A Birkhoff normal form theorem for some semilinear PDEs

D. Bambusi

Abstract In these lectures we present an extension of Birkhoff normal form theorem to some Hamiltonian PDEs. The theorem applies to semilinear equations with nonlinearity of a suitable class. We present an application to the nonlinear wave equation on a segment or on a sphere. We also give a complete proof of all the results.

1 Introduction

These lectures concern some qualitative features of the dynamics of semilinear Hamiltonian PDEs. More precisely we will present a normal form theorem for such equations and deduce some dynamical consequences. In particular we will deduce almost global existence of smooth solutions (in the sense of Klainerman [Kla83]) and a result bounding the exchange of energy among degrees of freedom with different frequency. In the case of nonresonant systems we will show that any solution is close to an infinite dimensional torus for times longer than any inverse power of the size of the initial datum. The theory presented here was developed in [Bam03, BG06, DS06, BDGS07, Gré06].

In order to illustrate the theory we will use as a model problem the nonlinear wave equation

$$u_{tt} - \Delta u + \mu^2 u = f(u) \ , \quad \mu \in \mathbb{R} \ , \tag{1}$$

on a d dimensional sphere or on $[0, \pi]$ with Neumann boundary conditions. In eq. (1), f is a smooth function having a zero of order 2 at the origin and Δ is the Laplace Beltrami operator.

D. Bambusi
Dipartimento di Matematica dell'Università, Via Saldini 50, 20133 Milano, Italy, e-mail: dario.bambusi@unimi.it

The theory of Birkhoff normal form is a particular case of the theory of close to integrable Hamiltonian systems. Concerning the extension to PDEs of Hamiltonian perturbation theory, the most celebrated results are the KAM type theorems due to Kuksin [Kuk87], Wayne [Way90], Craig–Wayne [CW93], Bourgain [Bou98, Bou05], Kuksin-Pöschel [KP96], Eliasson–Kuksin [EK06], Yuan [Yua06]. All these results ensure the existence of families of quasiperiodic solutions, so they only describe solutions lying on finite dimensional manifolds in an infinite dimensional phase space. On the contrary the result on which we concentrate here allows one to describe *all* small amplitude solutions of the considered systems. The price we pay is that the description turns out to be valid only over long but finite times.

A related research stream is the one carried on by Bourgain [Bou96a, Bou96b, Bou97, Bou00] who studied intensively the behavior of high Sobolev norms in close to integrable Hamiltonian PDEs. In particular he gave some lower estimates showing that in some cases high Sobolev norm can grow in an unbound way, and also some upper estimate showing that the nonlinearity can stabilize resonant systems, somehow in the spirit of Nekhoroshev's theorem.

The paper is organized as follows. First we present the classical Birkhoff normal form theorem for finite dimensional systems and we recall its proof (see sect. 2). Then we pass to PDEs. Precisely, in sect. 3 we first show that the nonlinear wave equation is an infinite dimensional Hamiltonian system (subsection 3.1) and then we present the problem met in trying to extend the normal form theorem to PDEs. Subsequently we give a heuristic discussion on how to solve such difficulties (see subsection 3.2). Then we give a precise formulation of our Birkhoff normal form theorem (sect. 4). This part contains only the statements of the results and is split into three subsection, in the first (subsection 4.1) we introduce the class of functions to which the theory applies and we study its properties. In the second subsection (subsection 4.2) we give the statement of the normal form theorem and deduce the main dynamical consequences. In the third subsection 4.3 we give the application to the considered model. Then, in sect. 5 we give a short discussion presenting the main open problems of the domain.

Finally sect. 6 contains the proofs of all the results. The subsections of this section are independent each other. We made an effort to give a paper which is essentially self contained. We also mention that the method introduced here in order to prove the property of localization of coefficients (the property defining our class of functions) is original.

2 Birkhoff's theorem in finite dimension

2.1 Statement.

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

Definition 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

$$H(p,q) := H_0(p,q) + H_P(p,q) \ , \tag{2}$$

where

$$H_0(p,q) = \sum_{l=1}^{n} \omega_l \frac{p_l^2 + q_l^2}{2} \ , \quad \omega_l \in \mathbb{R} \tag{3}$$

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

Remark 1. The equations of motion of (2) have the form

$$\dot{p}_l = -\omega_l q_l - \frac{\partial H_P}{\partial q_l} \tag{4}$$

$$\dot{q}_l = \omega_l p_l + \frac{\partial H_P}{\partial p_l} \tag{5}$$

Since H_P has a zero of order three, its gradient starts with quadratic terms. Thus, in the linear approximation the equations (4), (5) take the form

$$\begin{array}{c} \dot{p}_l = -\omega_l q_l \\ \dot{q}_l = \omega_l p_l \end{array} \Longrightarrow \ddot{q}_l + \omega_l^2 q_l = 0 \tag{6}$$

namely the system consists of n independent harmonic oscillators.

Definition 2. The vector field

$$X_H(p,q) := \left(-\frac{\partial H}{\partial q} \ , \ \frac{\partial H}{\partial p} \right) \tag{7}$$

is called the Hamiltonian vector field of H.

Theorem 1. *(Birkhoff) For any positive integer $r \geq 0$, there exist a neighborhood $\mathscr{U}^{(r)}$ of the origin and a canonical transformation $\mathscr{T}_r : \mathbb{R}^{2n} \supset \mathscr{U}^{(r)} \to \mathbb{R}^{2n}$ which puts the system (2) in Birkhoff Normal Form up to order r, namely*

$$H^{(r)} := H \circ \mathscr{T}_r = H_0 + Z^{(r)} + \mathscr{R}^{(r)} \tag{8}$$

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

$$|\mathscr{R}^{(r)}(z)| \leq C_r \|z\|^{r+3} \ , \quad \forall z \in \mathscr{U}^{(r)} \ ; \tag{9}$$

moreover, one has

$$\|z - \mathscr{T}_r(z)\| \leq C_r \|z\|^2 \ , \quad \forall z \in \mathscr{U}^{(r)} \ . \tag{10}$$

An inequality identical to (10) *is fulfilled by the inverse transformation* $\mathscr{T}_r^{-1}$.

If the frequencies are nonresonant at order $r+2$, *namely if*

$$\omega \cdot k \neq 0 \ , \quad \forall k \in \mathbb{Z}^n \ , \quad 0 < |k| \leq r+2 \tag{11}$$

the function $Z^{(r)}$ *depends on the actions*

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

only.

Remark 2. The remainder $\mathscr{R}^{(r)}$ is very small in a small neighborhood of the origin. In particular, it is of order ε^{r+3} in a ball of radius ε. It will be shown in sect. 4.2 that in typical cases $\mathscr{R}^{(r)}$ might have a relevant effect only after a time of order ε^{-r}.

2.2 Proof

The idea of the proof is to construct a canonical transformation putting the system in a form which is as simple as possible. 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). We are now going to describe more precisely this method.

Definition 3. We will denote by $\mathscr{H}_j$ the set of the real valued homogeneous polynomials of degree $j+2$.

Remark 3. Let $g \in \mathscr{H}_j$ be a homogeneous polynomial, then there exists a constant C such that

$$|g(z)| \leq C \|z\|^{j+2} \ . \tag{12}$$

The Hamiltonian vector field X_g of g is a homogeneous polynomial of degree $j+1$ and therefore one has

$$\left\|X_g(z)\right\| \leq C' \|z\|^{j+1} \tag{13}$$

with a suitable constant C'. The best constant such that (13) holds is usually called the norm of X_g and is denoted by $\left\|X_g\right\|$. Similarly one can define the norm of the polynomial g.

Remark 4. If the phase space is infinite dimensional then equations (12) and (13) are not automatic. They hold if and only if the considered polynomial are smooth.

Remark 5. Let $f \in \mathscr{H}_i$ and $g \in \mathscr{H}_j$ then, by the very definition of Poisson Brackets one has $\{f, g\} \in \mathscr{H}_{i+j}$.

2.2.1 Lie Transform

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

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

and denote by ϕ^t the corresponding flow.

Definition 4. The time one map $\phi := \phi^t\big|_{t=1}$ is called the *Lie transform generated by* χ. It is well known that ϕ is a canonical transformation.

We are now going to study the way a polynomial transforms when the coordinates are subjected to a Lie transformation.

Lemma 1. *Let $g \in \mathscr{H}_i$ be a polynomial and let ϕ be the Lie transform generated by a polynomial $\chi \in \mathscr{H}_j$ with $j \geq 1$. Define*

$$g_0 := g, \quad g_l = \frac{1}{l}\{\chi; g_{l-1}\} , \quad l \geq 1 , \tag{14}$$

then the Taylor expansion of $g \circ \phi$ is given by

$$g(\phi(z)) = \sum_{l \geq 0} g_l(z) , \tag{15}$$

for all z small enough.

Proof. Compute the Taylor expansion of $g \circ \phi^t$ with respect to time. Iterating the relation

$$\frac{d}{dt} g \circ \phi^t = \{\chi, g\} \circ \phi^t \tag{16}$$

one has

$$\frac{d^l}{dt^l} g \circ \phi^t = \underbrace{\{\chi, ...\{\chi, g\}}_{l \text{ times}} \circ \phi^t \tag{17}$$

which gives

$$g \circ \phi^t = \sum_{l \geq 0} t^l g_l . \tag{18}$$

Evaluating at $t = 1$ one gets (15). Since remark 5 implies $g_l \in \mathscr{H}_{i+lj}$, eq. (15) is the Taylor expansion of $g \circ \phi$ *as a function of the phase space variables z.* □

Remark 6. Corollary 4 below shows that the series (15) is convergent in a neighborhood of the origin small enough.

2.2.2 The homological equation

We are now ready to construct a canonical transformation normalizing the system up to terms of fourth order. Thus let $\chi_1 \in \mathscr{H}_1$ be the generating function of the Lie transform ϕ_1, and consider $H \circ \phi_1$, with H given by (2). Using (15) and (14) to compute the first terms of the Taylor expansion of $H \circ \phi$ one gets

$$H \circ \phi = H_0 + \{\chi_1, H_0\} + H_1 + \text{h.o.t}$$

where H_1 is the Taylor polynomial of degree three of H_P and h.o.t. denotes higher order terms.

We wont to construct χ_1 in such a way that

$$Z_1 := \{\chi_1, H_0\} + H_1 \tag{19}$$

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

$$\{\chi_1, H_0\} + H_1 = 0 \tag{20}$$

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

$$\mathscr{L}_0 : \mathscr{H}_1 \to \mathscr{H}_1 \tag{21}$$

$$\chi \mapsto \mathscr{L}_0 \chi := \{H_0, \chi\} \tag{22}$$

and rewrite equation (20) as $\mathscr{L}_0 \chi_1 = H_1$, which is a linear equation in the finite dimensional linear space of polynomials of degree 3. Thus, if one is able to diagonalize the operator $\mathscr{L}_0$; it is immediate to understand whether the equation (20) is solvable or not.

Remark 7. The operator $\mathscr{L}_0$ can be defined also on any one of the spaces $\mathscr{H}_j$, $j \geq 1$, it turns out that $\mathscr{L}_0$ 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 $\mathscr{L}_0$ in $\mathscr{H}_j$ with an arbitrary j.

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

$$\xi_l := \frac{1}{\sqrt{2}}(p_l + \mathrm{i}q_l) \; ; \; \eta_l := \frac{1}{\sqrt{2}}(p_l - \mathrm{i}q_l) \quad l \geq 1 \; . \tag{23}$$

in which the symplectic form takes the form $\sum_l \mathrm{i} \, d\xi_l \wedge d\eta_l$,[1]

Remark 8. In these complex variables the actions are given by

[1] 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

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

and

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

Remark 9. 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 10. 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 2. *Each element of the basis $\xi^J\eta^L$ is an eigenvector of the operator $\mathscr{L}_0$, the corresponding eigenvalue is* $\mathrm{i}(\omega\cdot(L-J))$.

Proof. Just remark that in terms of the variables ξ,η, the action of $\mathscr{L}_0$ is given by

$$\begin{aligned}\mathscr{L}_0 f = \{H_0, f\} &:= \sum_l \mathrm{i}\frac{\partial f}{\partial \xi_l}\frac{\partial H_0}{\partial \eta_l} - \mathrm{i}\frac{\partial f}{\partial \eta_l}\frac{\partial H_0}{\partial \xi_l} \\ &= \left(\mathrm{i}\sum_l \omega_l\left(\eta_l\frac{\partial}{\partial \eta_l} - \xi_l\frac{\partial}{\partial \xi_l}\right)\right) f \ .\end{aligned}$$

Then

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

and thus

$$\mathscr{L}_0\xi^J\eta^L = \mathrm{i}\omega\cdot(L-J)\xi^L\eta^J$$

which is the thesis. □

Thus we have that for each j the space $\mathscr{H}_j$ decomposes into the direct sum of the kernel K of $\mathscr{L}_0$ and its range R. In particular the Kernel is generated by the *resonant monomials*, namely

$$K = \mathrm{Span}(\xi^J\eta^L \in \mathscr{H}_i \ : (J,L) \in \mathrm{RS}) \tag{24}$$

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

and therefore the Poisson Brackets take the form

$$\{f,g\} := \mathrm{i}\sum_l\left(\frac{\partial g}{\partial \xi_l}\frac{\partial f}{\partial \eta_l} - \frac{\partial g}{\partial \eta_l}\frac{\partial f}{\partial \xi_l}\right) \ .$$

and

$$\mathrm{RS} := \{(J,L)\ :\ \omega\cdot(L-J)=0\} \tag{25}$$

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. *Let $f\in\mathscr{H}_j$ be a polynomial, write*

$$f(\xi,\eta)=\sum_{J,L} f_{JL}\xi^J\eta^L \tag{26}$$

and define

$$Z(\xi,\eta):=\sum_{(J,L)\in RS} f_{JL}\xi^J\eta^L\ ,\quad \chi(\xi,\eta):=\sum_{(J,L)\notin RS}\frac{f_{JL}}{\mathrm{i}\omega\cdot(L-J)}\xi^J\eta^L \tag{27}$$

then one has

$$Z=\{\chi,H_0\}+f\ . \tag{28}$$

and

$$\{Z,H_0\}\equiv 0\ . \tag{29}$$

Motivated by the above lemma we give the following definition.

Definition 5. A function Z will be said to be in normal form if, when written in terms of the variables ξ,η, it contains only resonant monomials, i.e. if writing

$$Z(\xi,\eta):=\sum_{(J,L)} Z_{JL}\xi^J\eta^L\ , \tag{30}$$

one has

$$Z_{JL}\neq 0 \Longrightarrow \omega\cdot(L-J)=0\ . \tag{31}$$

Remark 11. A property which is equivalent to (31) is $\{Z,H_0\}=0$, which has the advantage of being coordinate independent.

Remark 12. If the frequencies are nonresonant, namely if eq. (11) 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

$$\xi^J\eta^J=(\xi_1\eta_1)^{J_1}...(\xi_n\eta_n)^{J_n}\equiv I_1^{J_1}...I_n^{J_n}\ . \tag{32}$$

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

2.2.3 Proof of Birkhoff's theorem.

We proceed by induction. The theorem is trivially true for $r=0$. Supposing it is true for r we prove it for $r+1$. First consider the Taylor polynomial of degree $r+3$

of $\mathscr{R}^{(r)}$ and denote it by $H_{r+1}^{(r)} \in \mathscr{H}_{r+1}$. Let $\chi_{r+1} \in \mathscr{H}_{r+1}$ be the solution of the homological equation

$$\{\chi_{r+1}; H_0\} + H_{r+1}^{(r)} = Z_{r+1} \tag{33}$$

with Z_{r+1} in normal form. By lemma 3 such a χ_{r+1} exists. By corollary 4 below, χ_{r+1} generates an analytic flow. Use it to generate the Lie transform ϕ_{r+1} and consider $H^{(r+1)} := H^{(r)} \circ \phi_{r+1}$ and write it as follows

$$\begin{aligned} H^{(r)} \circ \phi_{r+1} &= H_0 + Z^{(r)} && (34)\\ &+ Z_{r+1} && (35)\\ &+ (Z^{(r)} \circ \phi_{r+1} - Z^{(r)}) && (36)\\ &+ H_0 \circ \phi_{r+1} - (H_0 + \{\chi_{r+1}; H_0\}) && (37)\\ &+ (\mathscr{R}^{(r)} - H_{r+1}^{(r)}) \circ \phi_{r+1} && (38)\\ &+ H_{r+1}^{(r)} \circ \phi_{r+1} - H_{r+1}^{(r)} \ . && (39) \end{aligned}$$

define $Z^{(r+1)} := Z^{(r)} + Z_{r+1}$. To prove that the terms (36-39) have a vector field with a zero of order at least $r+3$ use lemma 1 which ensures that each line is the remainder of a Taylor expansion (in the space variables) truncated at order $r+3$.

It remains to show that the estimate (10) of the deformation holds. Denote by R_{r+1} a positive number such that $B_{2R_{r+1}} \subset \mathscr{U}_s^{(r)}$, and remark that, by lemma 6, possibly reducing R_{r+1}, one has

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

and

$$\sup_{B_\rho} \|z - \phi_{r+1}(z)\| \le C\rho^{r+2} \ . \tag{40}$$

Define $\mathscr{T}_{r+1} := \mathscr{T}_r \circ \phi_{r+1}$ then one has

$$Id - \mathscr{T}_{r+1} = Id - \mathscr{T}_r \circ \phi_{r+1} = Id - \mathscr{T}_r + \mathscr{T}_r - \mathscr{T}_r \circ \phi_{r+1}$$

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

$$\begin{aligned} \|z - \mathscr{T}_{r+1}(z)\| &\le \|z - \mathscr{T}_r(z)\| + \|\mathscr{T}_r(z) - \mathscr{T}_r(\phi_{r+1}(z))\| \\ &\le C_r\rho^2 + \sup_{z\in B_{2\rho}} \|\mathrm{d}\mathscr{T}_r(z)\| \sup_{z\in B_\rho} \|z - \phi_{r+1}(z)\| \\ &\le C_r\rho^2 + C\rho^{r+2} \le C_{r+1}\rho^2 \end{aligned}$$

from which the thesis follows.

3 The case of PDEs

3.1 Hamiltonian formulation of the wave equation

Consider the nonlinear wave equation (1).

It is well known that the energy is a conserved quantity for (1). It is given by

$$H(u,v) := \int_D \left(\frac{v^2}{2} - \frac{u\Delta u}{2} + \frac{\mu^2 u^2}{2} \right) \mathrm{d}^d x + \int_D F(u) \mathrm{d}^d x \tag{41}$$

where $v := u_t$ and F is such that $-F' = f$, and D is either S^d (d-dimensional sphere) or $[0,\pi]$. The function H is also the Hamiltonian of the system and the corresponding Hamilton equations are given by

$$\dot{u} = \nabla_v H \ , \quad \dot{v} = -\nabla_u H \tag{42}$$

where $\nabla_u H$ is the L^2 gradient of H with respect to u, defined by

$$\langle \nabla_u H; h \rangle_{L^2} = \mathrm{d}_u H h \quad \forall h \in C^\infty(D) \tag{43}$$

where d_u is the differential with respect to the u variables. ∇_v is defined similarly.

To write (1) in the form (2) we have to introduce the basis of the eigenfunctions of the Laplacian.

In the case of $[0,\pi]$ the eigenfunctions are given by

$$\mathbf{e}_1 := \frac{1}{\sqrt{\pi}} \ , \quad \mathbf{e}_j := \frac{1}{\sqrt{\pi/2}} \cos((j-1)x) \ , \quad j \geq 2 \tag{44}$$

and the corresponding eigenvalues of $-\Delta$ are $\lambda_j = (j-1)^2$.

In the case of the d dimensional sphere the eigenvalues λ_j of $-\Delta$ are given by

$$\lambda_j = (j-1)(j+d-2) \ ; \tag{45}$$

moreover the j^{th} eigenvalue has multiplicity

$$l^*(j) := \binom{j+d-1}{d} \ .$$

We will denote by $\mathbf{e}_{jl}$ a basis of eigenfunctions of the Laplacian, which is orthonormal in L^2 and such that

$$-\Delta \mathbf{e}_{jl} = \lambda_j \mathbf{e}_{jl} \ , \quad j \geq 1 \ , \quad l = 1, ..., l^*(j) \ . \tag{46}$$

For example they can be chosen to be the spherical harmonics.

In both cases define ω_j, p_{jl} and q_{jl} by

$$\omega_j := \sqrt{\lambda_j + \mu^2} \tag{47}$$

$$u = \sum_{jl} \frac{q_{jl}}{\sqrt{\omega_j}} \mathbf{e}_{jl} \ , \quad v = \sum_{jl} \sqrt{\omega_j} p_{jl} \mathbf{e}_{jl} \tag{48}$$

with the convention that l takes only the value 1 in the case of $[0,\pi]$ (and that, in such a case it will not be written).

Then the Hamiltonian (41) takes the form (2) with

$$H_0 = \sum_j \sum_l \omega_j \frac{p_{jl}^2 + q_{jl}^2}{2} \tag{49}$$

and H_P is given by the second integral in (41) considered as a function of q_{jl}.

3.2 Extension of Birkhoff's theorem to PDEs: Heuristic ideas

In this section we will concentrate on the case of the nonlinear wave equation on $[0,\pi]$.

The main difficulty one meets in order to extend the theory of Birkhoff normal form to infinite dimensional systems rests in the denominators one meets in solving the homological equation, namely in the second of equations (27). Indeed, while in the finite dimensional case one has that the set of vectors with integer components having modulus smaller than a given r is finite, this is no more true in infinite dimensions.

It turns out that typically the denominators in (27) accumulate to zero already at order 4. An example of such a behavior is the following one. Consider $\omega_{j+1} := \sqrt{j^2 + \mu^2}$. For $l \geq 1$ consider the integer vector $K^{(l)}$ whose only components different from zero are given by $K_l = -2$, $K_{l-1} = 1$ $K_{l+1} = 1$; such a vector has modulus 4, and one has

$$K^{(l)} \cdot \omega = \omega_{l+1} + \omega_{l-1} - 2\omega_l = \sqrt{l^2 + \mu^2} + \sqrt{(l-2)^2 + \mu^2} - 2\sqrt{(l-1)^2 + \mu^2} \sim \frac{\mu^2}{l^3} \to 0$$

Thus Birkhoff theorem does not trivially extend to infinite dimensional systems.

However it turns out that in the case of PDEs the nonlinearity has a particular structure. As a consequence it turns out that most of the monomials appearing in the nonlinearity are small and do not need to be eliminated through the normalization procedure. To illustrate this behavior consider the map

$$H^s([0,\pi]) \ni u \mapsto u^2 \in H^s([0,\pi]) \ , \tag{50}$$

which is the first term of the nonlinearity of the nonlinear wave equation (1). The use of Leibniz formula together with interpolation inequality allows one to prove the so called Tame inequality, namely

$$\left\|u^2\right\|_s \leq C_s \left\|u\right\|_s \left\|u\right\|_1 \ . \tag{51}$$

The key point is that, if u has only high frequency modes then its H^1 norm is much smaller than the H^s norm. Indeed, assume that, for some large M one has

$$u = \sum_{k\geq M} \hat{u}_k \mathbf{e}_k \tag{52}$$

then one has

$$\left\|u\right\|_1^2 = \sum_{k\geq M} k^2 \left|\hat{u}_k\right|^2 = \sum_{k\geq M} \frac{k^{2s}}{k^{2(s-1)}} \left|\hat{u}_k\right|^2 \leq \frac{1}{M^{2(s-1)}} \left\|u\right\|_s^2 \ . \tag{53}$$

Collecting (53) and (51) one gets

$$\left\|u^2\right\|_s \leq C_s \frac{1}{M^{s-1}} \left\|u\right\|_s^2 \ , \tag{54}$$

which is very small if M and s are large. In order to exploit such a condition one can proceed as follows: given $u \in H^s$ split it into high frequency and low frequency terms, namely write

$$u^S := \sum_{|k|<M} \hat{u}_k \mathbf{e}_k \ , \quad u^L := \sum_{|k|\geq M} \hat{u}_k \mathbf{e}_k \tag{55}$$

then one has

$$u^2 = (u^S)^2 + 2u^S u^L + (u^L)^2 \ , \tag{56}$$

the norms of these terms are bounded respectively by

$$\left\|u\right\|_s^2 \ , \quad \left\|u\right\|_s^2 \ , \quad \frac{1}{M^{s-1}} \left\|u\right\|_s^2$$

from which one sees that the last term can be considered small and is not relevant to the dynamics. Thus one could avoid to eliminate such a term from the nonlinearity. Correspondingly one will not have to consider small denominators involving frequencies with many small indexes.

To be able to exploit the tame property one has to ensure that it persists under the operations involved in the construction of the normal form, namely the computation of Poisson Brackets and the solution of the Homological equation. Now the stability of the tame property under Poisson Brackets is easy to check, while the verification of the stability under solution of the Homological equation is difficult and at present not known. For this reason one has to perform a more careful analysis. It turns out that it is convenient to understand the structure of the coefficients of the nonlinearity that ensure the Tame property, and to show that such a structure is invariant under the construction. The theory we develop is a variant of that developed by Delort and Szeftel in [DS06].

4 A Birkhoff normal form theorem for semilinear PDEs

4.1 Maps with localized coefficients and their properties

Having in mind the case of the nonlinear wave equation in S^d, consider the space ℓ^2_s of the sequences $q \equiv \{q_{jl}\}_{j\geq 0}^{1\leq l\leq l^*(j)}$, such that

$$\|q\|_s^2 := \sum_{1\leq j} |j|^{2s} \sum_{l=1}^{l^*(j)} |q_{jl}|^2 < \infty \ , \tag{57}$$

with a suitable $l^*(j)$.

Then define the projectors Π_j by

$$\Pi_j q := \sum_l q_{jl} \mathbf{e}_{jl} \tag{58}$$

(sum only over l), and the spaces $E_j := \Pi_j \ell^2_s$, which are independent of s.

The spaces $\mathscr{P}_s := \ell^2_s \times \ell^2_s \ni (p,q)$ will be used as phase spaces. We will also use the spaces $\mathscr{P}_\infty := \cap_s \mathscr{P}_s$ and $\mathscr{P}_{-\infty} := \cup_{s\in\mathbb{R}} \mathscr{P}_s$. It is useful to treat the p's and the q's exactly on an equal footing so we will use the notation

$$z_{jl} = q_{jl} \ , \quad z_{-jl} := p_{jl} \ , \quad j \geq 1 \ ,$$

correspondingly we will denote by z the set of all the variables and we will use the projector

$$\Pi_{-j} z := \sum_l p_{jl} \mathbf{e}_{jl} \ , \quad j \geq 1 \ . \tag{59}$$

Given an element $z \in \mathscr{P}_s$, one can write it as

$$z = \sum_{j\neq 0} \Pi_j z \ , \tag{60}$$

so that one has

$$\|z\|_s^2 = \sum_{j\neq 0} |j|^{2s} \left\|\Pi_j z\right\|^2 \tag{61}$$

where we defined

$$\left\|\Pi_j z\right\|^2 := \sum_{l=1}^{l^*(j)} z_{jl}^2 \ . \tag{62}$$

Let $f:\mathcal{P}_\infty\to\mathbb{R}$ be a smooth polynomial functions homogeneous of degree r. We can associate to f a symmetric multilinear map $\tilde{f}$, defined by the property

$$f(z)=\widetilde{f}(\underbrace{z,...,z}_{r-\text{times}}) \tag{63}$$

then we can write

$$f(z)=\sum_{j_1,...,j_r}\widetilde{f}(\Pi_{j_1}z,...,\Pi_{j_r}z)\ . \tag{64}$$

We will assume suitable localization properties of the norm of $\widetilde{f}(\Pi_{j_1}z,...,\Pi_{j_r}z)$ as a function of the indexes $j_1,...,j_r$.

Definition 6. Given a multi-index $j\equiv(j_1,...,j_r)$, let $(j_{i_1},j_{i_2},j_{i_3}...,j_{i_r})$ be a reordering of j such that

$$|j_{i_1}|\geq|j_{i_2}|\geq|j_{i_3}|\geq...\geq|j_{i_r}|\ .$$

We define $\mu(j):=|j_{i_3}|$ and

$$S(j):=\mu(j)+||j_{i_1}|-|j_{i_2}||\ . \tag{65}$$

Definition 7. Let $f:\mathcal{P}_\infty\to\mathbb{R}$ be a homogeneous polynomial of degree r. Let $\widetilde{f}$ be the associated multilinear form. We will say that f has localized coefficients if there exists $\nu\in[0,+\infty)$ such that $\forall N\geq 1$ there exists C_N such that $\forall z\in\mathcal{P}_\infty$ and any choice of the indexes $j_1,...,j_r$ the following inequality holds

$$\left|\widetilde{f}(\Pi_{j_1}z,...,\Pi_{j_r}z)\right|\leq C_N\frac{\mu(j)^{\nu+N}}{S(j)^N}\left\|\Pi_{j_1}z\right\|...\left\|\Pi_{j_r}z\right\|\ . \tag{66}$$

Definition 8. A function $f\in C^\infty(\mathcal{U},\mathbb{R})$ with $\mathcal{U}\subset\mathcal{P}_\infty$ will be said to have localized coefficients if

i) all the terms of its Taylor expansion have localized coefficients.
ii) for any s large enough there exists a neighborhood $\mathcal{U}^{(s)}$ of the origin in $\mathcal{P}_s$ such that $X_f\in C^\infty(\mathcal{U}^{(s)},\mathcal{P}_s)$.

Remark 13. In the case of $[0,\pi]$ the property (7) turns out to really be a property of the coefficients of the expansion of the nonlinearity on the basis in which the quadratic part is diagonal. To understand this point consider the case of a homogeneous polynomial dependent on q only. Write $q=\sum_j q_j\mathbf{e}_j$, then one has

$$f(q)=\sum_j\widetilde{f}(\mathbf{e}_{j_1},...,\mathbf{e}_{j_r})q_{j_1}...q_{j_r}=:\sum_j f_{j_1,...,j_r}q_{j_1}...q_{j_r} \tag{67}$$

then (66) is equivalent to

$$\left|f_{j_1,...,j_r}\right|\leq C_N\frac{\mu(j)^{\nu+N}}{S(j)^N}\ ,\quad\forall N\geq 1 \tag{68}$$

It is useful to extend the definition to polynomial maps taking value in $\mathscr{P}_s$.

Definition 9. Let $F : \mathscr{P}_\infty \to \mathscr{P}_{-\infty}$ be a polynomial map of degree r and let $\widetilde{F}$ be the associated multilinear form. We will say that F has localized coefficients if there exists $\nu \in [0, +\infty)$ such that

$$\|\Pi_i \widetilde{F}(\Pi_{j_1} z, \dots, \Pi_{j_r} z)\| \le C_N \frac{\mu(i,j)^{\nu+N}}{S(i,j)^N} \left\|\Pi_{j_1} z\right\| \dots \left\|\Pi_{j_r} z\right\| , \tag{69}$$

$$\forall z \in \mathscr{P}_\infty , \quad \forall N \ge 1 \tag{70}$$

Here we denoted by (i, j) the multi-index $(i, j_1, \dots, j_r)$.

Remark 14. It is easy to see that if a polynomial function has localized coefficients, then its Hamiltonian vector field has localized coefficients.

Remark 15. By the very definition of the property of localization of coefficients it is clear that any (finite) linear combination of functions or maps with localized coefficients has localized coefficients.

Remark 16. As it will be clear from the theory of sects. 4.3, and 6.3 it is quite easy to verify the property of localization of the coefficients.

The main properties of polynomials with localized coefficients are their smoothness, their stability under composition, linear combination and solution of the homological equation. In this subsection we will just state the corresponding results that will be proved in the appendix.

First one has that the vector field of a polynomial with localized coefficients has the tame property.

Theorem 2. *Let $F : \mathscr{P}_\infty \to \mathscr{P}_{-\infty}$ be a polynomial of degree r with localized coefficients, then there exists s_0 such that for any $s \ge s_0$ it extends to a smooth map from $\mathscr{P}_s$ to itself, moreover the following estimate holds*

$$\|F(z)\|_s \le C \|z\|_s \|z\|_{s_0}^{r-1} . \tag{71}$$

Corollary 1. *Let f be a function with localized coefficients, then the result of theorem 2 holds for its vector field.*

The composition of maps with localized coefficients has localized coefficients. Precisely

Theorem 3. *Let $f : \mathscr{P}_\infty \to \mathbb{R}$ be a polynomial of degree r_1 with localized coefficients, and let $G : \mathscr{P}_\infty \to \mathscr{P}_{-\infty}$ be a polynomial of degree r_2 with localized coefficients, then the polynomial*

$$\mathrm{d}f(z) G(z) \tag{72}$$

has localized coefficients.

Thus the strategy in order to verify the property of localization of the coefficients is to verify it for a few simple maps and then to use the composition (72) to show that it holds for more general maps. The precise example we have in mind is that where $f(u)=\int u^3$ and $G(u)=u^2$, in which $\mathrm{d}f(u)G(u)=3\int u^4$. Hence by iteration one gets that all polynomials in u have localized coefficients if $\int u^3$ has.

Moreover the following corollary holds.

Corollary 2. *The Poisson Bracket of two functions with localized coefficients has localized coefficients.*

In order to develop perturbation theory we need a suitable nonresonance condition. This is given by the following definition.

Definition 10. Fix a positive integer r. The frequency vector ω is said to fulfill the *property (r–NR)* if there exist $\gamma>0$, and $\alpha\in\mathbb{R}$ such that for any N large enough one has

$$\left|\sum_{j\geq 1}\omega_j K_j\right| \geq \frac{\gamma}{N^\alpha}\ , \tag{73}$$

for any $K\in\mathbb{Z}^\infty$, fulfilling $0\neq |K|:=\sum_j |K_j|\leq r+2$, $\sum_{j>N}|K_j|\leq 2$.

It is easy to see that under this condition one can solve the Homological equation. The precise statement is given by the following lemma.

Lemma 4. *Let f be a homogeneous polynomial of degree less or equal than r having localized coefficients. Let H_0 be given by* (49) *and assume that the frequency vector fulfills the condition r-NR. Consider the Homological equation*

$$\{H_0,\chi\}+f=Z\ . \tag{74}$$

Its solution χ, Z defined by (27) *has localized coefficients. In particular χ has localized coefficients.*

4.2 Statement of the normal form theorem and its consequences

Using the above results it is very easy to prove a version of the Birkhoff normal form theorem for PDEs.

Definition 11. With reference to a system of the form (2) with H_0 given by (49), the quantity

$$J_j:=\sum_l \frac{p_{jl}^2+q_{jl}^2}{2} \tag{75}$$

is called the total action of the modes with frequency ω_j.

Theorem 4. *Fix $r \geq 1$, assume that the nonlinearity H_P has localized coefficients and that the frequencies fulfill the nonresonance condition (r-NR), then there exists a finite s_r a neighborhood $\mathscr{U}_{s_r}^{(r)}$ of the origin in $\mathscr{P}_{s_r}$ and a canonical transformation $\mathscr{T} : \mathscr{U}_{s_r}^{(r)} \to \mathscr{P}_{s_r}$ which puts the system in normal form up to order $r+3$, namely*

$$H^{(r)} := H \circ \mathscr{T} = H_0 + Z^{(r)} + \mathscr{R}^{(r)} \tag{76}$$

where $Z^{(r)}$ and $\mathscr{R}^{(r)}$ have localized coefficients and

(i) $Z^{(r)}$ is a polynomial of degree $r+2$ which Poisson commutes with J_j for all j's, namely $\left\{J_j; Z^{(r)}\right\} \equiv 0$;

(ii) $\mathscr{R}^{(r)}$ has a small vector field, i.e.

$$\left\|X_{\mathscr{R}^{(r)}}(z)\right\|_{s_r} \leq C \left\|z\right\|_{s_r}^{r+2} , \quad \forall z \in \mathscr{U}_{s_r}^{(r)} ; \tag{77}$$

(iii) one has

$$\left\|z - \mathscr{T}_r(z)\right\|_{s_r} \leq C \left\|z\right\|_{s_r}^{2} , \quad \forall z \in \mathscr{U}_{s_r}^{(r)} . \tag{78}$$

An inequality identical to (10) *is fulfilled by the inverse transformation $\mathscr{T}_r^{-1}$.*

(iv) For any $s \geq s_r$ there exists a subset $\mathscr{U}_s^{(r)} \subset \mathscr{U}_{s_r}^{(r)}$ open in $\mathscr{P}_s$ such that the restriction of the canonical transformation to $\mathscr{U}_s^{(r)}$ is analytic also as a map from $\mathscr{P}_s \to \mathscr{P}_s$ and the inequalities (77) *and* (78) *hold with s in place of s_r.*

The proof is deferred to sect. 6.2.

In order to deduce dynamical consequences we fix the number r of normalization steps; moreover, it is useful to distinguish between the original variables and the variables introduced by the normalizing transformation. So, we will denote by $z = (p,q)$ the original variables and by $z' = (p',q')$ the normalized variables, i.e. $z = \mathscr{T}_r(z')$. More generally we will denote with a prime the quantities expressed in the normalized variables.

Proposition 1. *Under the same assumptions of theorem 4, $\forall s \geq s_r$ there exists ε_{*s} such that, if the initial datum fulfills*

$$\varepsilon := \left\|z_0\right\|_s < \varepsilon_{*s}$$

then one has

(i)

$$\left\|z(t)\right\|_s \leq 4\varepsilon \quad \textit{for} \quad |t| \leq \frac{1}{\varepsilon^r} \tag{79}$$

(ii)

$$\sum_j j^{2s} \left|J_j'(t) - J_j'(0)\right| \leq C\varepsilon^{M+3} \quad \textit{for} \quad |t| \leq \frac{1}{\varepsilon^{r-M}} , \quad M < r \tag{80}$$

and

$$\sum_j j^{2s}\left|J_j(t)-J_j(0)\right| \leq C\varepsilon^3 \quad for \quad |t|\leq \frac{1}{\varepsilon^r} \ . \tag{81}$$

(iii) If for each j the space E_j is one dimensional, then there exists a smooth torus $\mathbb{T}_0$ such that, $\forall M\leq r$

$$d_s(z(t),\mathbb{T}_0)\leq C\varepsilon^{(M+3)/2} \ , \quad for \quad |t|\leq \frac{1}{\varepsilon^{r-M}} \tag{82}$$

where $d_s(.,.)$ is the distance in $\mathscr{P}_s$.

4.3 Application to the nonlinear wave equation.

The aim of this section is to verify the assumptions of theorem 4 in the model problems we are considering.

We start by the property of localization of the coefficients. The main step consists in verifying the property for the Hamiltonian function

$$f(u):=\int_D u^3(x)\mathrm{d}x \ ; \tag{83}$$

the corresponding multilinear form is given by

$$\widetilde{f}(u_1,u_2,u_3):=\int_D u_1(x)u_2(x)u_3(x)\mathrm{d}x \ , \tag{84}$$

so we have to estimate such a quantity when $u_j\in E_{n_j}$, namely the eigenspace of $-\Delta$ corresponding to the eigenvalue λ_{n_j}. We have the following theorem.

Proposition 2. *Let E_n be the eigenspace of $-\Delta$ associated to the eigenvalue λ_n, then for any $N\geq 1$ there exists C_N such that one has*

$$\left|\int_D u_{n_1}(x)u_{n_2}(x)u_{n_3}(x)\mathrm{d}x\right| \leq C_N\frac{\mu(n)^{N+\nu}}{S(n)^N}\left\|u_{n_1}\right\|_{L^2}\left\|u_{n_2}\right\|_{L^2}\left\|u_{n_3}\right\|_{L^2} \tag{85}$$

for all $u_{n_j}\in E_{n_j}$.

A simple strategy to obtain the proof consists in considering the quantity (84) as the matrix elements of index n_1,n_2 (on the basis of the eigenvectors of the Laplacian) of the operator of multiplication by u_{n_3}. The actual proof is deferred to sect.6.3.

Corollary 3. *The nonlinearity given by the second integral in* (41) *has localized coefficient.*

Proof. This is a consequence of proposition 2 and of theorem 3. Indeed a term in the Taylor expansion of $\int_D F(u)\mathrm{d}x$ is a multiple of

$$t_k(u):=\int_{S^d} u^k(x)\mathrm{d}x \tag{86}$$

and one has $t_k(u) = C\mathrm{d}t_2(u)T_{k-1}(u)$, where $T_{k-1}(u) = u^{k-1}$. Then such a quantity has localized coefficients by theorem 3. □

In order to apply theorem 4 to the nonlinear wave equation (1) there remains to verify the nonresonance condition $(r-NR)$. To this end consider the frequencies

$$\omega_{j+1} = \sqrt{j(j+d-1)+\mu^2} \tag{87}$$

then we have the following

Theorem 5. *There exists a zero measure set $S \subset \mathbb{R}$ such that, if $\mu \in \mathbb{R}-S$, then the frequencies* (87) *fulfill the condition* $(r-NR)$ *for any r.*

The proof was given in [Bam03] (see also [BG06]), and for the sake of completeness it is repeated in Appendix 6.4.

Thus the main theorem and its corollaries apply to the nonlinear wave equation both in the case of $[0,\pi]$ and in the case of the d dimensional sphere.

Remark 17. A particular consequence of this theory is that it allows one to ensure existence of smooth solutions of the nonlinear wave equation on the sphere for times of order ε^{-r}. It has to be emphasized that when $d>1$ local existence is ensured only in H^s, with $s>1$, so that the energy norm is useless in order to deduce estimate of the existence times of solutions. At present the method of Birkhoff normal form is the only one allowing one to improve the times given by the local existence theory.

5 Discussion

First I would like to mention that, as shown in [BG06], theorem 4 is a theorem that allows one to deal with quite general *semilinear* equations in *one space dimension*.

The limitation to semilinear equation is evident in theorem 4. Thus in particular all the equations with nonlinearity involving derivatives are excluded from the present theory. It would be of major interest to have a theory valid also for some quasilinear equations, since most physical models have nonlinearities involving derivatives. Very little is known on quasilinear problems. At present the only known result is that of [DS04] (and a recent extension by Delort), where only one step of normal form was performed for the quasilinear wave equation. It would be very interesting to understand how to iterate the procedure developed in such papers.

The limitation to 1-space dimension is more hidden. Actually it is hidden in the nonresonance condition. Indeed its verification is based on the asymptotic behavior of the frequencies: the nonresonance condition is typically satisfied only if the frequencies grow at infinity at least as $\omega_j \sim j$. As it was shown in the example of the nonlinear wave equation on the sphere, the possible multiplicity of the frequencies is not a problem. The theory easily extends to the case where the differences between couples of frequencies accumulate only at a discrete subset of $\mathbb{R}$. The understanding of the structure of the frequencies in higher dimension is surely a key point for the extension of the theory to higher dimensions.

Finally I would like to mention the fact that all known applications of the considered theory pertain equations on compact manifolds, however in principle the theory applies to smooth perturbations of linear system with discrete spectrum. A nice example of such a kind of systems is the Gross Pitaevskii equation. It would be interesting to show that such an equation fulfills the assumption of theorem 4. This could be interesting also in connection with the study of the blow up phenomenon.

6 Proofs

6.1 Proof of the properties of functions with localized coefficients

Lemma 5. *Let $z \in \mathscr{P}_s$ with $s > \nu + 1/2$ then there exists a constant C_s such that*

$$\sum_{j\neq 0} |j|^{\nu} \left\|\Pi_j z\right\| \leq C_s \left\|z\right\|_s \tag{88}$$

Proof. One has

$$\sum_{j\neq 0} |j|^{\nu} \left\|\Pi_j z\right\| \leq \sum_j |j|^s \frac{\left\|\Pi_j z\right\|}{|j|^{s-\nu}} \leq \sqrt{\sum_j \frac{1}{|j|^{2(s-\nu)}}} \sqrt{\sum_j |j|^{2s} \left\|\Pi_j z\right\|^2}$$

which is the thesis. $\square$

Proof of theorem 2. Write explicitly the norm of $F(z)$. One has

$$\|F(z)\|_s^2 = \sum_l |l|^{2s} \left\| \sum_{j_1,\dots,j_r} \Pi_l \tilde{F}(\Pi_{j_1} z, \dots, \Pi_{j_r} z) \right\|^2 . \tag{89}$$

In what follows, to simplify the notation we will write

$$a_j := \left\|\Pi_j z\right\| .$$

One has

$$\left\| \sum_{j_1,\dots,j_r} |l|^s \Pi_l \tilde{F}(\Pi_{j_1} z, \dots, \Pi_{j_r} z) \right\| \leq C \sum_{j_1,\dots,j_r} |l|^s \frac{\mu(j,l)^{\nu+N}}{S(j,l)^N} a_{j_1} \dots a_{j_r} \tag{90}$$

Since this expression is symmetric in $j_1, \dots j_r$ the r.h.s. of (90) is estimated by a constant times the sum restricted to ordered multi-indexes, namely indexes such that $|j_1| \leq |j_2| \leq \dots \leq |j_r|$. Moreover, in order to simplify the notations *we will restrict to* *the case of positive indexes*. To estimate (90) remark that for ordered multi-indexes one has

$$l\frac{\mu(j,l)}{S(j,l)} \le 2j_r\ . \tag{91}$$

Indeed, if $l \le 2j_r$ this is obvious ($\mu/S < 1$ by the very definition), while, if $l > 2j_r$ one has $S(j,l) \ge |l - j_r| > l/2$, and therefore

$$l\frac{\mu(j,l)}{S(j,l)} \le \mu(j,l) \le 2j_r\ .$$

Remark now that, by the definition of S one has

$$S(j,l) \ge \begin{cases} 1+|j_r - l| & \text{if } l \ge j_{r-1} \\ \mu(j,l) + j_r - j_{r-1} \ge l + j_r - j_{r-1} & \text{if } l < j_{r-1} \end{cases}$$

Thus define $\hat{S}(j,l) := \min\{1+|j_r - l|, l + j_r - j_{r-1}\}$ and remark that $S(j,l) \ge \hat{S}(j,l)$. Remark also that $\mu(j,l) \le j_{r-1}$. So it follows that (90) is smaller than (a constant times)

$$\sum_{j_1,\dots,j_r} j_r^s \frac{\mu(j,l)^{N'+\nu}}{\hat{S}(j,l)^{N'}} a_{j_1}\dots a_{j_r} \le \sum_{j_1,\dots,j_r} j_r^s \frac{j_{r-1}^{N'+\nu}}{\hat{S}(j,l)^{N'}} a_{j_1}\dots a_{j_r} \tag{92}$$

$$\le \|z\|_{s_1}^{r-2} \sum_{j_{r-1},j_r} j_r^s \frac{j_{r-1}^{N'+\nu}}{\hat{S}(j,l)^{N'}} a_{j_{r-1}}\ a_{j_r} \tag{93}$$

where we denoted $N' := N - s$ and we used lemma 5; we denoted by s_1 a number such that $s_1 > 1/2$. Inserting in (89) one gets

$$\|F(z)\|_s^2 \le \|z\|_{s_1}^{2(r-2)} \sum_l \left(\sum_{j_{r-1},j_r} j_r^s \frac{j_{r-1}^{N'+\nu}}{\hat{S}(j,l)^{N'}} a_{j_{r-1}}\ a_{j_r} \right)^2$$

$$= \|z\|_{s_1}^{2(r-2)} \sum_l \left(\sum_{j_{r-1}} j_{r-1}^{N'+\nu} a_{j_{r-1}} \sum_{j_r} j_r^s \frac{a_{j_r}}{\hat{S}(j,l)^{N'/2}}\ \frac{1}{\hat{S}(j,l)^{N'/2}} \right)^2$$

$$\le \|z\|_{s_1}^{2(r-2)} \sum_l \left(\sum_{j_{r-1}} j_{r-1}^{N'+\nu} a_{j_{r-1}} \sqrt{\sum_{j_r} j_r^{2s} \frac{a_{j_r}^2}{\hat{S}(j,l)^{N'}}} \sqrt{\sum_{j_r} \frac{1}{\hat{S}(j,l)^{N'}}} \right)^2$$

Now the last sum in j_r is finite provided $N' > 1$. Remark now that $\hat{S}(j,l) \ge \check{S}(j_r,l) := \min\{1+|l - j_r|, l\}$ (independent of j_{r-1}), and therefore the above quantity is estimated by a constant times

$$\|z\|_{s_1}^{2(r-2)}\sum_l\left(\sum_{j_{r-1}} j_{r-1}^{N'+\nu}a_{j_{r-1}}\sqrt{\sum_{j_r} j_r^{2s}\frac{a_{j_r}^2}{\breve{S}(j,l)^{N'}}}\right)^2 \tag{94}$$

$$= \|z\|_{s_1}^{2(r-2)}\sum_{j_r} j_r^{2s}a_{j_r}^2\sum_l\frac{1}{\breve{S}(j,l)^{N'}}\left(\sum_{j_{r-1}} j_{r-1}^{N'+\nu}a_{j_{r-1}}\right)^2 \tag{95}$$

$$\leq C\,\|z\|_s^2\,\|z\|_{s_1}^{2(r-2)}\,\|z\|_{s_0}^2 \tag{96}$$

where s_0 is such that $s_0 > N'+\nu+1/2$. Choosing $s_1 \leq s_0$ and estimating $\|z\|_{s_1}$ with $\|z\|_{s_0}$ one gets the thesis. $\square$

Proof of theorem 3. First remark that the multilinear form associated to the polynomial $\mathrm{d}f(z)G(z)$ is given by the symmetrization of

$$r_1\widetilde{f}(z^{(1)},\dots,z^{(r_1-1)},\widetilde{G}(z^{(r_1)},\dots,z^{(r_1+r_2-1)}))\ . \tag{97}$$

We will estimate the coefficients of this multilinear function. This will give the result. Forgetting the irrelevant constant r_1, the quantity to be estimated is

$$\widetilde{f}(\Pi_{j_1}z,\dots,\Pi_{j_{r_1-1}}z,\widetilde{G}(\Pi_{i_1}z,\dots,\Pi_{i_{r_2}}z)) \tag{98}$$

$$= \sum_l \widetilde{f}(\Pi_{j_1}z,\dots,\Pi_{j_{r_1-1}}z,\Pi_l\widetilde{G}(\Pi_{i_1}z,\dots,\Pi_{i_{r_2}}z)) \tag{99}$$

$$\leq C_{N,N'}\sum_l\frac{\mu^{\nu_1+N}(j,l)}{S(j,l)^N}\frac{\mu^{\nu_2+N'}(i,l)}{S(i,l)^{N'}}\left\|\Pi_{j_1}z\right\|\dots\left\|\Pi_{i_{r_2}}z\right\| \tag{100}$$

Thus it is enough to estimate

$$\sum_l\frac{\mu^{\nu_1+N}(j,l)}{S(j,l)^N}\frac{\mu^{\nu_2+N'}(i,l)}{S(i,l)^{N'}} \tag{101}$$

This is the heart of the proof.

In order to simplify the notation we will restrict to he case $r_1-1=r_2=r$. Due to the symmetry of this estimate we will restrict the case of ordered indexes, that can also be assumed to be positive, so that one has $j_r\geq j_{r-1}\geq\dots\geq j_1$ and similarly for i.

All along this proof we will use the notation

$$\tilde{S}(j) := j_r - j_{r-1} \equiv S(j)-\mu(j)$$

We have to distinguish two cases.

First case $j_r\geq i_r\geq j_{r-1}$.

The proof of this first case is (up to minor changes) equal to that given in [Gré06].

Take $N' = N$, then before estimating (101), we need to estimate the general term of the sum. So we collect a few facts on it.

The main relation we need is

$$\tilde{S}(i,j) \leq \tilde{S}(i,l) + \tilde{S}(j,l)\ . \tag{102}$$

This will be established by writing explicitly all the involved quantities as l varies. So, first remark that $\tilde{S}(i,j) = j_r - i_r$. Then one has

$$\tilde{S}(i,l) = \begin{cases} i_r - i_{r-1} & \text{if } l \leq i_{r-1} \\ |i_r - l| & \text{if } l > i_{r-1} \end{cases} \quad , \quad \tilde{S}(j,l) = \begin{cases} j_r - j_{r-1} & \text{if } l \leq j_{r-1} \\ |j_r - l| & \text{if } l > j_{r-1} \end{cases}$$

which gives

$$\tilde{S}(i,l) + \tilde{S}(j,l) = \begin{cases} i_r - i_{r-1} + j_r - j_{r-1} \geq j_r - j_{r-1} \geq j_r - i_r & \text{if } l \leq i_{r-1} \\ l - i_r + j_r - j_{r-1} \geq j_r - j_{r-1} \geq j_r - i_r & \text{if } i_{r-1} < l \leq j_{r-1} \\ |i_r - l| + |j_r - l| \geq |j_k - l| \geq j_k - i_k & \text{if } j_{k-1} < l \leq i_r \\ |i_r - l| + |j_r - l| = j_r - l + l - i_r & \text{if } i_r < l \leq j_r \\ |i_r - l| + |j_r - l| \geq l - i_r \geq j_r - i_r & \text{if } j_r \leq l \end{cases}$$

from this (102) follows.

One also has

$$\mu(j,l) \leq \mu(i,j)\ , \quad \mu(i,l) \leq \mu(i,j)\ . \tag{103}$$

Thus

$$\frac{S(i,j)}{\mu(i,j)} = 1 + \frac{\tilde{S}(i,j)}{\mu(i,j)} \leq 1 + \frac{\tilde{S}(i,l) + \tilde{S}(j,l)}{\mu(i,j)} \leq 1 + \frac{\tilde{S}(i,l)}{\mu(i,l)} + \frac{\tilde{S}(l,j)}{\mu(l,j)} < \frac{S(i,l)}{\mu(i,l)} + \frac{S(l,j)}{\mu(l,j)}$$

From this one has

$$\frac{\mu(i,j)}{S(i,j)} \leq \frac{1}{2} \min\left\{ \frac{\mu(i,l)}{S(i,l)}, \frac{\mu(l,j)}{S(l,j)} \right\}\ . \tag{104}$$

Separate the sum over those l such that $\frac{\mu(i,l)}{S(i,l)} > \frac{\mu(l,j)}{S(l,j)}$ and that over its complement. Let L_1 be the first set. Then one has

$$\begin{aligned} \sum_{l \in L_1} \frac{\mu^{\nu_1+N}(j,l)}{S(j,l)^N} \frac{\mu^{\nu_2+N}(i,l)}{S(i,l)^N} &\leq \sum_{l \geq 1} 2^{N-1-\varepsilon} \mu(j,l)^{\nu_1} \frac{\mu(i,j)^{N-1-\varepsilon}}{S(i,j)^{N-1-\varepsilon}} \frac{\mu(i,l)^{1+\varepsilon+\nu_2}}{S(i,l)^{1+\varepsilon}} \\ &\leq C \frac{\mu(i,j)^{N+\nu_1+\nu_2}}{S(i,j)^{N-1-\varepsilon}}\ . \end{aligned}$$

Acting in the same way for the case of L_1^c one concludes the proof in the first case.

Second case $j_r \geq j_{r-1} > i_r$. Here it is easy to see that (102) still holds. However, in some cases it happens that the equation

$$\mu(j,l) \leq \mu(i,j) \tag{105}$$

is violated. When (105) holds the proof of the first case extends also to the present case. So let us consider only the case where (105) is violated. We claim that in this case one has

$$\frac{\mu(j,l)}{\tilde{S}(i,l)} \leq 2\mu(i,j) \ . \tag{106}$$

To prove (106) we distinguish two cases

i) $j_{r-2} \leq i_r \leq j_{r-1} \leq j_r$.
   Then (105) is violated when $i_r < l \leq j_{r-1}$. In this case one has

$$\tilde{S}(i,l) = l - i_r \tag{107}$$

   It follows that

$$\frac{\mu(j,l)}{\tilde{S}(i,l)} \frac{1}{\mu(i,j)} = \frac{l}{l-i_r} \frac{1}{i_r}$$

   which is easily seen to be smaller than 2 (for example write $l = i_r + \delta$, then the relation becomes evident).

ii) $i_r < j_{r-2} \leq j_{r-1} \leq j_r$. Here (105) is violated when $j_{r-2} < l \leq j_{r-1}$. It is easy to see that also in this case (107) holds. Then

$$\frac{\mu(j,l)}{\tilde{S}(i,l)} \frac{1}{\mu(i,j)} = \frac{l}{l-i_r} \frac{1}{j_r} \leq \frac{l}{l-i_r} \frac{1}{i_r}$$

   from which (106) still follows.

It is now easy to conclude the proof. Take $N' = 2N + \nu_2$, then, using (106) one has

$$\begin{aligned} &\frac{\mu(i,l)^{\nu_1+2N+\nu_2}}{S(i,l)^{2N+\nu_2}} \frac{\mu(j,l)^{N+\nu_2}}{S(j,l)^N} \\ \leq &\frac{\mu(i,l)^{\nu_1+2N+\nu_2}}{S(i,l)^N} \left(\frac{\mu(j,l)}{S(i,l)}\right)^{N+\nu_2} \frac{1}{S(j,l)^N} \\ &\leq \frac{\mu(i,l)^{\nu_1+2N+\nu_2}}{S(i,l)^N} \frac{\mu(i,j)^{N+\nu_2}}{S(j,l)^N} \end{aligned}$$

From this, following the proof given in the first case it is easy to prove that

$$\frac{S(i,j)}{\mu(i,j)} \leq \frac{S(i,l)}{\mu(i,l)} + \frac{S(j,l)}{\mu(i,j)}$$

and to conclude the proof in the same way as in the first case. □

Proof of lemma 4. Consider the polynomial f and expand it in Taylor series. Introduce now the complex variables (23). Remark that this is a linear change of variable so it does not change the degree of a polynomial. Remark that the change of variables does not mix the different spaces $E_j \times E_j$. It follows that if a polynomial has localized coefficients in terms of the real variables p, q it has also localized coefficients when written in terms of the complex variables, i.e. it fulfills (66) with z_j

which is either ξ_j or η_j. Remark that the converse is also true. Now, Z is the sum of some of the coefficients of f so it is clear that its coefficients are still localized. In order to estimate χ, remark first that, in the particular case where

$$f(z) \equiv \widetilde{f}(\Pi_{j_1}\xi, ..., \Pi_{j_{r_1}}\xi, \Pi_{l_1}\eta, ..., \Pi_{l_{r_2}}\eta)$$

(no summation over j, l) one has

$$\{H_0, f\} = i(\omega_{j_1} + ... + \omega_{j_{r_1}} - \omega_{l_1} - ... - \omega_{l_{r_2}}) f \tag{108}$$

It follows that in the case of general f the function χ solving the homological equation can be rewritten as

$$\chi(\xi, \eta) := \sum_{jl} \frac{\widetilde{f}(\Pi_{j_1}\xi, ..., \Pi_{j_{r_1}}\xi, \Pi_{l_1}\eta, ..., \Pi_{l_{r_2}}\eta)}{i(\omega_{j_1} + ... + \omega_{j_{r_1}} - \omega_{l_1} - ... - \omega_{l_{r_2}})} \tag{109}$$

where the sum runs over the indexes such that the denominators do not vanish. Now, it is easy to verify that by condition (r-NR) the denominators are bounded from below by $\gamma/\mu(j,l)^{\alpha}$. So χ fulfills the estimate (66) with ν substituted by $\nu + \alpha$, if f does with ν. □

6.2 *Proof of the Birkhoff normal form theorem 4 and of its dynamical consequences.*

In this section we will fix s large enough and work in $\mathscr{P}_s$. Here $B_R \subset \mathscr{P}_s$ will denote the open ball of radius R with center at the origin in $\mathscr{P}_s$. Moreover all along this section $\mathscr{H}_j$ will denote the set of homogeneous polynomials of degree $j+2$ **having a Hamiltonian vector field which is smooth as map from $\mathscr{P}_s$ to itself.** Finally, along this section we will omit the index s from the norm, thus we will simply denote $\|.\| := \|.\|_s$.

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

Lemma 6. *Let $\chi \in \mathscr{H}_j$, ($j \geq 1$) be a polynomial. Denote by ϕ^t the flow of the corresponding vector field. Denote also*

$$\bar{t} = \bar{t}(R, \delta) := \inf_{z \in B_R} \sup \left\{ t > 0 : \phi^t(z) \in B_{R+\delta} \text{ and } \phi^{-t}(z) \in B_{R+\delta} \right\}$$

(minimum escape time of $\phi^t(z)$ from $B_{R+\delta}$). Then one has

$$\bar{t} \geq \frac{\delta}{2\left\|X_\chi\right\| R^{j+1}} \tag{110}$$

where $\left\|X_\chi\right\|$ is the norm defined in remark 3. Moreover for any t, such that $|t| \leq \overline{t}$ and any $z \in B_R$ one has

$$\left\|\phi^t(z) - z\right\| \leq |t| R^{j+1} \left\|X_\chi\right\| \tag{111}$$

Proof. First remark that, by the definition of $\overline{t}$ one has that there exists $\overline{z} \in B_R$ such that $\left\|\phi^{\pm\overline{t}}(\overline{z})\right\| = R + \delta$. Assume by contradiction $\overline{t} < \frac{\delta}{2\left\|X_\chi\right\| R^{j+1}}$, then, since for any t with $|t| < \overline{t}$ one has $\phi^t(\overline{z}) \in B_{R+\delta}$. It follows that

$$\begin{aligned}\left\|\phi^{\overline{t}}(\overline{z})\right\| &\leq \|\overline{z}\| + \left\|\phi^{\overline{t}}(\overline{z}) - \overline{z}\right\| = \|\overline{z}\| + \left\|\int_0^{\overline{t}} \frac{d}{ds}\phi^s(\overline{z})\mathrm{d}s\right\| \\ &\leq R + \int_0^{\overline{t}} \left\|X_\chi(\phi^s(\overline{z}))ds\right\| \leq R + |\overline{t}| R^{j+1} \left\|X_\chi\right\| ,\end{aligned}$$

from which $R + \delta \leq R + \delta/2$ which is absurd. $\square$

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

Corollary 4. *Fix arbitrary R and δ, then the map*

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

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

Proof of theorem 4. The proof proceeds as in the finite dimensional case. The only fact that has to be ensured is that at any step the functions involved in the construction have localized coefficients. By lemma 4 the solution χ_{r+1} of the homological equation (33) has localized coefficients. Thus, by theorem 2 its vector field is smooth on a space $\mathscr{P}_{s_{r+1}}$. This determines the index s_{r+1} of the space with minimal smoothness in which the transformation $\mathscr{T}_{r+1}$ is defined. By corollary 4 χ_{r+1} generates an analytic flow. As in the finite dimensional case we use it to generate the Lie transform. Then $H^{(r+1)}$ is still given by eqs. (36-39). Remark now that given a Hamiltonian function f, the Hamiltonian vector field of $f \circ \phi_{r+1}$ is given by

$$X_{f\circ\phi_{r+1}}(z) = \mathrm{d}\phi_{r+1}^{-1}(\phi_{r+1}(z)) X_f(\phi_{r+1}(z)) \tag{112}$$

so that the Hamiltonian vector fields of the terms (35), (36), (38), (39) are smooth. To ensure the smoothness of the vector field of (37) write

$$\ell(z) := H_0 \circ \phi - H_0 - \{\chi_{r+1}, H_0\}$$

and remark that

$$H_0(\phi_{r+1}(z)) - H_0(z) = \int_0^1 \frac{d}{dt} H_0(\phi_{r+1}^t(z))dt$$
$$= \int_0^1 \{\chi_{r+1}, H_0\}(\phi_{r+1}^t(z))dt = \int_0^1 (H_{r+1}^{(r)}(\phi_{r+1}^t(z)) - Z_{r+1}(\phi_{r+1}^t(z))dt \ ,$$

where we used the homological equation to calculate $\{\chi_{r+1}, H_0\}$. Denote again $G := H_{r+1}^{(r)} - Z_{r+1}$, then one has

$$\ell(z) = \int_0^1 (G(\phi_{r+1}^t(z)) - G(z))dt \ ,$$

from which the smoothness of the vector field of (37) immediately follows. Since the Taylor expansion of the terms (36-39) can be computed using (14), by corollary 15 one has that all these functions have localized coefficients. Then, as in the finite dimensional case the terms (36-39) have a vector field with a zero of order at least $r+3$ which ensures the estimate of the remainder.

We show now that the normal form $Z^{(r)}$ commutes with all the J_j. To this end remark that, by construction, the normal form contains only resonant monomials, i.e. monomials $\xi^L \eta^J$ with

$$0 = \sum_{jl} \omega_j (J_{jl} - L_{jl}) = \sum_j \omega_j \left(\sum_l (J_{jl} - L_{jl}) \right) \ . \tag{113}$$

Now the nonresonance condition implies

$$\left(\sum_l (J_{jl} - L_{jl}) \right) = 0 \quad \forall j \ .$$

It follows

$$\{J_j, \xi^L \eta^J\} = \mathrm{i} \left[\sum_l (J_{jl} - L_{jl}) \right] \xi^L \eta^J = 0 \tag{114}$$

which is the wonted property.

Finally the estimate (78) of the deformation can be obtained exactly as in the finite dimensional case. □

Proof of proposition 1. We start by (i). Assume that ε is so small that $B_{3\varepsilon} \subset \mathscr{U}_s^{(r)}$; perform the normalizing transformation. Remark that, by (78), one has $z_0' \in B_{2\varepsilon} \subset \mathscr{U}_s^{(r)}$. Define $F(z) := \sum_j |j|^{2s} J_j \equiv \|z\|_s^2$, then, as far as $\|z'(t)\|_s \le 3\varepsilon$ one has

$$\begin{aligned} \left| F(z'(t)) - F(z_0') \right| &= \left| \int_0^t \left\{ H^{(r)}, F \right\} (z'(s)) ds \right| \\ &\le \int_0^t \left| \left\{ \mathscr{R}^{(r)}, F \right\} (z'(s)) \right| ds \le |t| C \varepsilon^{r+3} \le C \varepsilon^3 \end{aligned} \tag{115}$$

where the last inequality holds for the times (79). To conclude the proof of (79) it is enough to show that, for the considered times one actually has $z'(t) \in B_{3\varepsilon}$. To this end we follow the scheme of the proof of Lyapunov's theorem: define

$$\overline{t}:=\sup\left\{t>0\ :\ \left\|z'(t)\right\|_s<3\varepsilon \text{ and } \left\|z'(-t)\right\|_s<3\varepsilon\right\}$$

To fix ideas assume that the equality is realized for $t=\overline{t}$ Assume by contradiction that $\overline{t}<\varepsilon^{-r}$, then one can use (115) which gives

$$\left\|z'(\overline{t})\right\|^2=9\varepsilon^2=F(\overline{t})\leq F(z_0')+\left|F(z'(t))-F(z_0')\right|\leq 4\varepsilon^2+C\varepsilon^3\ , \tag{116}$$

which is impossible for ε small enough.

We come to (ii). First remark that

$$\dot{J}_j'=\sum_l\left(-p_{jl}'\frac{\partial\mathscr{R}^{(r)}}{\partial q_{jl}'}+q_{jl}'\frac{\partial\mathscr{R}^{(r)}}{\partial p_{jl}'}\right)\ ,$$

so that

$$\sum_j j^{2s}\left|\dot{J}_j'\right|=\sum_{jl}j^{2s}\left|-p_{jl}'\frac{\partial\mathscr{R}^{(r)}}{\partial q_{jl}'}+q_{jl}'\frac{\partial\mathscr{R}^{(r)}}{\partial p_{jl}'}\right| \tag{117}$$

$$\leq\left(\sum_{jl}j^{2s}({p_{jl}'}^2+{q_{jl}'}^2)\right)^{1/2}\left(\sum_{jl}j^{2s}\left(\left|\frac{\partial\mathscr{R}^{(r)}}{\partial q_{jl}'}\right|^2+\left|\frac{\partial\mathscr{R}^{(r)}}{\partial p_{jl}'}\right|^2\right)\right)^{1/2} \tag{118}$$

$$\leq\left\|z'\right\|_s\left\|X_{\mathscr{R}^{(r)}}(z')\right\|_s\leq C\left\|z'\right\|_s^{r+3} \tag{119}$$

which implies (80).

To prove (81) write

$$\left|J_j(t)-J_j(0)\right|\leq\left|J_j(z(t))-J_j(z'(t))\right|+\left|J'(t)-J'(0)\right|+\left|J_j(z_0)-J_j(z_0')\right|\ . \tag{120}$$

The contribution of the middle term is estimated by (80). To estimate the contribution of the first and the last term write

$$j^{2s}\left|q_{jl}^2-{q_{jl}'}^2\right|\leq j^{2s}\left(2|q_{jl}'|\ |q_{jl}-q_{jl}'|+|q_{jl}-q_{jl}'|^2\right) \tag{121}$$

adding the corresponding estimate for the p variables and summing over jl one gets the thesis.

We come to (iii). In the considered case J_j reduces to I_j, so the actions are individually conserved. In this proof we omit the index l which would take only the value 1. Denote $\overline{I_j}':=\frac{{p_j'}^2(0)+{q_j'}^2(0)}{2}$ and define the torus

$$\mathbb{T}_0':=\left\{z'\in\mathscr{P}_s\ :\ I_j(z')=\overline{I_j}\ \right\}$$

One has

$$d(z'(t),\mathbb{T}_0')\leq\left[\sum_j j^{2s}\left|\sqrt{I_j'(t)}-\sqrt{\overline{I_j}}\right|^2\right]^{1/2} \tag{122}$$

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

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

Thus, using (119), one has that

$$\left[d(z'(t),\mathbb{T}'_0)\right]^2 \leq \sum_j j^{2s}|I'_j(t)-\overline{I}_j| \leq C\varepsilon^{M+3}$$

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

$$d(z(t),\mathbb{T}_0) = d(\mathscr{T}_r(z'(t)),\mathscr{T}_r(\mathbb{T}'_0)) \leq Cd(z'(t),\mathbb{T}'_0) \leq C\varepsilon^{\frac{M+3}{2}}\ .$$

□

6.3 *Proof of proposition 2 on the verification of the property of localization of coefficients*

In this subsection we will prove the property of localization of coefficients for the function $u \mapsto \int u^3$ in the case where the basis used for the definition (66) is the basis of the eigenfunction of general second order elliptic operator. Thus the present theory directly applies also to the case of the equation

$$u_{tt} - u_{xx} + Vu = f(x,u)$$

with Neumann boundary conditions on $[0,\pi]$. The case of Dirichlet boundary conditions can also be covered by a minor variant (indeed in such a case the function $u \mapsto \int u^3$ has to be substituted by the function $u \mapsto \int u^4$.

Thus consider a second order elliptic operator P, which is L^2 self adjoint. This means that we assume that in any coordinate system there exist smooth functions $V_\alpha(x)$, $\alpha \in \mathbb{N}^d$ such that $P = \sum_{|\alpha|\leq 2} V_\alpha \partial^\alpha$, where we used a vector notation for the derivative. Moreover we will assume that

$$\|u\|_{s+2} \leq \|Pu\|_s\ .$$

Then, by L^2 symmetry, one gets

$$\|u\|_s \leq \left\|P^{s/2}u\right\|_0\ , \tag{123}$$

(where $P^{s/2}$ is defined spectrally). We will denote by $D(P^k)$ the domain of P^k. Finally, denote by λ_n the sequence of the eigenvalues of P counted without multiplicity (i.e. in such a way that $\lambda_{n+1} > \lambda_n$). We will assume that the eigenvalues of P behave as $\lambda_n \sim n^2$. We will denote by E_n the eigenspace of P relative to λ_n.

Let A be a linear operator which maps $D(P^k)$ into itself for all $k \geq 0$, and define the sequence of operators

$$A_N := [P, A_{N-1}] \ , \quad A_0 := A \ . \tag{124}$$

Lemma 7. *Let P be as above and let $u_j \in E_{n_j}$. Then, for any $N \geq 0$ one has*

$$|\langle Au_1; u_2\rangle| \leq \frac{1}{|\lambda_{n_1} - \lambda_{n_2}|^N} |\langle A_N u_1; u_2\rangle| \tag{125}$$

Proof. One has

$$\begin{aligned}\langle A_1 u_1; u_2\rangle &= \langle [A,P] u_1; u_2\rangle = \langle APu_1; u_2\rangle - \langle PAu_1; u_2\rangle \\ &= \lambda_{n_1} \langle Au_1; u_2\rangle - \langle Au_1; Pu_2\rangle = (\lambda_{n_1} - \lambda_{n_2}) \langle Au_1; u_2\rangle\end{aligned}$$

Eq. (125) follows applying the above equality to the operator $A_N := [P, A_{N-1}]$ and using an induction argument. $\square$

To conclude the proof we have to estimate the matrix elements of A_N, i.e. the r.h.s. of (125). To this end we need a few remarks and lemma.

Remark 18. Consider two d-dimensional multi-indexes α and β and define

$$\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!} \ ,$$

with the convention that it is 0 if $\beta_j > \alpha_j$ for some j. One has

$$\partial^\alpha (uv) = \sum_\beta \binom{\alpha}{\beta} \partial^\beta u \partial^{\alpha-\beta} v \ . \tag{126}$$

Remark 19. Let $A := a(x)\partial^\alpha$ and $B := b(x)\partial^\beta$ with a and b smooth functions. Then one has

$$[A, B] = \sum_{\gamma_j \leq \alpha_j + \beta_j} \left[a \binom{\beta}{\gamma} \partial^\gamma b - b \binom{\alpha}{\gamma} \partial^\gamma a \right] \partial^{\alpha+\beta-\gamma} \ . \tag{127}$$

Lemma 8. *Choose a coordinate system, let $A = a_0(x)$ be a multiplication operator, then one has*

$$A_N = \sum_{|\alpha| \leq N} c_\alpha^{(N)} \partial^\alpha \tag{128}$$

with $c_\alpha^{(N)}$ of the form

$$c_\alpha^{(N)} = \sum_{|\beta| \leq 2N - |\alpha|} V_{\alpha\beta}^{(N)}(x) \partial^\beta a_0 \tag{129}$$

and $V_{\alpha\beta}^{(N)}$ which are C^∞ functions depending only on the functions V_α defining the operator P.

Proof. First remark that by (127), the operator A_N is a differential operator of order N. By induction, using (127) one easily sees that the coefficients of such an operator are linear combinations of the derivatives of a_0. To show (129) we proceed by induction. The result is true for $N=0$. Then use equation (127) to compute

$$\left[V_\alpha \partial^\alpha; c_\beta^{(N)} \partial^\beta\right] = \sum_{\gamma_j \le \alpha_j + \beta_j} \left[V_\alpha \binom{\beta}{\gamma} \partial^\gamma c_\beta^{(N)} - c_\beta^{(N)} \binom{\alpha}{\gamma} \partial^\gamma V_\alpha\right] \partial^{\alpha+\beta-\gamma} \tag{130}$$

Consider the first term in the square bracket which is the one involving more derivatives of $c_\beta^{(N)}$. Since $c_\beta^{(N)}$ depends on $\partial^\delta a_0$ with $|\delta| \le 2N - |\beta|$, one has that $\partial^\gamma c_\beta^{(N)}$ depends only on the derivatives $\partial^\delta a_0$ with $|\delta| \le 2N - |\beta| + |\gamma|$; in order to conclude the proof we have to show that this is smaller than $2(N+1) - (|\alpha| + |\beta| - |\gamma|)$, a fact which is true since $|\alpha| \le 2$. □

Remark 20. Let $u_n \in E_n$ then by (123) one has

$$\|u_n\|_s \le C n^s \|u_n\|_0$$

Remark 21. Let $u_n \in E_n$ with $\|u_n\|_0 = 1$, and b_α be a smooth function ($\alpha \in \mathbb{N}^d$), then one has for any $\nu_0 > d/2$ one has

$$\|b_\alpha \partial^\alpha u_n\|_0 \le C_{\nu_0} \|b_\alpha\|_{\nu_0} n^{|\alpha|} \tag{131}$$

Remark 22. Let $u_n \in E_n$ with $\|u_n\|_0 = 1$, and let

$$b_\alpha := V_{\alpha\beta}^{(N)}(x) \partial^\beta u_n \tag{132}$$

(with some β) then one has

$$\|b_\alpha\|_{\nu_0} \le C n^{\nu_0 + |\beta|} \tag{133}$$

with a C that depends on $V_{\alpha\beta}^{(N)}$.

End of the proof of proposition 2. Assume that $n_3 \le n_2 \le n_1$ so that $\mu(n) = n_3$ and $S(n) = n_3 + n_1 - n_2$. Write the l.h.s. of (85) as

$$|\langle A u_{n_2}; u_{n_1} \rangle| \tag{134}$$

with A the multiplication operator by u_{n_3}. Using (125) this is smaller than

$$\frac{1}{|n_1^2 - n_2^2|^N} \|A_N u_{n_2}\|_{L^2} \|u_{n_1}\|_{L^2} \ . \tag{135}$$

To estimate $\|A_N u_{n_2}\|_{L^2}$ we use (128) and estimate each term separately. By Sobolev embedding theorem, one term is estimated by

$$\left\|c_\alpha^{(N)}\partial^\alpha u_{n_2}\right\| \leq C\left\|c_\alpha^{(N)}\right\|_{v_0}\|\partial^\alpha u_{n_2}\|$$

$v_0 > d/2$. Using (129), (131), (133) one gets

$$\left\|c_\alpha^{(N)}\right\|_{v_0} \leq C\left\|u_{n_3}\right\|_{2N+v_0-|\alpha|} \leq Cn_3^{2N+v_0-|\alpha|}\left\|u_{n_3}\right\|_{L^2} \ .$$

where we used the ellipticity of P. We thus get that the l.h.s. of (85) is estimated by

$$C\sum_{|\alpha|\leq N} n_3^{2N+v_0-|\alpha|}n_2^{|\alpha|}\frac{1}{|n_1^2-n_2^2|^N}\left\|u_{n_1}\right\|_{L^2}\left\|u_{n_2}\right\|_{L^2}\left\|u_{n_3}\right\|_{L^2}$$

A part from a constant, the sum of the coefficients in front of the norms is estimated by

$$n_3^{2N+v_0}\left(\frac{n_2}{n_3}\right)^N\frac{1}{|n_1^2-n_2^2|^N} = \left(\frac{n_2}{n_1+n_2}\right)^N\frac{n_3^{v_0+N}}{|n_1-n_2|^N} \leq \frac{n_3^{v_0+N}}{|n_1-n_2|^N} \tag{136}$$

To conclude the proof just remark that $n_3 = \mu$, $S = \mu + (n_1 - n_2)$ and that if $n_3 > n_1 - n_2$ then the inequality (85) is trivially true. On the contrary, if $n_3 \leq (n_1 - n_2)$ the r.h.s. of (136) is smaller than

$$n_3^{v_0+N}\frac{2}{(n_3+|n_1-n_2|)^N}$$

which concludes the proof. □

6.4 *Proof of theorem 5 on the nonresonance condition*

The proof follows the proof of theorem 6.5 of [Bam03] (see also [BG06]). We repeat the main steps for completeness. Fix r once for all and denote by C any constant depending only on r. The value of C can change from line to line. Finally we will denote $m := \mu^2$.

Lemma 9. *For any $K \leq N$, consider K indexes $j_1 < ... < j_K \leq N$; consider the determinant*

$$D := \begin{vmatrix} \omega_{j_1} & \omega_{j_2} & \dots & \omega_{j_K} \\ \frac{d\omega_{j_1}}{dm} & \frac{d\omega_{j_2}}{dm} & \dots & \frac{d\omega_{j_K}}{dm} \\ . & . & \dots & . \\ . & . & \dots & . \\ \frac{d^{K-1}\omega_{j_1}}{dm^{K-1}} & \frac{d^{K-1}\omega_{j_2}}{dm^{K-1}} & \dots & \frac{d^{K-1}\omega_{j_K}}{dm^{K-1}} \end{vmatrix} \tag{137}$$

One has

$$D = C\left(\prod_l \omega_{i_l}^{-2K+1}\right)\left(\prod_{1\le l<k\le K}(\lambda_{j_l}-\lambda_{j_k})\right) \ge \frac{C}{N^{2K^2}}. \tag{138}$$

Proof. One has

$$\frac{d^j\omega_i}{dm^j} = \frac{(2j-1)!}{2^{j-1}(j-1)!2^j}\frac{(-1)^j}{(\lambda_i+m)^{j-\frac{1}{2}}}\ . \tag{139}$$

Substitute (139) in the r.h.s. of (9), factorize from the $l-th$ column the term $(\lambda_{j_l}+m)^{1/2}$, and from the $j-th$ row the term $\frac{(2j-3)!}{2^{j-2}(j-2)!2^j}$. The determinant becomes

$$C\left[\prod_{l=1}^{K}\omega_{j_l}\right]\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_{j_1} & x_{j_2} & x_{j_3} & \dots & x_{j_K} \\ x_{j_1}^2 & x_{j_2}^2 & x_{j_3}^2 & \dots & x_{j_K}^2 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ x_{j_1}^{K-1} & x_{j_2}^{K-1} & x_{j_3}^{K-1} & \dots & x_{j_K}^{K-1} \end{vmatrix} \tag{140}$$

where we denoted by $x_j := (\lambda_j+m)^{-1} \equiv \omega_j^{-2}$. The last determinant is a Vandermond determinant given by

$$\prod_{1\le l<k\le K}(x_{j_l}-x_{j_k}) = \prod_{1\le l<k\le K}\frac{\lambda_{j_k}-\lambda_{j_l}}{\omega_{j_l}^2\omega_{j_k}^2} = \left(\prod_{1\le l<k\le K}(\lambda_{j_l}-\lambda_{j_k})\right)\prod_{l=1}^{K}\omega_{j_l}^{-2K}\ . \tag{141}$$

Using the asymptotic of the frequencies one gets also the second of (138). □

Next we need the lemma from appendix B of [BGG85], namely

Lemma 10. *Let $u^{(1)},\dots,u^{(K)}$ be K independent vectors with $\left\|u^{(i)}\right\|_{\ell^1} \le 1$. Let $w\in\mathbb{R}^K$ be an arbitrary vector, then there exist $i\in[1,\dots,K]$, such that*

$$|u^{(i)}\cdot w| \ge \frac{\|w\|_{\ell^1}\det(u^{(1)},\dots,u^{(K)})}{K^{3/2}}\ .$$

Combining Lemmas 9 and 10 we deduce

Corollary 5. *Let $w\in\mathbb{R}^\infty$ be a vector with K components different from zero, namely those with index $i_1,\dots,i_K$; assume $K\le N$, and $i_1<\dots<i_K\le N$. Then, for any $m\in[m_0,\Delta]$ there exists an index $i\in[0,\dots,K-1]$ such that*

$$\left|w\cdot\frac{d^i\omega}{dm^i}(m)\right| \ge C\frac{\|w\|_{\ell^1}}{N^{2K^2+2}} \tag{142}$$

where ω is the frequency vector.

From [XYQ97] we have.

Lemma 11. *(Lemma 2.1 of [XYQ97]) Suppose that $g(\tau)$ is m times differentiable on an interval $J \subset \mathbb{R}$. Let $J_h := \{\tau \in J \ : \ |g(\tau)| < h\}$, $h > 0$. If on J, $\left|g^{(m)}(\tau)\right| \geq d > 0$, then $|J_h| \leq Mh^{1/m}$, where*

$$M := 2(2+3+...+m+d^{-1}) \ .$$

For any $k \in \mathbb{Z}^N$ with $|k| \leq r$ and for any $n \in \mathbb{Z}$, define

$$\mathscr{R}_{kn}(\gamma,\alpha) := \left\{ m \in [m_0,\Delta] \ : \ \left|\sum_{j=1}^{N} k_j\omega_j + n\right| < \frac{\gamma}{N^\alpha} \right\} \tag{143}$$

Applying lemma 11 to the function $\sum_{j=1}^{N} k_j\omega_j + n$ and using corollary 5 we get as in [Bam99] lemma 8.4

Corollary 6. *Assume $|k|+|n| \neq 0$, then*

$$|\mathscr{R}_{kn}(\gamma,\alpha)| \leq C(\Delta - m_0)\frac{\gamma^{1/r}}{N^\varsigma} \tag{144}$$

with $\varsigma = \frac{\alpha}{r} - 2r^2 - 2$.

Lemma 12. *Fix $\alpha > 2r^3 + r^2 + 5r$. For any positive γ small enough there exists a set $\mathscr{I}_\gamma \subset [m_0,\Delta]$ such that $\forall m \in \mathscr{I}_\gamma$ one has that for any $N \geq 1$*

$$\left|\sum_{j=1}^{N} k_j\omega_j + n\right| \geq \frac{\gamma}{N^\alpha} \tag{145}$$

for all $k \in \mathbb{Z}^N$ with $0 \neq |k| \leq r$ and for all $n \in \mathbb{Z}$. Moreover,

$$\left|[m_0,\Delta] - \mathscr{I}_\gamma\right| \leq C\gamma^{1/r} \ . \tag{146}$$

Proof. Define $\mathscr{I}_\gamma := \bigcup_{nk} \mathscr{R}_{nk}(\gamma,\alpha)$. Remark that, from the asymptotic of the frequencies, the argument of the modulus in (145) can be small only if $|n| \leq CrN$, By (144) one has

$$\left|\bigcup_k \mathscr{R}_{nk}(\gamma,\alpha)\right| \leq \sum_k |\mathscr{R}_k(\gamma,\alpha)| < C\frac{N^r(\Delta - m_0)\gamma^{1/r}}{N^\varsigma} \ ,$$

summing over n one gets an extra factor rN. Provided α is chosen according to the statement, one has that the union over N is also bounded and therefore the thesis holds. $\square$

Lemma 13. *For any γ positive and small enough, there exist a set $\mathscr{J}_\gamma$ satisfying, $|[m_0,\Delta] - \mathscr{J}_\gamma| \to 0$ when $\gamma \to 0$, and a real number α' such that for any $m \in \mathscr{J}_\gamma$ one has for $N \geq 1$*

$$\left| \sum_{j=0}^{N} \omega_j k_j + \varepsilon_1 \omega_j + \varepsilon_2 \omega_l \right| \geq \frac{\gamma}{N^{\alpha'}} \tag{147}$$

for any $k \in \mathbb{Z}^N$, $\varepsilon_i = 0, \pm 1$, $j \geq l > N$, *and* $|k| + |\varepsilon_1| + |\varepsilon_2| \neq 0$, $|k| \leq r+2$.

Proof. We consider only the case now the case $\varepsilon_1 \varepsilon_2 = -1$ which is the most complicate. One has

$$\omega_j - \omega_l = j - l + a_{jl} \text{ with } \left|a_{jl}\right| \leq \frac{C}{l} \tag{148}$$

So the quantity to be estimated reduces to

$$\sum_{j=0}^{N} \omega_j k_j \pm n \pm a_{jl} \ , \quad n := j - l$$

If $l > 2CN^{\alpha}/\gamma$ then the a_{jl} term represents an irrelevant correction and therefore the lemma follows from lemma 12. In the case $l \leq 2CN^{\alpha}/\gamma$ one reapplies the same lemma with $N' := 2CN^{\alpha}/\gamma$ in place of N and $r' := r+2$ in place of r. $\square$

To obtain theorem 5 just define $\mathscr{J} := \bigcap_{r \geq 1} \bigcup_{\gamma > 0} \mathscr{J}_\gamma$ and remark that its complement is the union of a numerable infinity of sets of zero measure.

References

[Bam99] D. Bambusi. On long time stability in Hamiltonian perturbations of non-resonant linear PDEs. *Nonlinearity*, 12:823–850, 1999.

[Bam03] D. Bambusi. Birkhoff normal form for some nonlinear PDEs. *Comm. Math. Physics*, 234:253–283, 2003.

[BDGS07] D. Bambusi, J.M. Delort, B. Grébert, and J. Szeftel. Almost global existence for Hamiltonian semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds. *Comm. Pure Appl. Math.*, 2007.

[BG06] D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.*, 135(3):507–567, 2006.

[BGG85] G. Benettin, L. Galgani, and A. Giorgilli. A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems. *Celestial Mech.*, 37:1–25, 1985.

[Bou96a] J. Bourgain. Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations. *Geometric and Functional Analysis*, 6:201–230, 1996.

[Bou96b] Jean Bourgain. On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. *Internat. Math. Res. Notices*, (6):277–304, 1996.

[Bou97] J. Bourgain. On growth in time of Sobolev norms of smooth solutions of nonlinear Schrödinger equations in $\mathbf{R}^D$. *J. Anal. Math.*, 72:299–310, 1997.

[Bou98] J. Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equation. *Ann. Math.*, 148:363–439, 1998.

[Bou00] J. Bourgain. On diffusion in high-dimensional Hamiltonian systems and PDE. *J. Anal. Math.*, 80:1–35, 2000.

[Bou05] J. Bourgain. *Green's function estimates for lattice Schrödinger operators and applications*, volume 158 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2005.

[CW93] W. Craig and C. E. Wayne. Newton's method and periodic solutions of nonlinear wave equations. *Comm. Pure Appl. Math.*, 46:1409–1498, 1993.

[DS04] J. M. Delort and J. Szeftel. Long–time existence for small data nonlinear Klein–Gordon equations on tori and spheres. *Internat. Math. Res. Notices*, 37:1897–1966, 2004.

[DS06] J.-M. Delort and J. Szeftel. Long-time existence for semi-linear Klein-Gordon equations with small Cauchy data on Zoll manifolds. *Amer. J. Math.*, 128(5):1187–1218, 2006.

[EK06] H. L. Eliasson and S. B. Kuksin. KAM for non-linear Schrödinger equation. *Preprint*, 2006.

[Gré06] B. Grébert. Birkhoff normal form and Hamiltonian PDES. *Preprint.*, 2006.

[Kla83] S. Klainerman. On "almost global" solutions to quasilinear wave equations in three space dimensions. *Comm. Pure Appl. Math.*, 36:325–344, 1983.

[KP96] S. B. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. Math.*, 143:149–179, 1996.

[Kuk87] S. B. Kuksin. Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum. *Funct. Anal. Appl.*, 21:192–205, 1987.

[Way90] C. E. Wayne. Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. *Comm. Math. Physics*, 127:479–528, 1990.

[XYQ97] J. Xu, J. You, and Q. Qiu. Invariant tori for nearly integrable Hamiltonian systems with degeneracy. *Math. Z.*, 226:375–387, 1997.

[Yua06] X. Yuan. A KAM theorem with applications to partial differential equations of higher dimensions. *Commun. Math. Phys* to appear (Preprint 2006).