An introduction to Birkhoff normal form

Dario Bambusi*
Dipartimento di Matematica, Universitá di Milano
via Saldini 50, 20133 Milano (Italy)

19.11.14

1 Introduction

The aim of this note is to present an introduction to Birkhoff normal form and to its use for the study of the dynamics of a Hamiltonian system close to an elliptic equilibrium point. Recall that Birkhoff's theorem ensures the existence of a canonical transformation putting a Hamiltonian system in normal form up to a remainder of a given order. The dynamics of the system in normal form depends on the resonance relations fulfilled by the frequencies. In the nonresonant case such a system is integrable and therefore its dynamics is very simple. In the resonant case new phenomena occur. Typically there is exchange of energy among the oscillators (beats). The phenomenon will be illustrated by studying a resonant system with two degrees of freedom.

2 Birkhoff Theorem

On the phase space $\mathbb{R}^{2n}$ consider a smooth Hamiltonian system H having an equilibrium point at zero.

Definition 2.1. The equilibrium point is said to be *elliptic* if there exists a canonical system of coordinates (p,q) (possibly defined only in a neighborhood of the origin) in which the Hamiltonian takes the form

$$H(p,q) := H_0(p,q) + P(p,q) \ , \tag{2.1}$$

where

$$H_0(p,q) = \sum_{l=1}^{n} \omega_l \frac{p_l^2 + q_l^2}{2} \ , \quad \omega_l \in \mathbb{R} \tag{2.2}$$

and P is a smooth function having a zero of order 3 at the origin.

*E-mail: bambusi@mat.unimi.it

Remark 2.2. The equations of motion of (2.1) take the form

$$\dot{p}_l = -\omega_l q_l - \frac{\partial P}{\partial q_l} \tag{2.3}$$

$$\dot{q}_l = \omega_l p_l + \frac{\partial P}{\partial p_l} \tag{2.4}$$

Since P has a zero of order three, its gradient starts with quadratic terms. Thus, in the linear approximation the equations (2.3,2.4) take the form

$$\begin{array}{c} \dot{p}_l = -\omega_l q_l \\ \dot{q}_l = \omega_l p_l \end{array} \Longrightarrow \ddot{q}_l + \omega_l^2 q_l = 0 \tag{2.5}$$

namely the system consists of n independent harmonic oscillators.

Example 2.3. Consider a Hamiltonian system of the form

$$\sum_{l=1}^{n} \frac{y_l^2}{2} + V(x)$$

where (y, x) are canonical variables. Assume that V has a minimum at the origin. Then by standard theory there exist canonical variables (normal modes) $(\tilde{p}, \tilde{q})$ in which the Hamiltonian takes the form

$$\sum_{l=1}^{n} \frac{\tilde{p}_l^2 + \omega_l^2 \tilde{q}_l^2}{2} + \tilde{P}(\tilde{q})$$

Then the canonical change of variables

$$p_l = \frac{\tilde{p}_l}{\sqrt{\omega_l}} \ , \quad q_l = \tilde{q}_l \sqrt{\omega_l}$$

reduces the Hamiltonian to the form (2.1).

In the case of example 2.3 the frequencies ω_l are positive numbers, while this not required in the general definition 2.1. A relevant example in which negative frequencies appear is that of the Lagrangian equilibrium of the circular three body problem.

In the following we will denote by $x = (p, q)$ the whole set of variables.

Theorem 2.4. *(Birkhoff) For any positive integer $N \geq 0$, there exist a neighborhood $\mathcal{U}_N$ of the origin and a canonical transformation $\mathcal{T}_N : \mathbb{R}^{2n} \supset \mathcal{U}_N \to \mathbb{R}^{2n}$ which puts the system (2.1) in Birkhoff Normal Form up to order N, namely such that*

$$H^{(N)} := H \circ \mathcal{T}_N = H_0 + Z^{(N)} + \mathcal{R}^{(N)} \tag{2.6}$$

where $Z^{(N)}$ is a polynomial of degree $N + 2$ which Poisson commutes with H_0, namely $\left\{H_0; Z^{(N)}\right\} \equiv 0$ and $\mathcal{R}^{(N)}$ is small, i.e.

$$|\mathcal{R}^{(N)}(x)| \leq C_N \left\|x\right\|^{N+3} \ , \quad \forall x \in \mathcal{U}_N \ ; \tag{2.7}$$

moreover, one has

$$\|x - \mathcal{T}_N(x)\| \leq C_N \|x\|^2 \ , \quad \forall x \in \mathcal{U}_N \ . \tag{2.8}$$

An inequality identical to (2.8) *is fulfilled by the inverse transformation* $\mathcal{T}_N^{-1}$.

If the frequencies are nonresonant namely if

$$\omega \cdot k \neq 0 \ , \quad \forall k \in \mathbb{Z}^n \backslash \{0\} \ , \tag{2.9}$$

I will show that the function $Z^{(N)}$ depends on the actions

$$I_j := \frac{p_j^2 + q_j^2}{2}$$

only.

In the resonant case the normal form is more complicated and new phenomena occur. They will be illustrated in sect. 5.

Remark 2.5. The remainder $\mathcal{R}^{(N)}$ is very small in a small neighborhood of the origin. In particular, it is of order ϵ^{N+3} in a ball of radius ϵ. We will show in sect. 4 that in typical cases $\mathcal{R}^{(N)}$ might have a relevant effect only after a time of order ϵ^{-N}.

3 Proof of Birkhoff's theorem

3.1 Formal theory

The idea is to construct a canonical transformation putting the system in a form which is as simple as possible (as we will see, this is the normal form). More precisely one constructs a canonical transformation pushing the non normalized part of the Hamiltonian to order four, followed by a transformation pushing it to order five and so on. Each of the transformations is constructed as the time one flow of a suitable auxiliary Hamiltonian function (Lie transform method).

Definition 3.1. We will denote by $\mathcal{P}_j$ the set of the homogeneous polynomials of degree $j+2$ on $\mathbb{R}^{2n}$.

Remark 3.2. Let $g \in \mathcal{P}_j$ be a homogeneous polynomial, then there exists a constant C such that

$$|g(x)| \leq C \|x\|^{j+2} \ . \tag{3.1}$$

Denote by X_g the Hamiltonian vector field of g then X_g is a homogeneous polynomial of degree $j+1$ and therefore one has

$$\|X_g(x)\| \leq C' \|x\|^{j+1} \tag{3.2}$$

with a suitable constant C'. The best constant such that (3.2) holds is usually called the norm of X_g and denoted by $\|X_g\|$. Similarly one can define the norm of the polynomial g, but in the following we will not need it.

Remark 3.3. Let $f \in \mathcal{P}_i$ and $g \in \mathcal{P}_j$ then, by the very definition of Poisson Brackets one has $\{f, g\} \in \mathcal{P}_{i+j}$.

3.2 Lie Transform

Let $\chi \in \mathcal{P}_j$ be a polynomial function, consider the corresponding Hamilton equations, namely

$$\dot{x} = X_\chi(x) \ ,$$

and denote by ϕ^t the corresponding flow.

Definition 3.4. The time one map $\phi := \phi^t\big|_{t=1}$ is called *Lie transform generated by* χ. It is clear by construction that ϕ is a canonical transformation.

We are now going to study the way a polynomial transforms when the coordinate are subjected to a Lie transformation. Let $g \in \mathcal{P}_i$ be a polynomial and let ϕ be the Lie transform generated by a polynomial $\chi \in \mathcal{P}_j$ with $j \geq 1$. To compute the Taylor expansion of $g \circ \phi$ first remark that

$$\frac{d}{dt} g \circ \phi^t = \{\chi, g\} \circ \phi^t \tag{3.3}$$

so that, iterating one has

$$\frac{d^l}{dt^l} g \circ \phi^t = \underbrace{\{\chi, ...\{\chi, g\}}_{l \text{ times}} \circ \phi^t \tag{3.4}$$

Thus the Taylor expansion of $g \circ \phi^t$ with respect to the time variable is given by

$$g \circ \phi^t = \sum_{l \geq 0} t^l g_l \tag{3.5}$$

where the g_l's are iteratively defined by

$$g_0 := g \, , \quad g_l = \frac{1}{l} \{\chi; g_{l-1}\} \ , \quad l \geq 1 \ . \tag{3.6}$$

Evaluating at $t = 1$ one gets[1]

$$g \circ \phi = \sum_{l \geq 0} g_l \ . \tag{3.7}$$

Remark 3.5. Eq. (3.7) is the Taylor expansion of $g \circ \phi$ at the origin *as a function of the phase space variables* x. This is an immediate consequence of remark 3.3 which implies

$$g_l \in \mathcal{P}_{i+lj} \tag{3.8}$$

so that the expansion (3.7) is an expansion in homogeneous polynomials of increasing order, i.e. the Taylor expansion.

[1] As a consequences of corollary 3.17 below, provided R is small enough, the map ϕ is analytic for $(t, x) \in (-3/2, 3/2) \times B_R$, where B_R is the ball of radius R centered at the origin. It follows that the series (3.7) converges in B_R.

3.3 Looking for the generating function

We are now ready to construct a canonical transformation normalizing the system up to terms of fourth order. Thus let $\chi_1 \in \mathcal{P}_1$ be the generating of the Lie transform ϕ_1, and consider $H \circ \phi_1$, with H given by (2.1). Using (3.7) and (3.6) to compute the first terms of the Taylor expansion of $H \circ \phi$ one gets

$$H \circ \phi = H_0 + \{\chi_1, H_0\} + P_1 + \text{h.o.t}$$

where P_1 is the Taylor polynomial of order three of P and h.o.t. denotes higher order terms.

We wont to construct χ_1 in such a way that

$$Z_1 := \{\chi_1, H_0\} + P_1 \tag{3.9}$$

turns out to be as simple as possible. Obviously the simplest possible form is $Z_1 = 0$. Thus we begin by studying the equation

$$\{\chi_1, H_0\} + P_1 = 0 \tag{3.10}$$

for the unknown polynomial χ_1. To study this equation define the *homological operator*

$$\begin{aligned} \mathcal{L}_0 : \mathcal{P}_1 &\to \mathcal{P}_1 && (3.11)\\ \chi &\mapsto \mathcal{L}_0\chi := \{H_0, \chi\} && (3.12) \end{aligned}$$

and rewrite equation (3.10) as $\mathcal{L}_0\chi_1 = P_1$, which is a linear equation in the linear space of polynomials of degree 3. Thus, if one is able to diagonalize the operator $\mathcal{L}_0$; it is immediate to understand whether the equation (3.10) is solvable or not.

Remark 3.6. The operator $\mathcal{L}_0$ can be defined also on any one of the spaces $\mathcal{P}_j$, $j \geq 1$, it turns out that $\mathcal{L}_0$ maps polynomials of a given degree into polynomials of the same degree. This is important for the iteration of the construction. For this reason we will study $\mathcal{L}_0$ in $\mathcal{P}_j$ with an arbitrary j.

It turns out that it is quite easy to diagonalize the homological operator in anyone of the spaces $\mathcal{P}_j$. To this end consider the complex variables

$$\xi_l := \frac{1}{\sqrt{2}}(p_l + \mathrm{i}q_l) \; ; \; \eta_l := \frac{1}{\sqrt{2}}(p_l - \mathrm{i}q_l) \quad l \geq 1 \; . \tag{3.13}$$

in which the symplectic form takes the form $\sum_l \mathrm{i}\; d\xi_l \wedge d\eta_l$,[2]

[2]This means that the transformation is not canonical, however, in these variables all the theory remains unchanged except for the fact that the equations of motions have to be substituted by

$$\dot{\xi}_l = \mathrm{i}\frac{\partial H}{\partial \eta_l} \; , \quad \dot{\eta}_l = -\mathrm{i}\frac{\partial H}{\partial \xi_l} \; ,$$

and therefore the Poisson Brackets take the form

$$\{f, g\} := \mathrm{i}\frac{\partial g}{\partial \xi_l}\frac{\partial f}{\partial \eta_l} - \mathrm{i}\frac{\partial g}{\partial \eta_l}\frac{\partial f}{\partial \xi_l}$$

summed over l.

Remark 3.7. In these complex variables the actions are given by

$$I_l = \xi_l \eta_l \ .$$

and

$$H_0(\xi,\eta) = \sum_{l=1}^{n} \omega_l \xi_l \eta_l$$

$\square$

Remark 3.8. Consider a homogeneous polynomial f of the variables (p,q), then it is a homogeneous polynomial of the same degree also when expressed in terms of the variables (ξ,η).

Remark 3.9. The monomials $\xi^J \eta^L$ defined by

$$\xi^J \eta^L := \xi_1^{J_1} \xi_2^{J_2} ... \xi_n^{J^n} \eta_1^{L_1} .... \eta_n^{L_n}$$

form a basis of the space of the polynomials.

Lemma 3.10. *Each element of the basis $\xi^J \eta^L$ is an eigenvector of the operator $\pounds_0$, the corresponding eigenvalue is* $\mathrm{i}(\omega \cdot (L-J))$.

Proof. Just remark that in terms of the variables ξ,η, the action of $\pounds_0$ is given by

$$\begin{aligned} \pounds_0 f = \{H_0, f\} &:= \sum_l \mathrm{i}\frac{\partial f}{\partial \xi_l}\frac{\partial H_0}{\partial \eta_l} - \mathrm{i}\frac{\partial f}{\partial \eta_l}\frac{\partial H_0}{\partial \xi_l} \\ &= \left(\mathrm{i}\sum_l \omega_l \left(\eta_l \frac{\partial}{\partial \eta_l} - \xi_l \frac{\partial}{\partial \xi_l}\right)\right) f \ . \end{aligned}$$

Then

$$\eta_l \frac{\partial}{\partial \eta_l} \xi^J \eta^L = L_l \xi^J \eta^L$$

and thus

$$\pounds_0 \xi^J \eta^L = \mathrm{i}\omega \cdot (L-J) \xi^L \eta^J$$

which is the thesis. $\square$

Thus we have that for each j the space $\mathcal{P}_j$ decomposes into the direct sum of the kernel K of $\pounds_0$ and its range R. In particular the Kernel is generated by the *resonant monomials*, namely

$$K = \mathrm{Span}(\xi^J \eta^L \in \mathcal{P}_i \ : (J,L) \in \mathrm{RS}) \tag{3.14}$$

and

$$\mathrm{RS} := \left\{(J,L) \in \mathbb{N}^{2n} \ : \ \omega \cdot (L-J) = 0\right\} \tag{3.15}$$

is the set of the resonant indexes. Obviously the range is generated by the space monomials $\xi^J \eta^L$ with J, L varying in the complement of the resonant set.

Thus it is easy to obtain the following important lemma.

Lemma 3.11. *Let $P \in \mathcal{P}_j$ be a polynomial, write*

$$P(\xi, \eta) = \sum_{J,L} P_{JL} \xi^J \eta^L \tag{3.16}$$

and define

$$Z(\xi, \eta) := \sum_{(J,L)\in RS} P_{JL} \xi^J \eta^L \ , \quad \chi(\xi, \eta) := \sum_{(J,L)\notin RS} \frac{P_{JL}}{i\omega \cdot (L - J)} \xi^J \eta^L \tag{3.17}$$

then one has

$$Z = \{\chi, H_0\} + P \ . \tag{3.18}$$

and

$$\{Z, H_0\} \equiv 0 \ . \tag{3.19}$$

Remark 3.12. Since both χ and Z are polynomials, the inequalities (3.1,3.2) hold for their moduli and their vector fields.

Motivated by the above lemma we give the following definition.

Definition 3.13. A function Z will be said to be in normal form if it contains only resonant monomials, i.e. if writing

$$Z(\xi, \eta) := \sum_{(J,L)} Z_{JL} \xi^J \eta^L \ , \tag{3.20}$$

one has

$$Z_{JL} \neq 0 \Longrightarrow \omega \cdot (L - J) = 0 \ . \tag{3.21}$$

Remark 3.14. A property which is equivalent to (3.21) is $\{Z, H_0\} = 0$, which has the advantage of being coordinate independent.

Remark 3.15. If the frequencies are nonresonant, namely if eq. (2.9) holds, then the set of the indexes (J, L) such that $\omega \cdot (L - J) = 0$ reduces to the set $J = L$. Thus the resonant monomials are only the monomials of the form

$$\xi^J \eta^J = (\xi_1 \eta_1)^{J_1} ... (\xi_n \eta_n)^{J_n} \equiv I_1^{J_1} ... I_n^{J_n} \ . \tag{3.22}$$

Thus in the nonresonant case a function Z is in normal form if and only if it is a function of the actions only.

Now, one can use the above lemma with $P := P_1$ in order to construct the function χ_1 putting the system in normal form up to order four. We will describe the procedure in detail in the next section.

3.4 Rigorous theory

From now on we will denote by $B_R \subset \mathbb{C}^{2n}$ the closed ball of radius R with center at the origin.

First we estimate the domain where the Lie transform generated by a polynomial $\chi \in \mathcal{P}_j$, $(j \geq 1)$ is well defined.

Lemma 3.16. *Let $\chi \in \mathcal{P}_k$, $(k \geq 1)$ be a polynomial. Denote by ϕ^t the flow of the corresponding vector field. Denote also*

$$\bar{t} = \bar{t}(R, \delta) := \inf_{x \in B_R} \sup \{t > 0 : \phi^s(z) \in B_{R+\delta} \ \forall |s| \leq t \in B_{R+\delta}\}$$

(minimum escape time of $\phi^t(x)$ from $B_{R+\delta}$). Assume $\delta \leq R$, then one has

$$\bar{t} \geq \frac{\delta}{\|X_\chi\| \, (2R)^{k+1}} \tag{3.23}$$

where $\|X_\chi\|$ is the norm defined in remark 3.2. Moreover for any t, such that $|t| \leq \bar{t}$ and any $x \in B_R$ one has

$$\left\|\phi^t(x) - x\right\| \leq |t|(2R)^{k+1} \, \|X_\chi\| \tag{3.24}$$

Proof. First remark that, by the definition of $\bar{t}$ one has that there exists $\bar{x} \in B_R$ such that $\left\|\phi^{\pm\bar{t}}(\bar{x})\right\| = R+\delta$. To be determined we assume that the time realizing the equality is positive. Assume by contradiction $\bar{t} < \frac{\delta}{\|X_\chi\|(2R)^{k+1}}$, then, since for any t with $|t| < \bar{t}$ one has $\phi^t(\bar{x}) \in B_{R+\delta}$, it follows that

$$\left\|\phi^{\bar{t}}(\bar{x})\right\| \leq \|\bar{x}\| + \left\|\phi^{\bar{t}}(\bar{x}) - \bar{x}\right\| = \|\bar{x}\| + \left\|\int_0^{\bar{t}} \frac{d}{ds}\phi^s(\bar{x})\mathrm{d}s\right\|$$
$$\leq R + \int_0^{\bar{t}} \|X_\chi(\phi^s(\bar{x}))ds\| \leq R + |\bar{t}|(R+\delta)^{k+1} \, \|X_\chi\| \ .$$

It follows $\delta < \delta$ which is absurd. □

Since χ is analytic together with its vector field (it is a polynomial), then one has the following corollary.

Corollary 3.17. *Fix arbitrary R and δ, then the map*

$$\begin{aligned} \phi : B_\sigma \times B_R &\to B_{R+\delta} \ , \quad \sigma := \frac{\delta}{\|X_\chi\| \, (2R)^{k+1}} \\ (t, x) &\mapsto \phi^t(x) \end{aligned}$$

is analytic. Here, by abuse of notation, we denoted by B_σ also the ball of radius σ contained $\mathbb{C}$.

Proof of Birkhoff's theorem. We proceed by induction. The theorem is trivially true for $N = 0$. Supposing it is true for N we prove it for $N+1$. First consider the Taylor polynomial of degree $N+3$ of $\mathcal{R}^{(N)}$ and denote it by

$P_1^{(N)} \in \mathcal{P}_{N+1}$. We look for the generating function $\chi_{N+1} \in \mathcal{P}_{N+1}$ of the Lie transform ϕ_{N+1} which normalizes up to order $N+4$ the Hamiltonian. Consider $H^{(N+1)} := H^N \circ \phi_{N+1}$ and write it as follows

$$\begin{aligned} H^{(N)} \circ \phi_{N+1} &= H_0 + Z^{(N)} && (3.25)\\ &+ \{\chi_{N+1}; H_0\} + P_1^{(N)} && (3.26)\\ &+ (Z^{(N)} \circ \Phi_{N+1} - Z^{(N)}) && (3.27)\\ &+ H_0 \circ \phi_{N+1} - (H_0 + \{\chi_{N+1}; H_0\}) && (3.28)\\ &+ (\mathcal{R}^{(N)} - P_1^{(N)}) \circ \phi_{N+1} && (3.29)\\ &+ P_1^{(N)} \circ \phi_{N+1} - P_1^{(N)} \ . && (3.30) \end{aligned}$$

We use lemma 3.11 to construct χ_{N+1} in such a way that

$$Z_{N+1} := \{\chi_{N+1}, H_0\} + P_1^{(N)}$$

is in normal form, and we define $Z^{(N+1)} := Z^{(N)} + Z_{N+1}$. Then one has to prove that all the terms (3.27-3.30) have a zero of order at least $N+4$. This is an immediate consequence of the smoothness of each term and of remark 3.5 which ensures that each line is the remainder of a Taylor expansion (in the space variables) truncated at order $N+3$.

It remains to show that the estimate (2.8) of the deformation holds. Denote by R_{N+1} a positive number such that $B_{2R_{N+1}} \subset \mathcal{U}_N$, and remark that, by lemma 3.16, possibly reducing R_{N+1}, one has

$$\phi_{N+1} : B_\rho \to B_{2\rho} \ , \quad \forall \rho \le R_{N+1}$$

and

$$\sup_{B_\rho} \|x - \phi_{N+1}(x)\| \le C\rho^{N+2} \ . \tag{3.31}$$

Define $\mathcal{T}_{N+1} := \mathcal{T}_N \circ \phi_{N+1}$ then one has

$$I - \mathcal{T}_{N+1} = I - \mathcal{T}_N \circ \phi_{N+1} = I - \mathcal{T}_N + \mathcal{T}_N - \mathcal{T}_N \circ \phi_{N+1}$$

and thus, for any $x \in B_\rho$ with ρ small enough, we have

$$\begin{aligned} \|x - \mathcal{T}_{N+1}(x)\| &\le \|x - \mathcal{T}_N(x)\| + \|\mathcal{T}_N(x) - \mathcal{T}_N(\phi_{N+1}(x))\| \\ &\le C_N\rho^2 + \sup_{x \in B_{2\rho}} \|\mathrm{d}\mathcal{T}_N(x)\| \sup_{x \in B_\rho} \|x - \phi_{N+1}(x)\| \\ &\le C_N\rho^2 + C\rho^{N+2} \le C_{N+1}\rho^2 \end{aligned}$$

from which the thesis follows. □

4 Dynamics

In what follows we fix the number N of normalization steps; moreover, it is useful to distinguish between the original variables and the variables introduced by the normalizing transformation. So, we denote by $x = (p, q)$ the original variables and by $x' = (p', q')$ the normalized variables, i.e. $x = \mathcal{T}_N(x')$. More generally we will denote with a prime the quantities expressed in the normalized variables.

Remark 4.1. Given an initial datum x'_0 consider the corresponding solution $x'(t)$, then one can use the normal form result only as long as $x'(t)$ belongs to the neighborhood $\mathcal{U}_N$ in which the normal form holds.

It follows we will use a related quantity defined as follows. Given $x_0 \in B_R$ we define

$$T_e(x_0, R) := \inf \{T = |t| \ : \ x(t) \notin B_R\} \ . \tag{4.1}$$

Remark 4.2. By the theorem of existence and uniqueness of solutions one has $T_e > 0$.

In all the interesting cases we will bound from below T_e, at least for initial data in ball smaller than B_R.

4.1 The general case

Consider the simplified system in which the remainder is neglected, namely the system with Hamiltonian

$$H_s = H_0 + Z^{(N)} \ . \tag{4.2}$$

To start with we show that integrals of motion of H_s are approximatively constant along the solutions of the complete Hamiltonian (2.1).

Theorem 4.3. *Let F be a polynomial having a zero of order k at the origin. Assume it is an integral of motion of H_s, then there exists R_* (independent of F) such that, if the initial datum fulfills*

$$\epsilon := \|x_0\| < R_*$$

then along the corresponding solution of (2.1) *one has*

$$|F(t) - F(0)| \le C\epsilon^{k+1} \quad \text{for} \quad |t| \le \min\left\{\frac{1}{\epsilon^N}, T_e(x_0, 2\epsilon)\right\} \ , \tag{4.3}$$

where $F(t) := F(x(t))$.

Proof. Assume that ϵ is so small that $B_{3\epsilon} \subset \mathcal{U}_N$; perform the normalizing transformation. Remark that for $|t| < T_e(x_0, 2\epsilon)$ one has $x(t) \in B_{2\epsilon}$; by (2.8) one also has $x'(t) \in B_{3\epsilon}$. In particular, provided ϵ is small enough, one has $x'(t) \in \mathcal{U}_N$. Consider now

$$|F(x(t)) - F(x_0)| \le |F(x(t)) - F(x'(t))| + |F(x'(t)) - F(x'_0)| + |F(x_0) - F(x'_0)| \ . \tag{4.4}$$

The first term at r.h.s. is estimated as follows

$$|F(x_0) - F(x'_0)| \le \sup_{x \in B_{3\epsilon}} \|\mathrm{d}F(x)\| \, \|x_0 - x'_0\| \le C\epsilon^{k-1}\epsilon^2$$

where we used the fact that the differential of a polynomial having a zero of order k has a zero of order $k-1$ and we also used the estimate (2.8). The third term is estimated in the same way.

Concerning the middle term, since for all the considered times $x'(t) \in \mathcal{U}_N$, one has

$$|F(x'(t)) - F(x_0)| = \left| \int_0^t \left\{ H^{(N)}, F \right\} (x'(s)) ds \right| \leq \int_0^t \left| \left\{ \mathcal{R}^{(N)}, F \right\} (x'(s)) \right| ds$$
$$\leq |t| C\epsilon^{N+3+k-2} \leq C\epsilon^{k+1}$$

where the last inequality holds for the times (4.3). $\square$

Remark 4.4. By construction, at least H_0 is an integral of motion of the simplified system, thus the theorem applies in particular to the case $F = H_0$.

Remark 4.5. By the estimate (4.4) one also has that, defining $F'(x) := F \circ \mathcal{T}^{-1}$, the following estimate holds

$$|F'(x(t)) - F'(x_0)| \leq C\epsilon^{k+1+M} \text{ for } \quad |t| \leq \frac{1}{\epsilon^{N-M}} \tag{4.5}$$

4.2 Nonresonant case

We come now to the nonresonant case, namely we assume here that (2.9) holds. In this case, as pointed out at the end of section 2, the normal form depends on the actions

$$I_l' = \frac{p_l'^2 + q_l'^2}{2} \tag{4.6}$$

only. Therefore the actions are integrals of motion for the normalized system. As a consequence it is also possible to give a lower bound on the escape time T_e obtaining the following proposition.

Proposition 4.6. *Assume $\omega \cdot k \neq 0$ for any $k \in \mathbb{Z}^n$, with $0 \neq |k| \leq N+2$; then the following holds true: there exists R_* such that, if the initial datum fulfills*

$$\epsilon := \|x_0\| < R_*$$

then, along the corresponding solution $x(t)$ one has

$$\|x(t)\| \leq 2\epsilon \quad for \quad |t| \leq \frac{1}{\epsilon^N} \tag{4.7}$$

Moreover along the same solution one has

$$|I_l'(t) - I_l'(0)| \leq C\epsilon^{M+3} \quad for \quad |t| \leq \frac{1}{\epsilon^{N-M}} \ , \quad M < N \tag{4.8}$$

and

$$|I_l(t) - I_l(0)| \leq C\epsilon^3 \quad for \quad |t| \leq \frac{1}{\epsilon^N} \ . \tag{4.9}$$

Proof. Following the scheme of the proof of Lyapunov's theorem we show that $T_e(x_0, 2\epsilon) > \epsilon^{-N}$. Define $F(x) := \sum I_k \equiv \|x\|^2$, and assume by contradiction that $T_e < \epsilon^{-N}$. Let $\bar{x}_0$ be the initial datum giving rise to a solution which escapes from $B_{2\epsilon}$ at time T_e; by theorem 4.3, one has

$$4\epsilon^2 = \|x(T_e)\|^2 = F(T_e) \le F(x_0) + |F(T_e) - F(x_0)| \le \epsilon^2 + C\epsilon^3 \ , \tag{4.10}$$

which is impossible if $\epsilon < C^{-1}$. $\square$

Corollary 4.7. *There exists a smooth torus $\mathbb{T}_0$ such that, $\forall M \le N$*

$$d(x(t), \mathbb{T}_0) \le C\epsilon^{(M+3)/2} \ , \quad \text{for} \quad |t| \le \frac{1}{\epsilon^{N-M}} \tag{4.11}$$

where $d(.,.)$ is the distance in $\mathbb{R}^{2n}$.

Proof. First remark that in the normalized coordinates one has

$$\dot{I}'_j = -p'_j \frac{\partial \mathcal{R}^{(N)}}{\partial q'_j} + q'_j \frac{\partial \mathcal{R}^{(N)}}{\partial p'_j} \ ,$$

so that

$$\sum_j \left| \dot{I}'_j \right| = \sum_j \left| -p'_j \frac{\partial \mathcal{R}^{(N)}}{\partial q'_j} + q'_j \frac{\partial \mathcal{R}^{(N)}}{\partial p'_j} \right| \tag{4.12}$$

$$\le \left(\sum_j ({p'_j}^2 + {q'_j}^2) \right)^{1/2} \left(\sum_j \left(\left| \frac{\partial \mathcal{R}^{(N)}}{\partial q'_j} \right|^2 + \left| \frac{\partial \mathcal{R}^{(N)}}{\partial p'_j} \right|^2 \right) \right)^{1/2} \le \|x'\| \, \|X_{\mathcal{R}^{(N)}}(x')\| \tag{4.13}$$

Denote $\bar{I}_j{}' := \frac{{p'_j}^2(0) + {q'_j}^2(0)}{2}$ and define the torus

$$\mathbb{T}'_0 := \left\{ x' \in \mathbb{R}^{2n} \ : \ I_j(x') = \bar{I}_j \ , j = 1, ..., n \right\}$$

One has

$$d(x'(t), \mathbb{T}'_0) \le \left[\sum_j \left| \sqrt{I'_j(t)} - \sqrt{\bar{I}_j} \right|^2 \right]^{1/2} \tag{4.14}$$

Notice that for $a, b \ge 0$ one has,

$$\left| \sqrt{a} - \sqrt{b} \right| \le \sqrt{|a - b|} \ .$$

Thus, using (4.13), one has that

$$[d(x'(t), \mathbb{T}'_0)]^2 \le \sum_j |I'_j(t) - \bar{I}_j| \le C\epsilon^{M+3}$$

Define now $\mathbb{T}_0 := \mathcal{T}_N(\mathbb{T}'_0)$ then, since $\mathcal{T}_N$ is Lipschitz one has

$$d(x(t), \mathbb{T}_0) = d(\mathcal{T}_N(x'(t)), \mathcal{T}_N(\mathbb{T}'_0)) \le C d(x'(t), \mathbb{T}'_0) \le C\epsilon^{\frac{M+3}{2}} \ .$$

$\square$

5 A resonant example

In this section we will study in detail the dynamics of the normal form of a system whose linear part is composed by two oscillators with the same frequency. Thus we take

$$H_0 = \omega \left(\frac{p_1^2 + q_1^2}{2} + \frac{p_2^2 + q_2^2}{2} \right) . \tag{5.1}$$

Introduce the complex variables (ξ, η) and compute the first nontrivial term, of the normal form. To this end we compute the resonant monomials. First we study the resonance relations between the frequencies.

Remark 5.1. For $K = (K_1, K_2)$, one has

$$\omega_1 K_1 + \omega_2 K_2 = \omega(K_1 + K_2) = 0 \Longrightarrow K_1 = -K_2 \Longrightarrow |K| = 2|K_1|$$

so that the modulus of K has to be even, and therefore there are no resonant monomials of order 3.

The resonant monomials of order 4 are of the form $\xi^L \eta^J$ with $K_1 \equiv L_1 - J_1 = L_2 - J_2 \equiv K_2$ and $|L| + |J| = 4$. So it easy to construct all of them that have to satisfy

$$L_1 + L_2 = J_1 + J_2 \ , \quad L_1 + L_2 = 2$$

We enumerate below the resonant indexes and the corresponding resonant monomials

$$\begin{array}{ccccc} L_1 & L_2 & J_1 & J_2 & \\ 0 & 2 & 0 & 2 & \xi_2^2\eta_2^2 = I_2^2 \\ 0 & 2 & 1 & 1 & \xi_2^2\eta_1\eta_2 = I_2\xi_2\eta_1 \\ 0 & 2 & 2 & 0 & \xi_2^2\eta_1^2 \\ 1 & 1 & 0 & 2 & I_2\xi_1\eta_2 \\ 1 & 1 & 1 & 1 & I_1 I_2 \\ 1 & 1 & 2 & 0 & I_1\xi_1\eta_2 \\ 2 & 0 & 0 & 2 & \xi_1^2\eta 2^2 \\ 2 & 0 & 1 & 1 & I_1\xi_1\eta_2 \\ 2 & 0 & 2 & 0 & I_1^2 \end{array} \tag{5.2}$$

The most general *real* Hamiltonian that one can form with the above monomials is given by

$$Z^{(2)}(\xi, \eta) = aI_2^2 + bI_2(\xi_2\eta_1 + \eta_2\xi_1) + c(\xi_2^2\eta_1^2 + \xi_1^2\eta_2^2) \tag{5.3}$$

$$+dI_2I_1 + eI_1(\xi_2\eta_1 + \eta_2\xi_1) + fI_1^2 \tag{5.4}$$

with arbitrary a, b, c, d, e, f. In order to understand the behavior of the system $H_0 + Z^{(2)}$ it is useful to introduce action angle variables by the change of variables

$$\xi_l = \sqrt{I_l} e^{i\varphi_l} \ , \quad \eta_l = \sqrt{I_l} e^{-i\varphi_l} \ .$$

Then one has

$$\begin{aligned} Z^{(2)}(I, \varphi) &= a_1 I_1^2 + a_2 I_1 I_2 + a_3 I_2^2 \\ &+ b_1 I_1 I_2 \cos 2(\varphi_1 - \varphi_2) + b_2(c_1 I_1 + c_2 I_2)\sqrt{I_1 I_2} \cos(\varphi_1 - \varphi_2) \end{aligned}$$

Since the Hamiltonian depends on the angles only through the combination $\varphi_1 - \varphi_2$ it is useful to make the canonical coordinate change

$$J_1 = I_1 \ , \quad J_2 = I_1 + I_2 \tag{5.5}$$

$$\psi_1 = \varphi_1 - \varphi_2 \ , \quad \psi_2 = \varphi_2 \tag{5.6}$$

in which H_s takes the form

$$H_s = \omega J_2 + a_1' J_1^2 + a_2' J_1 J_2 + a_3' J_2^2 + b J_1 (J_2 - J_1) \cos 2\psi_1 \tag{5.7}$$

$$(\tilde{c}_1 J_1 + \tilde{c}_2 J_2) \sqrt{J_1 (J_2 - J_1)} \cos \psi_1 \tag{5.8}$$

from which it is obvious that $J_2 = H_0'$ is an integral of motion. Then one can pass to the reduced system in which J_2 is considered as a parameter and study the dynamics of the resulting system with one degree of freedom which is integrable. In particular the level surface of H_s are orbits of the corresponding system. The dynamics depends on the values of the different parameters. Here, in order to put into evidence a behaviors which is typical, we study in detail a particular case.

Thus fix $a_1' = a_2' = a_3' = b = 0$ so that the nonlinear part of the Hamiltonian reduces to (5.8). Moreover we fix $\tilde{c}_1 = \tilde{c}_2 = 1$. Then we study the reduced system on the surface $J_2 = 1$. Its Hamiltonian is thus given by

$$H_r := (J_1 + 1) \sqrt{J_1 (1 - J_1)} \cos \psi_1 \tag{5.9}$$

and its Hamilton equations are given by

$$\dot{J}_1 = (J_1 + 1) \sqrt{J_1 (1 - J_1)} \sin \psi_1 \ , \quad \dot{\psi}_1 = f(J_1) \cos \psi_1 \tag{5.10}$$

with a suitable $f(J_1)$ whose form is not important here. Remark also that since $J_2 = 1$ is the total harmonic energy the variable J_1 which represents the harmonic energy of the first oscillator is subjected to the limitation $0 \leq J_1 \leq 1$. Concerning the angle we study the system for $-\pi/2 < \psi_1 \leq 3\pi/2$. From the equations of motion it is clear that the lines $J_1 = 1$ and $J_1 = 0$ are invariant (on these lines $\dot{J}_1 = 0$) and similarly for the vertical lines $\psi_1 = -\pi/2$ and $\psi_1 = \pi/2$. Moreover on the first vertical line the flow is downward, while on the second one it is upward. Thus, by continuity the phase portrait of the system is the one in the figure `to be inserted`. One sees that there are orbits in which J_1 pass from an arbitrary small value to a value arbitrary close to 1 and then comes back. These are nonlinear beats in which the energy pass from one oscillator and back.

The solutions of the complete system have a similar behavior. To give a precise statement normalize the Hamiltonian up to order $N + 4$. Then the simplified system turns out to have the form (4.2) with

$$Z^{(N)} = Z^{(2)} + \sum_{k=3}^{N} Z_k \ , \quad Z_k \in \mathcal{P}_k \ ,$$

which therefore is a perturbation of $Z^{(2)}$. Both $Z^{(N)}$ and H_0 are integrals of motion of H_s, and thus the intersection of their level surface are invariant manifolds for H_s and, thank to theorem 4.3 approximate invariant manifolds for the complete dynamics. Thus one can prove a result similar to theorem 4.3, precisely one has the following theorem.

Theorem 5.2. *There exists R_* such that, if the initial datum fulfills*

$$\epsilon := \|x_0\| < R_*$$

then along the corresponding solution one has

$$|H_0(t) - H(0)| \leq C\epsilon^3 \ , \quad \left| Z^{(2)}(t) - Z^{(2)}(0) \right| \leq C\epsilon^5 \ , \quad |t| \leq \frac{1}{\epsilon^N} \ . \tag{5.11}$$

The main interest of this result rests in the fact that it ensures that over very long time the dynamics remains close to the dynamics of $H_0 + Z^{(2)}$ (that we just studied), which is an integrable system, but an integrable system with invariant tori topologically different from the invariant tori of H_0.