_1$, λ_2, $1/\lambda_2$ on the complex plane are the ones plotted in figure B.4. Another consequence of the proposition is that the characteristic polynomial may be written in the following form

$$p(\lambda) = \lambda^4 + \alpha\lambda^3 + \beta\lambda^2 + \alpha\lambda + 1 \tag{B.2.6}$$

An important fact is that the stability only depends on the coefficients α and β of this polynomial.

In astronomy and in celestial mechanics, it is usual to define the *stability indices* of the map (and so, of the corresponding orbit) by

$$b_i = \lambda_i + \frac{1}{\lambda_i}, \tag{B.2.7}$$

$(i = 1, 2)$, and it can be immediately seen that these indices are related with the coefficients of the characteristic polynomial (B.2.6), α, β through

$$\alpha = -(b_1 + b_2), \quad \beta = 2 + b_1 b_2. \tag{B.2.8}$$

From these two relations it follows that the stability indices are given by the solutions of

$$x^2 + \alpha x + (\beta - 2) = 0, \tag{B.2.9}$$

which are,

$$b_{1,2} = \frac{-\alpha \pm \sqrt{\Delta}}{2}$$

where we have introduced the quantity

$$\Delta = \alpha^2 - 4\beta + 8. \tag{B.2.10}$$

By the definition of the stability indices (B.2.7), it is seen immediately that the eigenvalues λ_1, $1/\lambda_1$ are the solutions of the two degree equation

$$x^2 - b_1 x + 1 = 0, \tag{B.2.11}$$

given by

$$\lambda_1, \lambda_1^{-1} = \frac{b_1 \pm \sqrt{b_1^2 - 4}}{2},$$

and in the same manner, the other reciprocal pair λ_2, $1/\lambda_2$ are the solutions of

$$x^2 - b_2 x + 1 = 0, \tag{B.2.12}$$

and given by

$$\lambda_2, \lambda_2^{-1} = \frac{b_2 \pm \sqrt{b_1^2 - 4}}{2}.$$

Then, the problem of finding the roots of the characteristic polynomial (and hence, the eigenvalues) is reduced to solving the three two order equations (B.2.9), (B.2.11) and (B.2.12). The discriminant Δ of (B.2.9) is zero for $\beta = \alpha^2/4 + 2$. In the plane (α, β), this is the equation of a parabola with its apex at $(\alpha, \beta) = (0, 2)$, and the β axis the symmetry axis.

The equations for b_1 and b_2 have zero discriminant along the straight lines $\beta = 2\alpha - 2$ and $\beta = -2\alpha - 2$ respectively. These are tangents to the parabola $\Delta = 0$ at the points $(4, 6)$ and $(-4, 6)$. The parabola, together with these two straight lines bound seven regions in the plane (α, β). They correspond to the seven possible distributions of the eigenvalues with respect the unit circle in the complex plane plotted in the figures B.4(a) to B.4(g).

Thus, periodic orbits (or equivalently, the fixed points of their associated Poincaré maps) can be classified by the position in the *Broucke diagram* (figure B.3 above) of their characteristic polynomial coefficients (α, β) –see (B.2.6)–. Depending upon the region in the diagram the point (α, β) belongs to, the periodic orbit may be stable or present six different types of instability.

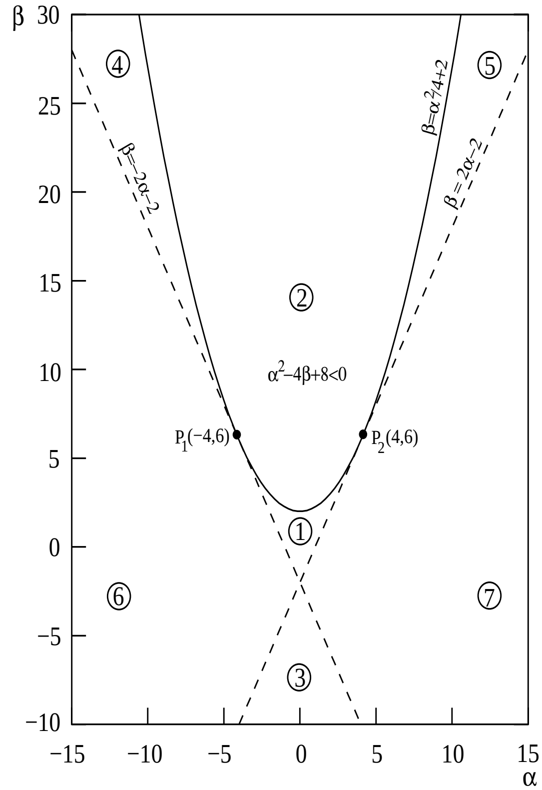


Figure B.3: Broucke Diagram.

Next, we shall briefly describe the relation between the distribution of the (nontrivial) eigenvalues of the periodic orbit with respect the unit circle, and their representation in the Broucke diagram. For a more complete account, see Broucke (1969).

B.2.3 Regions in the Broucke diagram

The connection between the seven regions on the diagram of figure B.3 and the seven types of stability depicted in the figure B.4, can be described briefly as follows:

Region 1. $\Delta > 0$ and $b_1, b_2 \in \mathbb{R}$ with $b_1^2 < 4$, $b_2^2 < 4$. Then $\lambda_1, 1/\lambda_1$ are complex conjugates, and so are $\lambda_2, 1/\lambda_2$. The four eigenvalues lie on the unit circle. We have thus *stability*.

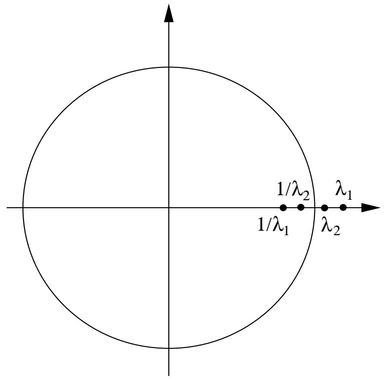
Region 2. $\Delta < 0$ and b_1, b_2 are complex conjugates. The four multipliers are complex and lie outside the unit circle. Moreover $\overline{\lambda_1} = \lambda_2$. This type of instability is called *complex instability*.

Region 3. $\Delta > 0$ and $b_1^2 > 4$, $b_2^2 > 4$. All the eigenvalues are real, but $\lambda_1, 1/\lambda_1$ have signs which are opposite to the signs of $\lambda_2, 1/\lambda_2$. This is known as *even-odd instability*.

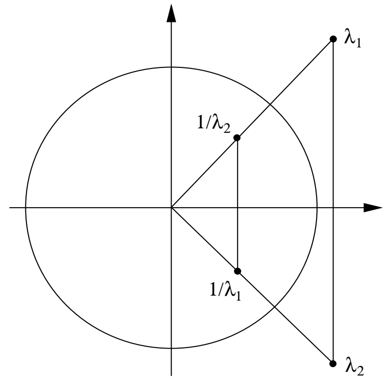
Region 4. $\Delta > 0$ b_1, b_2 are reals, positive and $b_1^2 > 4$, $b_2^2 > 4$. There are four positive eigenvalues. This is the *even-even instability*.

Region 5. $\Delta > 0$. b_1, b_2 reals and negative with $b_1^2 > 4$ and $b_2^2 > 4$. There are four negative eigenvalues. This is the *odd-odd instability*.

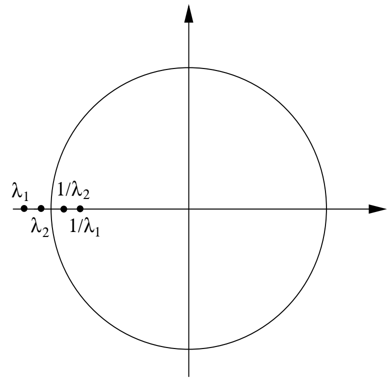
Region 6. $\Delta > 0$. b_1 and b_2 are real with $b_1^2 > 4$ and $b_2^2 < 4$. b_1 is real and positive and $b_1 > 2$ and $\lambda_1, 1/\lambda_1$ are real and positive. The other pair $\lambda_2, 1/\lambda_2$ are complex conjugates and lie on the unit circle. This kind of instability is called *even semi-instability*.



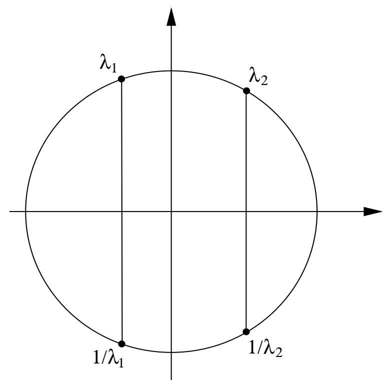
(a) Region four



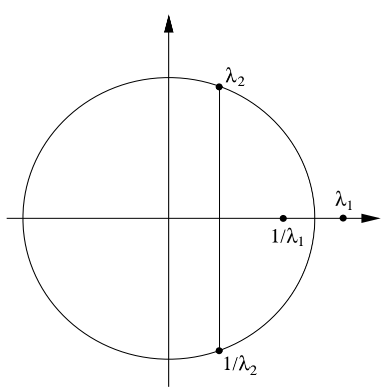
(b) Region two



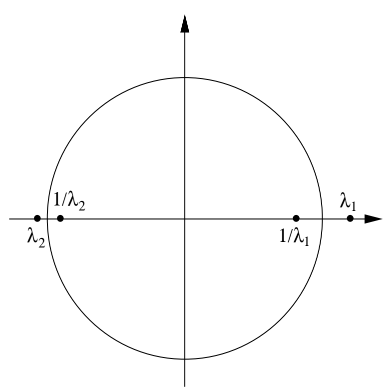
(c) Region five



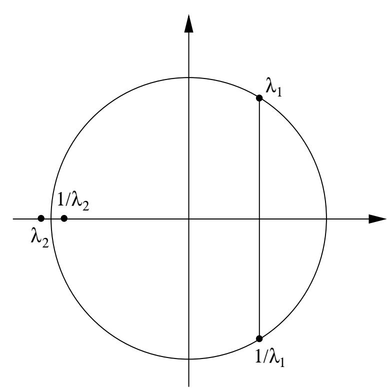
(d) Region one



(e) Region six



(f) Region three



(g) Region seven

Figure B.4: Configuration of the roots with respect to the unit circle on the complex plane for each stability region in the Broucke Diagram.

Region 7. $\Delta > 0$. b_1, b_2 are real with $b_1^2 < 4$ and $b_2^2 > 4$. b_2 is real negative and < -2. The eigenvalues $\lambda_2, 1/\lambda_2$ are real and negative. $\lambda_1, 1/\lambda_1$ are complex conjugates on the unit circle. This type of instability is called the *odd semi-instability*.

Remark B.19. For complex instability, $\Delta < 0$, and the stability indices b_1 and b_2 are complex. Therefore it is advisable (see Pfenniger, 1985a) to use instead

$$c_1 = \frac{1}{2}(b_1 + b_2) = \mathrm{Re}\left(\lambda_1 + \frac{1}{\lambda_1}\right) = \frac{\alpha}{2}, \tag{B.2.13}$$

which measures the even or odd character of the complex instability, and

$$c_2 = \frac{1}{2}|b_1 - b_2| = \left|\mathrm{Im}\left(\lambda_1 + \frac{1}{\lambda_1}\right)\right| = \frac{|\Delta|^{\frac{1}{2}}}{2}, \tag{B.2.14}$$

which measures the degree of complex instability. ♠

B.3 The transformation algorithm

In chapter 1, we shall simplify the Hamiltonian function using changes of coordinates. The key point is that such transformations must preserve the structure of the Hamiltonian equations, so they must be canonical or symplectic transformations as defined in section B.1.

A practical way –from a computational point of view–, for generating canonical transformations is based on the fact that the flow of a Hamiltonian system at a fixed time is a symplectic map (see theorem B.8).

Consider a real (or complex) analytic function G, defined on a domain Ω of $\mathbb{R}^{2n}$ (or $\mathbb{C}^{2n}$) and the one-parameter family of transformations $\phi_t^G : \Omega \to \Omega$, $t \in \mathbb{R}$, verifying:

(i) $\phi_0^G = Id$ (the identity map),

(ii) $\frac{d}{dt}\phi_t^G = J_n \mathrm{grad}\, G \circ \phi_t^G$, for all $t \in \mathbb{R}$.

Then, $\{\phi_t^G\}_{t\in\mathbb{R}}$, is the *one-parameter group of symplectic transformations* generated by the function G. The element corresponding to $t = 1$, ϕ_1^G , is the *symplectic transformation generated time-one flow* (of the Hamiltonian G). Therefore, for any analytic $f : \Omega \to \mathbb{R}$ (or $\mathbb{C}$), one may Taylor-expand at $t = 0$ to obtain a *formal* development of the transformed function $f \circ \phi_1^G$, i. e.,

$$f \circ \phi_1^G = f + \frac{d}{dt}(f \circ \phi_t^G)\Big|_{t=0} + \frac{1}{2!}\,\frac{d^2}{dt^2}(f \circ \phi_t^G)\Big|_{t=0} + \frac{1}{3!}\,\frac{d^3}{dt^3}(f \circ \phi_t^G)\Big|_{t=0} + \cdots, \tag{B.3.1}$$

Now, on the space of all real (or complex) functions defined on Ω, we define the linear operator: $L_G = \{\cdot, G\}$ and, recursively,

$$\begin{aligned} L_G^0 f &= f, \\ L_G^k f &= L_G\left(L_G^{k-1} f\right), \qquad \text{for } k = 1, 2, \cdots, \end{aligned}$$

then, induction shows that,

$$\frac{d^k}{dt^k}\left(f \circ \phi_t^G\right) = \left(L_G^k f\right) \circ \phi_t^G. \tag{B.3.2}$$

Thus, the above expansion of $f \circ \phi_1^G$ can be expressed as next lemma shows.

Lemma B.20. *Under the conditions specified above on the functions f and G, the transformed of the function f through the time-one flow ϕ_1^G $f \circ \phi_1^G$, can be cast into*

$$f \circ \phi_1^G = \sum_{k=0}^{r-1} \frac{1}{k!} L_G^k f + \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} (L_G^r f) \circ \phi_t^G \, dt,$$

for any $r = 1, 2, \ldots$

Proof. If follows from the Taylor expansion (B.3.1), up to order $r-1$, adding the remainder in its integral form and with the time derivatives substituted by (B.3.2). $\square$

Remark B.21. Note that, in particular applying the above lemma to the coordinate functions (i. e., taking $f = x_i$, for $i = 1, \ldots, 2n$) one determines the components of ϕ_1^G itself.♠

The method just described is the most elementary version of the so called *Lie series methods* to generate canonical transformations. These are particularly well-suited for mechanized treatment, since they only require computation of Poisson brackets. The interested reader can find, in Jorba (1999), a practical implementation of a specialized software package able to deal with these questions.

Corollary B.22. *With the same assumptions of lemma B.20, but accept $f = f_0 + f_1$, then*

$$f \circ \phi_1^G = f + \{f - f_1, G\} + \int_0^1 \left(\{f_1, G\} + (1-t) \{\{f - f_1, G\}, G\} \right) \circ \phi_t^G \, dt \qquad \text{(B.3.3)}$$

Proof. Directly from the lemma for $r = 2$, setting $\{f, G\} = \{f - f_1, G\} + \{f_1, G\}$ and taking into account that,

$$\{f_1, G\} = \int_0^1 \frac{d}{dt} \left((t-1) \left(L_G f_1 \right) \circ \phi_t^G \right) \, dt = \int_0^1 \left(L_G f_1 + (t-1) L_G^2 f_1 \right) \circ \phi_t^G \, dt,$$

one arrives to the result of the lemma. $\square$

Formula (B.3.3) will be used in chapter 3 to derive bounds for norm of the the "bad terms" –i. e., those avoiding certain kind of solutions–, in the Lie-transformed Hamiltonians appearing along the iterative steps of the KAM method. The function G is the *generating function* or the *generator* of the transformation. In our context, the function to transform will be an initially given Hamiltonian, $H^{(0)}$. Therefore, one may ask for G such that the transformed new Hamiltonian $H^{(1)} = H^{(0)} \circ \phi_1^G$, will be –tied to some predefined criteria–, simpler than the initial one. This is, essentially, the idea on which the *normal form* and *normalizing transformation* computations are based upon. The aim of the example below is just to illustrate these concepts.

Example B.23. Let $H(\boldsymbol{\xi}, \boldsymbol{\eta})$ be an n degree of freedom real analytic Hamiltonian, so $\boldsymbol{\xi}^* = (\xi_1, \ldots, \xi_n)$ and $\boldsymbol{\eta}^* = (\eta_1, \ldots, \eta_n)$. Furthermore, we suppose that H can be expanded as $H = H_2 + H_3 + \ldots + H_k + \ldots$, where H_k, is an homogeneous polynomial of degree $k > 2$ in the variables $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{C}^{2n}$. So

$$H_k(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sum_{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = k} h_{\boldsymbol{l},\boldsymbol{m}} \, \boldsymbol{\xi}^{\boldsymbol{l}} \boldsymbol{\eta}^{\boldsymbol{m}}, \qquad \text{(B.3.4)}$$

and the following standard notation is used: $\boldsymbol{u}^{\boldsymbol{l}} = \prod_{j=1}^{n} u_j^{l_j}$, $\boldsymbol{u} \in \mathbb{C}^n$, while $|\cdot|_1$ denotes the norm $|\boldsymbol{r}|_1 = |r_1| + \cdots + |r_n|$.

From the development (B.3.4), it follows that $(\boldsymbol{\xi}, \boldsymbol{\eta}) = (\mathbf{0}, \mathbf{0})$ is an equilibrium point of the Hamiltonian system (since grad $H(\mathbf{0}, \mathbf{0}) = \mathbf{0}$). In order to simplify, we shall suppose that this equilibrium point is non-degenerate *elliptic*; i. e., that $\operatorname{Spec}(J\, D^2 H(\mathbf{0}, \mathbf{0})) = \{\mathrm{i}\omega_1, \ldots, \mathrm{i}\omega_n\}$, with $w_j \in \mathbb{R}$, $\omega_j \neq \omega_k$ for $j \neq k$ and $\mathrm{i} = \sqrt{-1}$. In such cases (see Arnol'd, 1974), by means of a real linear canonical transformation, $\binom{\boldsymbol{x}}{\boldsymbol{y}} = S\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$ with $S^* J S = J$, the initial Hamiltonian may be transformed to its real linear normal form,

$$H'(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2} \sum_{j=1}^{n} \omega_j \left(x_j^2 + y_j^2 \right) + H_3'(\boldsymbol{x}, \boldsymbol{y}) + H_4'(\boldsymbol{x}, \boldsymbol{y}) + \ldots \tag{B.3.5}$$

Moreover, to get simpler *homological equations* (see below), it is useful to introduce the following (complex) linear symplectic change,

$$x_j = \frac{q_j + i p_j}{\sqrt{2}}, \qquad y_j = \frac{i q_j + p_j}{\sqrt{2}}, \tag{B.3.6}$$

$j = 1, \ldots, n$. With this complex change, the Hamiltonian (B.3.5) transforms to,

$$H^{(0)}(\boldsymbol{q}, \boldsymbol{p}) = \sum_{j=1}^{n} i \omega_j q_j p_j + \sum_{k>2} H_k^{(0)}(\boldsymbol{q}, \boldsymbol{p}), \tag{B.3.7}$$

where $H_k^{(0)}(\boldsymbol{q}, \boldsymbol{p})$, as in (B.3.4) are homogeneous polynomials of degree k in $\boldsymbol{q}, \boldsymbol{p}$; so, as before, we write:

$$H_k^{(0)}(\boldsymbol{q}, \boldsymbol{p}) = \sum_{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = k} h_{\boldsymbol{l},\boldsymbol{m}}^{(0)}\, \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}, \tag{B.3.8}$$

and $H_2^{(0)}$ will stand for the quadratic part of $H^{(0)}$, i. e.,

$$H_2^{(0)}(\boldsymbol{q}, \boldsymbol{p}) = \sum_{j=1}^{n} i \omega_j\, q_j p_j.$$

The Hamiltonian (B.3.7) is known as the *complexified Hamiltonian*. An explicit description of the linear normalization process which leads to $H^{(0)}$ can be found in Siegel and Moser (1971), chap. 2, § 15.

The inverse of the change (B.3.6) is

$$q_j = \frac{x_j - i y_j}{\sqrt{2}}, \quad p_j = \frac{y_j - i x_j}{\sqrt{2}} \tag{B.3.9}$$

so, when x_j, y_j are real, the complex conjugates of the complex positions $\boldsymbol{q}$ and momenta $\boldsymbol{p}$ satisfy, $\overline{q}_j = -i p_j$, $\overline{p}_j = -i q_j$ for $j = 1, \ldots, n$. This induces the following symmetries on the coefficients in (B.3.8),

$$\overline{h}_{\boldsymbol{l},\boldsymbol{m}}^{(0)} = i^{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1} h_{\boldsymbol{m},\boldsymbol{l}}^{(0)}. \tag{B.3.10}$$

Remark B.24. Before continuing with the nonlinear normalization of the Hamiltonian, it is worth mentioning two essential properties of the Poisson brackets.

P1. If f and g are homogeneous polynomials of degrees r and s respectively, the degree of their Poisson bracket is $\deg\{f,g\} = r+s-2$, and

P2. if the expansions of f and g satisfy the symmetry (B.3.10), so does $\{f,g\}$.

By this last property, it is assured that, after the nonlinear reduction process, the change in (B.3.9) will transform the final Hamiltonian into a real analytic one. ♠

In principle, it is possible to "remove" (in the sense we specify below) the monomials on $H_3^{(0)}$ by taking a generating function $G_3(\boldsymbol{q},\boldsymbol{p}) = \sum g_{\boldsymbol{l},\boldsymbol{m}}\boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}$, with the coefficients $g_{\boldsymbol{l},\boldsymbol{m}}$ ($|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3$), suitably chosen. First, by (B.20) the transformed Hamiltonian $H^{(1)} = H^{(0)} \circ \phi_1^{G_3}$ will be given by

$$H^{(1)} = H_2^{(0)} + \{H_2^{(0)}, G_3\} + H_3^{(0)} + \dots \tag{B.3.11}$$

Note that, by the first of the properties on the remark B.24, the dots hold terms of degree greater than 3, i. e., there are no more terms of degree 3 than the sum: $\{H_2^{(0)}, G_3\} + H_3^{(0)}$. Ideally, we ask $H^{(1)}$ not to contain terms of degree 3 (in this sense we want to *remove* these terms). Hence, G_3 should satisfy the following *homological equations*:

$$\{H_2^{(0)}, G_3\} + H_3^{(0)} = 0.$$

Let $\mathfrak{A}_k$ be the space of (complex) homogeneous polynomials of degree k in the variables $(\boldsymbol{q},\boldsymbol{p}) \in \mathbb{C}^{2n}$ and define the operator $L_{H_2^{(0)}} = \{\cdot, H_2^{(0)}\}$. More precisely,

$$\begin{array}{rcl} L_{H_2^{(0)}} : \mathfrak{A}_k & \to & \mathfrak{A}_k \\ f & \mapsto & L_{H_2^{(0)}} f = \{f, H_2^{(0)}\}, \end{array}$$

(we abbreviate $L_{H_2^{(0)}} = L$ in the text).

Remark B.25. Note that, by (B.1) $L_{H^{(0)}} f = \mathcal{L}_{X_{H^{(0)}}} f$, where $X_{H^{(0)}}$ is the Hamiltonian vector field associated to the function $H^{(0)}$. ♠

With this operator, the homological equations above, may be written as,

$$L_{H_2^{(0)}} G_3 = H_3^{(0)}$$

but for this equations to be compatible, it is necessary that $H_3^{(0)} \in \mathrm{Range}\,(L)$. This does not happen, in general, due to the presence of *resonant monomials* in $H_3^{(0)}$. A monomial $f = a_{\boldsymbol{l},\boldsymbol{m}}\boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}$, $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = k$, is said to be *resonant*, if $Lf = 0$ (equivalently, if $f \in \mathrm{Ker}\,(L)$). It is thus necessary to add a compatibility term, $Z_3 \in \mathrm{Ker}(L)$, satisfying $H_3^{(0)} - Z_3 \in \mathrm{Range}\,(L)$. Therefore, the homological equations (B.23) for the degree $k = 3$; must be completed in the form,

$$L_{H_2^{(0)}} G_3 + Z_3 = H_3^{(0)}. \tag{B.3.12}$$

We define the resonance modulus associated to $\boldsymbol{\omega}^* = (\omega_1, \dots, \omega_n)$, the vector of the frequencies, as $\mathfrak{R} = \{\boldsymbol{r} \in \mathbb{Z}^n : \langle \boldsymbol{r}, \boldsymbol{\omega} \rangle = 0\}$ (the angular brackets $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j=1}^n u_j v_j$, will be used for the ordinary scalar product). Thus, it becomes clear how to express the compatibility term Z_3,

$$Z_3(\boldsymbol{q},\boldsymbol{p}) = \sum_{\substack{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=3\\ \boldsymbol{m}-\boldsymbol{l}\in\mathfrak{R}}} h^{(0)}_{\boldsymbol{l},\boldsymbol{m}}\, \boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}, \tag{B.3.13}$$

When the homological equations (B.3.12) are written down explicitly, using the developments of G_3, $H_3^{(0)}$ and Z_3, and computing the Poisson bracket, on can realize that

$$\sum_{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=3} \mathrm{i}\langle\boldsymbol{\omega},\boldsymbol{m}-\boldsymbol{l}\rangle g_{\boldsymbol{l},\boldsymbol{m}}\, \boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}} = \sum_{\substack{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=3\\ \boldsymbol{m}-\boldsymbol{l}\notin\mathfrak{R}}} -h^{(0)}_{\boldsymbol{l},\boldsymbol{m}}\, \boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}.$$

This gives, for the unknowns $g_{\boldsymbol{l},\boldsymbol{m}}$, an algebraic linear diagonal system in the space $\mathbb{C}^d \simeq \mathfrak{A}_3$, $d = \binom{2n+2}{3}$; so, easily, one obtains

$$g_{\boldsymbol{l},\boldsymbol{m}} = \frac{-h^{(0)}_{\boldsymbol{l},\boldsymbol{m}}}{\mathrm{i}\langle\boldsymbol{\omega},\boldsymbol{m}-\boldsymbol{l}\rangle}; \tag{B.3.14}$$

with $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3$ and indeed, $\boldsymbol{l}-\boldsymbol{m} \notin \mathfrak{R}$, so the divisors do not vanish. With this choice for the coefficients in G_3, the transformed Hamiltonian, does not hold three degree terms, but the resonant ones, contained in Z_3. Therefore,

$$H^{(1)} = H^{(0)} \circ \phi_1^{G_3} = H_2 + Z_3 + H_4^{(1)} + \dots, \tag{B.3.15}$$

where $H_4^{(1)} + \dots$ denotes the transformed terms of degree greater than three.

The general step

In the previous paragraphs, we have given an account of the linear reduction and some details of the first step in the nonlinear –or in the "normal form"–, reduction process of the initial Hamiltonian. Suppose now that the same has been repeated up to degree $k > 3$. So at the k-th step we have a Hamiltonian $H^{(k)} = H^{(0)} \circ \phi_1^{G_3} \circ \cdots \circ \phi_1^{G_{k+2}}$,

$$H^{(k)} = H_2^{(0)} + Z_3 + \ldots + Z_{k+2} + \sum_{s>k+2} H_s^{(k)}, \tag{B.3.16}$$

with $H_s^{(k)}$ homogeneous polynomials of degree $s > k$, in $(\boldsymbol{q},\boldsymbol{p})$,

$$H_s^{(k)}(\boldsymbol{q},\boldsymbol{p}) = \sum_{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=s} h^{(k)}_{\boldsymbol{l},\boldsymbol{m}}\, \boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}.$$

In the $(k+1)$-th step we want to remove the terms of degree $k+3$ of $H^{(k)}$. We apply the change $\phi_1^{G_{k+3}}$ to obtain a new Hamiltonian, $H^{(k+1)} = H^{(k)} \circ \phi_1^{G_{k+3}}$. By the Lie transform formula (B.20), and writing only up to degree $k+3$,

$$H^{(k+1)} = H_2^{(0)} + Z_3 + \cdots + Z_{k+3} + \{H_2^{(0)}, G_{k+3}\} + H_{k+3}^{(k)} + \dots,$$

for, again –with (B.20) in mind–, it is just a check on the degrees, to realize that the only terms of degree $k+3$ present in this new Hamiltonian are those in the sum $\{H_2^{(0)}, G_{k+3}\}+H_{k+3}^{(k)}$. In consequence, the homological equations will be,

$$L_{H_2^{(0)}} G_{k+3} + Z_{k+3} = H_{k+3}^{(k)}, \tag{B.3.17}$$

with, in the same way as for $k=3$, taking Z_k an homogeneous polynomial holding the resonant terms of $H_{k+3}^{(k)}$,

$$Z_{k+3}(\boldsymbol{q},\boldsymbol{p}) = \sum_{\substack{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=k+3\\ \boldsymbol{m}-\boldsymbol{l}\in\mathfrak{R}}} h_{\boldsymbol{l},\boldsymbol{m}}^{(k)}\, \boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}, \tag{B.3.18}$$

The solutions of (B.3.17) will have, of course, the same form of (B.3.14),

$$g_{l,m} = \frac{-h_{\boldsymbol{l},\boldsymbol{m}}^{(k)}}{\mathrm{i}\langle\boldsymbol{\omega},\boldsymbol{m}-\boldsymbol{l}\rangle}; \tag{B.3.19}$$

now, with $|\boldsymbol{l}|_1+|\boldsymbol{m}|_1 = k+3$, but identically: $\boldsymbol{l}-\boldsymbol{m}\notin\mathfrak{R}$. Therefore, if we make r steps of the nonlinear reduction process, the resulting Hamiltonian splits into

$$H^{(r)} = Z^{(r)} + R^{(r)}, \tag{B.3.20}$$

where the *remainder* $R^{(r)}$ contains terms of degree $> r+2$, while

$$Z^{(r)} = Z_2 + Z_3 + \ldots + Z_{r+2},$$

(with $Z_2 = H_2^{(0)}$), is the (complex) *normal form* up to order $r+2$ of the Hamiltonian. It holds the quadratic part plus the resonant terms $Z_3,\ldots,Z_{r+2}$ from degree 3 up to $r+2$. Nevertheless, sometimes, the term *normal form* applies, by extension, to the whole transformed Hamiltonian (B.3.20).

Remark B.26. From (B.3.14) it follows that the generating functions coefficients $g_{\boldsymbol{l},\boldsymbol{m}}$, $|\boldsymbol{l}|_1+|\boldsymbol{m}|_1 = 3$, satisfy the symmetries (B.3.10). By Induction and taking into account the second property of remark (B.24), it is possible to see that, at any degree $3\le k\le r+2$, $g_{\boldsymbol{l},\boldsymbol{m}} = i^{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1} g_{\boldsymbol{m},\boldsymbol{l}}$. Hence, the same is true for the coefficients of $H^{(r)}$. This implies that we can obtain a *real* normalized Hamiltonian, applying to (B.3.20) the change (B.3.9). ♠

Even in the simplest case, when $\mathfrak{R}=\emptyset$, there appear *inevitable* resonances when $\boldsymbol{l}=\boldsymbol{m}$. Then, the terms in $Z^{(r)}$ take the form,

$$\begin{array}{rcl} Z_{2s}(\boldsymbol{q},\boldsymbol{p}) & = & \sum_{|\boldsymbol{l}|_1=s} h_{\boldsymbol{l},\boldsymbol{l}}^{(2s-3)}\boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{l}}, \\ & & \\ Z_{2s-1} & \equiv & 0, \end{array} \qquad 2\le s\le \lfloor r/2\rfloor+1,$$

where $\lfloor x\rfloor$ denotes the greatest integer function of $x\in\mathbb{R}$ (i. e, $\lfloor x\rfloor := \max\{z\in\mathbb{Z} : z\le x\}$). If –as pointed in last remark–, we apply first the linear symplectic change (B.3.9), and then introduce polar canonical coordinates in the form,

$$x_j = \sqrt{2I_j}\cos\theta_j, \qquad y_j = -\sqrt{2I_j}\sin\theta_j \tag{B.3.21}$$

($j = 1, \dots, n$), we obtain a real normalized Hamiltonian,

$$\widetilde{H}^{(r)}(\boldsymbol{I}, \boldsymbol{\theta}) = \sum_{j=1}^{n} \omega_j I_j + \sum_{j=1}^{\lfloor r/2 \rfloor + 1} \widetilde{Z}_{2j}(\boldsymbol{I}) + \widetilde{R}^{(r)}(\boldsymbol{I}, \boldsymbol{\theta}),$$

with $\boldsymbol{I} = (I_1, \dots, I_n)$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$, and $\widetilde{Z}_{2j}(\boldsymbol{I})$ are real homogeneous polynomials of degree j in $I_1, \dots, I_n$.

An interesting point is that the normal form, i. e., the first two sums in the expression above,

$$\widetilde{Z}^{(r)} = \sum_{j=1}^{n} \omega_j I_j + \sum_{j=1}^{[r/2]+1} \widetilde{Z}_{2j}, \tag{B.3.22}$$

does not depend on the angular variables θ_j, so if we skip the remainder off and consider the Hamiltonian system given by (B.3.22),

$$\dot{\theta}_j = \frac{\partial \widetilde{Z}^{(r)}}{\partial I_j}, \qquad \dot{I}_j = 0, \tag{B.3.23}$$

for $j = 1. \dots, n$, this system is immediately integrable, with solutions:

$$I_j = I_j^0 = \text{contant}, \quad \theta_j = \Omega_j(\boldsymbol{I}^0)\, t + \theta_j^0, \tag{B.3.24}$$

with the frequencies, $\Omega_j = \frac{\partial \widetilde{Z}^{(r)}}{\partial I_j} = \omega_j + \dots ; j = 1, \dots, n$. Therefore, the trajectory of the phase point $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ winds an n dimensional invariant torus defined by the first integrals $I_j = I_j^0 = \frac{1}{2}(x_j^2 + y_j^2)$. $\diamond$

Normal forms around equilibrium points were studied by Birkhoff (1927). In fact, in the texts, the normal form (B.3.20) is known as the Birkhoff's normal form. In the description left here, it has been obtained applying successive canonical changes, each of them constructed to remove the terms one degree higher than the preceding one. The final transformation is a product of the r transformations: $\Psi^{(r)} = \phi_1^{G_3} \circ \phi_1^{G_4} \circ \dots \circ \phi_1^{G_{r+2}}$.

There is, however, a vast literature for *Lie transformation algorithms*. See Deprit (1969) and the extensions of Kamel (1970); Henrard (1970a,b,c) for non Hamiltonian systems; or also the books of Chow and Hale (1982), chapter 12 and Meyer and Hall (1992) chapter 7.

Without digging deeper into the details: given a generating function G, which can be expanded as a sum $G = \sum_{k \geq 3} G_k$, there are algorithms which allows us to construct a canonical transformation T_G, such that if f is a function of $(\boldsymbol{q}, \boldsymbol{p})$, the transformed function $T_G f$ will be defined by

$$T_G f = \sum_{k \geq 1} F_k,$$

where the terms f_k are obtained recursively, that is, beginning with $F_1 = f_1$, for $k > 1$ is $F_k = F_k(G_1, \dots, G_{k+1}; f_0, \dots, f_k)$, so each term in the sum can be obtained from the generating function and the previous computed terms.

In particular in chapters 1 and 2 we shall use an algorithm of this type: the Giorgilli–Galgani algorithm, (see references there). With this short outline of the transformation theory, we close this appendix.

Bibliography

Abraham, R., and Marsden, J. E.: 1978, *Foundations of Mechanics*. 2nd edition (1978). Benjamin/Cummings, Reading, Massachussetts.

Abraham, R., Marsden, J. E., and Ratiu, T.: 1983, *Manifolds, Tensor Analysis and Applications*. 2nd edition (1988). Springer-Verlag, Berlin, Heidelberg, New York.

Andronov, A. A., Khaikin, S. E., and Witt, A. A.: 1959, *Theory of Oscillators*. Dover Publications, Inc. New York, 1987, unaltered republication of the edition published by Pergamon Press, Oxford, in 1966 as Volume Four of *The International Series of Monographs in Physics*.

Arnol'd, V. I.: 1963, Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. Russian Math. Surveys **18**(5), 9–36.

Arnol'd, V. I.: 1963, Small denominators and problems of stability of motion in classical and celestial mechanics. Russian Math. Surveys **18**(6), 85–191.

Arnol'd, V. I.: 1974, *Méthodes Mahtématiques de la Mécanique Classique*. Traduction française Editions Mir (1976). Mir, Moscou.

Arnol'd, V. I.: 1978, *Chapitres supplémentaires de la théorie des équations différentielles ordinaires*. Traduction française Editions Mir 1980. Réimpression (1984). Mir, Moscou.

Arnol'd, V. I. (ed.): 1988, *Dynamical Systems III*. Encyclopaedia of Mathematical Sciences, Vol. III. Springer-Verlag Berlin, Heidelberg.

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J. M.: 1984, A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. Nuovo Cimento B **79**(2), 201–203.

Birkhoff, G. D.: 1927, *Dynamical Systems*. 2nd edition (1966). American Mathematical Society, Providence, Rhode Island. Freely available from the web site of the American Mathematical Society, $\mathcal{AMS}$, `http://www.ams.org/online_bks/coll9`

Bourgain J.: 1997, On Melnikov's persistency problem. Math. Res. Lett. **4**, 445–458.

Bridges, T. J.: 1990, Bifurcation of periodic solutions near a collision of eigenvalues of opposite signature. Math. Proc. Cambridge Philos. Soc. **108**, No. 3, 575–601.

Bridges, T. J.: 1991, Stability of periodic solutions near a collision of eigenvalues of opposite signature. Math. Proc. Cambridge Philos. Soc. **109**, No. 2, 375–403.

Bridges, T. J. and Furter, J. E.: 1993, *Singularity Theory and Equivariant Symplectic maps*. Lecture Notes in Math. 1558. Springer-Verlag Berlin, New York.

Bridges, T. J., and Cushman, R. H.: 1993, Unipotent normal forms for symplectic maps. Physica D **65**, No. 3, 211–241.

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.

Broer, H., Huitema, G. B., and Sevryuk, M. B.: 1996, *Quasi-Periodic Motions in Families of Dynamical Systems: Order Amidst Chaos*, Lecture Notes in Math. **1645**. Springer-Verlag Berlin, Heidelberg, New York, etc.

Broucke, R.: 1969, Stability of periodic orbits in the elliptic restricted three body problem. AIAA Journal **7**, No. 3, 1003–1009.

de Bruijn, N. G.: 1958, *Asymptotic Methods in Analysis*. North Holland Publishing Co., Amsterdam.

Bruno, A. D.: 1989, Normalization of a Hamiltonian system near a cycle or a torus. Russian Math. Surveys **44** No. 2, 53–89.

Bruno, A. D.: 1994, *The restricted 3-body problem: plane periodic orbits*. Walter de Gruyter & Co., Berlin.

Burgoyne, N., and Cushman, R.: 1974, Normal forms for real linear Hamiltonian systems with purely imaginary eigenvalues. Celestial Mech. **8**, 435–443.

Chow, S.-N., and Hale, K.: 1982, *Methods of Bifurcation Theory*. 2nd edition (1996). Springer-Verlag, Berlin, Heidelberg, New York.

Contopoulos, G., and Giorgilli, A.: 1988, Bifurcations and complex instability in a 4-dimensional symplectic mapping. Meccanica **23**, No. 1, 19–28.

Contopoulos, G., and Barbanis, B.: 1994, Periodic orbits and their bifurcations in a 3-D system. Celestial Mech. and Dynam. Astronom. **59**, No. 3, 279–300.

Corless, R. M., Gonnet, G. H., Hare, D. E., Jeffrey, D. J., and Knuth, D. E.: 1996, On the Lambert W function. Advances Computational Maths **5**, 329–359.

Cushman, R., and Van der Meer, J.-C.: 1990, The Hamiltonian Hopf bifurcation in the Lagrange top. In: *Differential Geometric Methods in Mathematical Physics*, H. D. Doebner and D. J. Henning (eds.), Lecture Notes in Math. **1416**. Springer, Berlin.

Delshams, A.: 1994, *Sistemes Hamiltonians i Teoria de Pertorbacions*. Postgraduate course, teaching and class notes. Manuscript.

Delshams, A., Gutiérrez, P.: 1996, Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system. J. Differential Equations 131, no. 2, 277–303.

Deprit, A.: 1969, Canonical transformations depending on a small parameter. Celestial Mech. **1**, 12-30.

Elphick, C., Tirapegui, E., Brachet, M. E., Coullet, P., and Ioss, G.: 1987, A simple global characterization for normal forms of singular vector fields. Physica **29**D, 95–127.

Fröberg, C.-E,: 1973, *Introduction to Numerical Analysis*. Addison-Wesley, Reading, Massachussets.

Froeschlé, C.: 1971, On the number of isolating integrals in systems with three degrees of freedom. Astrophys. Space Sci. **14**, 110–117.

Froeschlé, C.: 1972, Numerical study of a four-dimensional mapping. Astronom. and Astrophys. **16**, 172–189.

Giorgilli, A., and Galgani, L.: 1978, Formal integrals for an autonomous Hamiltonian system near an equilibrium point. Celestial Mech. **17**, 267–280.

Giorgilli, A., and Galgani, L.: 1985, Rigurous estimates for the series expansions of Hamiltonian perturbation theory. Celestial Mech. **37**, 95–112.

Giorgilli, A., Delshams, A., Fontich, E., Galgani, L., and Simó, C.: 1989, Effective Stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem. Journal of Differential Equations **77**, 167–198.

Goldstein, H.: 1980, *Classical Mechanics*. Addison-Wesley. Reading, Massachussets.

Gómez, G., Jorba, A., Masdemont, J., and Simó, C.: 1993, Study of Poincaré Maps for Orbits Near Lagrangian Points. ESOC contract 9711/91/D/IM(SC), Final Report.

Hadjidemetriou, J. D.: 1985, Stability of resonant orbits in planetary systems. In: *Resonances in the Motion of Planets, Satellites and Asteroids*, S. Ferraz-Mello and W. Sessin (eds.), Univ. São Paulo, 1–26.

Haller, G.: 1999, *Chaos near resonance*. Springer-Verlag New York, Berlin, Heidelberg.

Heggie, D.: 1985, Bifurcation at complex instability. Celestial Mech. **35**, 357–382.

Henrard, J.: 1970, Periodic orbits emanating from a resonant equilibrium. Celestial Mech. **1**, 437–466.

Henrard, J.: 1970, On a perturbation theory using Lie transforms. Celestial Mech. **3**, 107–120.

Henrard, J.: 1970, Liapounov's center theorem for resonant equilibrium. Journal of Differential Equations **14**(3), 431–441.

Howard, J. E., Lichtenberg, A. J., Lieberman, M. A., and Cohen, R. H.: 1986, Four-dimensional mapping model for two frequency electron cyclotron resonance heating. Physica D **20**, 259–284.

Jorba, A.: 1999, A Methodology for the numerical computation of normal forms, centre manifolds and first integrals of Hamiltonian systems. Experimental Mathematics, 8(2), 155–195.

Jorba, A., and Villanueva, J.: 1997, On the persistence of lower dimensional invariant tori under quasi-periodic perturbations. J. Nonlinear Sci., **7**, 427–473.

Jorba, A., and Villanueva, J.: 1997, On the normal behavior of partially elliptic lower-dimensional tori of Hamiltonian systems. Nonlinearity, **10**, 783–822.

Jorba, A., and Villanueva, J.: 1998, Numerical computation of normal forms around some periodic orbits of the restricted three-body problem. Physica D **114**, 197–229.

Kamel, A.: 1970, Perturbation methods in the theory of nonlinear oscillations. Celestial Mech. **3**, 90–106.

Kolmogorov, A. N.: 1979, Preservation of conditionally periodic movements with small change in the Hamiltonian function. In *Stochastic Behavior in Classical and Quantum Hamiltonian Systems* (G. Casati and J. Ford, Eds.), Lecture Notes in Physics, **93**, 51–56. Springer-Verlag, Berlin, New York.

Lochack, P., and Meunier, C.: 1988, *Multiphase Averaging for Classical Systems.* Springer-Verlag, New York.

Marsden, J. E., and Ratiu, S. T.: 1999, *Introduction to Mechanics and Symmetry.* 2nd edition (1999). Springer-Verlag, Berlin, Heidelberg, New York.

Martinet, L.: 1984, 1/1 Resonances in barred galaxies. Astronom. and Astrophys. **134**, 373–386.

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

Meyer, K. R., and Hall, G. H.: 1992, *Introduction to Hamiltonian Dynamical Systems and the N-body Problem.* Springer-Verlag, Berlin, Heidelberg, New York.

Meyer, K. R.: 1998, The evolution of the stable and unstable manifold of an equilibrium point. Celestial Mech. and Dynam. Astronom. **70**, 159–165.

Moser, J.: 1967, Convergent series expansions for quasiperiodic motions. Math. Ann. **169**(1), 136–176.

Nekhoroshev, N. N.: 1977, An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems I. Russian Math. Surveys **32**, No. 6, 1–65.

Nekhoroshev, N. N.: 1979, An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems II. Trudy. Sem. Petrovsk. **5**, 5–50. Topics in Modern Mathematics. Petrovskii seminar. No. 5, O. A. Oleinik Ed. Consultants Boreau (Plenum), New York, London (1985), 1–58.

Ollé, M., and Pfenniger, D.: 1998, Vertical orbital structure around the Lagrangian points in barred galaxies. Link with the secular evolution of galaxies. Astronom. and Astrophys. **334**, 829–839.

Ollé, M., and Pfenniger, D.: 1999, Bifurcation at complex instability. In: *Hamiltonian systems with three or more degrees of freedom*, C. Simó (ed.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **533**, Kluwer Acad. Publ., Dordrecht, 518–522.

Ollé, M., and Pacha, J. R.: 1999, The 3D elliptic restricted three-body problem: periodic orbits which bifurcate from limiting restricted problems. Complex instability. Astronom. and Astrophys. **351**, 1149–1164.

Palis, J., and Melo, W.: 1982, *Geometric Theory of Dynamical Systems.* (1988) Collæge Press, Beijing.

Perry, A. D., and Wiggins, S.: 1994, KAM tori are very sicky: rigorous lower bounds on the time to move away from an invariant Lagrangian torus with linear flow. Physica D **71**, 102–121.

Pfenniger, D.: 1985a, Numerical study of complex instability I. Mappings. Astronom. and Astrophys. **150**, 97–111.

Pfenniger, D.: 1985b, Numerical study of complex instability II. Barred galaxy bulges. Astronom. and Astrophys. **150**, 112–128.

Pfenniger, D.: 1990, Stability of the Lagrangian points in stellar bars. Astronom. and Astrophys. **230**, 55–66.

Pöschel, J.: 1989, On elliptic lower dimensional tori in Hamiltonian systems. Math. Z. **202**(4), 559–608.

Puig, J.: 2001, Reducibility of linear equations with quasi-periodic coefficients. A survey. Freely available at the Mathematical Physics Preprint Archive of the University of Texas-Austin, `http://www.ma.utexas.edu/mp_arc-bin/mpa?yn=02-246`

Puig Adam, P.: 1939, Demostración simplificada de la fórmula de de Moivre-Stirling y acotación geométrica del error. Revista de Matemática Hispano-Americana. Serie 3ª, tomo I. Madrid.

Rüssmann, H.: 1975, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus, in: J. Moser (ed.), *Dynamical systems, theory and applications*, LNP **38**, 598–624.

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**, No. 1-3, 155–176.

Siegel, C. L., and Moser, J. K.: 1971, *Lectures on Celestial Mechanics*. Springer-Verlag, Berlin, Heidelberg, New York.

Simó, C.: 1989, Estabilitat de sistemes hamiltonians. Memorias de la Real Academia de Ciencias y Artes de Barcelona, **48**. 303-348.

Smale, S., and Hirsch, M.: 1974, *Differential Equations, Dynamical Systems and Linear Albgebra*. Academic Press, New York and London.

Sotomayor, J.: 1979, *Llições de equações Diferenciais Ordinárias*. IMPA-CNPq, Rio de Janeiro.

Sevryuk, M. B.: 1997, Excitation of elliptic normal modes of invariant tori in Hamiltonian systems. Topics in singularity theory, 209–218, Amer. Math. Soc. Transl. Ser. 2, 180, Amer. Math. Soc., Providence, RI.

Sevryuk, M. B.: 1999, The lack-of-parameters problem in the KAM theory revisited, in *Hamiltonian Systems with Three or More Degrees of Freedom*, C. Simó (ed.), Hamiltonian Systems with Three or More Degrees of Freedom (3$\mathcal{DHAM}$95), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., Vol. 533 (1999), pp. 568-577, Kluwer, Dordrecht, Holland.

Szebehely, V.: 1967, *Theory of Orbits*. Academic Press, New York and London.

Van der Meer, J.-C.: 1985, *The Hamiltonian Hopf Bifurcation*. Lecture Notes in Math. 1160. Springer-Verlag Berlin, New York.

Whitney, H.: 1934, Analytic extensions of differentiable functions defined in closed sets. Trans. Amer. Math. Soc. **36**, 63–89.

Wiggins, S., and Rudnev, M.: 1997, KAM theory near multiplicity one resonant surfaces in perturbations of a a-priori stable Hamiltonian systems. Nonlinear Science **7**(2), 177–209.

Williamson, J.: 1937, On the normal forms of linear canonical transformations in dynamics. Amer. J. Math. **59**, 599–617.

Zehnder, E.: 1975, Generalized implicit function theorems with applications to some small divisor problems I. Comm. Pure Appl. Math. **28**, 91–140.