BIRKHOFF NORMAL FORM FOR SOME NONLINEAR PDEs

Dario BAMBUSI

Dipartimento di Matematica,
Università degli studi di Milano
Via Saldini 50, 20133 MILANO, Italy.

Abstract. We consider the problem of extending to PDEs Birkhoff normal form theorem on Hamiltonian systems close to nonresonant elliptic equilibria. As a model problem we take the nonlinear wave equation

$$u_{tt} - u_{xx} + g(x, u) = 0 \ , \tag{0.1}$$

with Dirichlet boundary conditions on $[0, \pi]$; g is an analytic skewsymmetric function which vanishes for $u = 0$ and is periodic with period 2π in the x variable. We prove, under a nonresonance condition which is fulfilled for most g's, that for any integer M there exists a canonical transformation that puts the Hamiltonian in Birkhoff normal form up to a reminder of order M. The canonical transformation is well defined in a neighbourhood of the origin of a Sobolev type phase space of sufficiently high order. Some dynamical consequences are obtained. The technique of proof is applicable to quite general equations in one space dimension.

1. Introduction

Consider a finite dimensional Hamiltonian system having an elliptic equilibrium point; it is well known that there exist canonical coordinates in which the Hamiltonian takes the form

$$H = h_0 + f$$

where

$$h_0(p, q) = \sum_j \omega_j \frac{p_j^2 + q_j^2}{2} \ , \tag{1.1}$$

and $f(p,q) = \mathcal{O}(\|(p,q)\|^3)$. Moreover, if the frequencies are nonresonant, then for any integer $M \geq 4$ there exists a canonical transformation $\mathcal{T}$ (defined in a neighborhood of the origin) which puts the system in Birkhoff normal form up to order M, i.e. such that

$$H \circ \mathcal{T} = h_0 + Z + \mathcal{R} \tag{1.2}$$

where Z depends only on the action variables $I_j = (p_j^2 + q_j^2)/2$, and $\mathcal{R} = \mathcal{O}(\|(p,q)\|^M)$ is a remainder which is small close to the origin.

In the present paper we investigate the existence of such a transformation for some infinite dimensional Hamiltonian systems describing nonlinear partial differential equations (PDEs). As a model problem we consider the nonlinear wave equation

$$u_{tt} - u_{xx} + g(x,u) = 0 \ , \tag{0.1}$$

with Dirichlet boundary conditions on $[0,\pi]$, where g is a function which extends analytically to the region $|u| < R_1$ and $|\mathrm{Im}\, x| < \sigma_1$, where R_1, σ_1 are positive numbers, and fulfills

$$g(x,0) = 0 \ , \quad g(x+2\pi, u) = g(x,u) \ , \quad g(-x,-u) = -g(x,u) \ . \tag{1.3}$$

We will prove that if g fulfills a suitable nonresonance condition (which has full measure in a sense that will be made precise), then for any $M \geq 4$ and any positive (small) R there exists a canonical transformation which puts the Hamiltonian of the system in the form (1.2). The transformation is well defined in an open ball of radius R of Sobolev–type phase spaces of sufficiently high order, and the norm of the hamiltonian vector field of $\mathcal{R}$ is smaller than CR^M in the same ball, with C a positive constant independent of R.

As a corollary we obtain an estimate of high Sobolev norms of the solution which is valid for very long times (in the spirit of [8]) and also a very precise qualitative description of the dynamics for such times. In particular we obtain (in the spirit of [15,16]) that all the actions I_j of the linearized system are approximately constants of motion for times of the order of the norm of the initial datum to the power $-M$.

We recall that perturbation theory for Hamiltonian PDEs has attracted great interest during the last fifteen years, and in particular KAM theory has been partially extended by proving that most of the *finite dimensional* invariant tori of an integrable (e.g. linear) system survive a small perturbation [12,13,11,7,9,18,19,10]. However, it is clear that, although very interesting, the solutions lying on these tori correspond to exceptional initial data. Concerning solutions starting outside such tori only a few results are known [1,2,6,3,8]. In particular, the papers [1,2] deal with perturbations of resonant systems and prove a long time stability property of periodic solutions or finite dimensional tori in the topology induced by the energy norm; such results are based on resonant averaging combined with other techninques. The papers [6,3] deal with perturbations of nonresonant systems and their results allow to control, for long times, the dynamics corresponding to all initial data in a ball of a phase space of smooth functions; however the topology in which the dynamics is controlled is much weaker than that of the original phase space (see remark 2.4). Such latter results are obtained by exploiting the linear stability of finite dimensional tori. Finally we quote the remarkable paper [8] where a stability property of most small amplitude infinite dimensional tori of some nonlinear Schrödinger equations is obtained. This result is strongly related to the main one proved here.

In the present paper we obtain a precise qualitative description of all the solutions starting in a ball of a Sobolev type phase space, in particular we obtain a control of the (phase space) norm of the solution, and moreover we prove that each solution remains close, in a topology slightly weaker than that of the phase spase, to an infinite dimensional torus (see corollary 2.3). We point out that the technique of proof of the present paper is very general, while that of [8] (which is completely different from ours) exploits some particular features of the NLS. Moreover, we think that the approach of the present paper could be the starting point for further developments in the same way as Birkhoff normal form is for finite dimensional systems (following for example [13,14,17]).

The paper is organized as follows: in sect. 2 we give a precise statement of the results for the equation (0.1), in sect. 3 we present the idea of the proof and we discuss some possible extensions. In sect. 4 we state our abstract normal form theorem. In sect. 5 we give the proof of such an abstract result. This section is divided in three subsection: in the first one we study the properties of the class of functions allowed as nonlinearities by our theory; in the second one we study the behaviour of such functions under canonical transformations generated by functions of the same class; in the third subsection we iteratively construct the normalizing transformation. Sect. 6 is devoted to the application to the nonlinear wave equation (0.1) and is divided in two subsections: in the first one we give equation (0.1) the form needed for the application of the abstract theorem; in the second subsection we prove that for most values of a suitable parameter the frequencies fulfill the nonresonance condition needed to apply the theorem.

2. Statement of the results

To start with we precise our assumptions. Write g as

$$g(x,u)=g_0(x)u+\bar g(x,u)\ ,\quad g_0(x)=m+V(x)\ ,\quad \bar g(x,u)=\mathcal{O}(|u|^2)$$

where m is the average over $[-\pi,\pi]$ of g_0 and V has zero average. Consider now the Sturm Liouville problem

$$-\partial_{xx}\varphi_j+V\varphi_j=\mu_j\varphi_j\ ,\quad \varphi_j(0)=\varphi_j(\pi)=0\ ; \tag{2.1}$$

it is well known that, since V is analytic and symmetric, the solutions $\{\varphi_j\}_{j\geq 1}$ of (2.1) form an orthogonal basis of L^2 that we will assume to be normalized, and that they are analytic and skewsymmetric. *We assume that the eigenvalues μ_j are distinct* (the case of multiple eigenvalues can also be dealt with very easily, but the statement is more complicate). Remark that the eigenvalues μ_j define the frequencies of oscillation of the linear system by

$$\omega_j:=\sqrt{\mu_j+m}\ . \tag{2.2}$$

It is well known that (0.1) is Hamiltonian with Hamiltonian function given by

$$H = h_0 + f \ , \tag{2.3}$$

where

$$h_0(v,u) := \int_0^\pi \frac{[v(x)]^2 + [u_x(x)]^2 + V(x)[u(x)]^2 + m[u(x)]^2}{2} dx \ ,$$
$$f(u) := \int_0^\pi G(x, u(x)) dx \ ;$$

$v := u_t$ is the momentum conjugated to u and $G(x,u)$ is such that $\partial_u G = \bar{g}$. To define formally the phase space consider the Sobolev space H^s of the functions $u \in L^2$ having s weak derivatives in L^2. For any integer $s \geq 2$ we define the phase space $\mathcal{F}_s$ as the space of the functions $(u,v) \in \mathcal{F}_s := H^s \oplus H^{s-1}$ fulfilling the compatibility conditions

$$\begin{aligned} u^{(2j)}(0) &= u^{(2j)}(\pi) = 0 \ , \quad 0 \leq j \leq \frac{s-1}{2} \ , \\ v^{(2j)}(0) &= v^{(2j)}(\pi) = 0 \ , \quad 0 \leq j \leq \frac{s}{2} - 1 \ , \end{aligned} \tag{2.4}$$

(the upper index in parentheses denoting derivative of the corresponding order). We endow $\mathcal{F}_s$ by the norm

$$\|(u,v)\|^2_{\mathcal{F}_s} := \int_0^\pi \left[(\partial_x^{s-1} v)^2 + (\partial_x^s u)^2 + u^2 + v^2 \right] dx \ ;$$

we will denote by $B_s(R)$ the ball of radius R in $\mathcal{F}_s$ centered at the origin.

Remark 2.1. *It is well known that the Cauchy problem for equation (0.1) is locally well posed in $\mathcal{F}_s$, and that it is also globally well posed for small enough initial data.*

To introduce the normal modes and the linear actions consider the expansion of u and v on the eigenfunctions φ_j:

$$u(x) = \sum_{j \geq 1} u_j \varphi_j(x) \ , \quad v(x) = \sum_{j \geq 1} v_j \varphi_j(x) \ ,$$

and define the j–th action of the linearized system by

$$I_j \equiv I_j(u,v) := \frac{v_j^2 + \omega_j^2 u_j^2}{2\omega_j}$$

Having fixed two positive numbers $m_1 < m_2$, such that $\mu_j + m_1 > 0$ for all $j \geq 1$ we will consider the parameter m as varying in the interval $\mathcal{I} := [m_1, m_2]$.

We have the following

Theorem 2.2. *Fix a positive integer $M \geq 4$, then there exists a subset $\mathcal{J} \subset \mathcal{I}$ with measure equal to $m_2 - m_1$ and a positive (large) integer s_* with the following properties: for any $m \in \mathcal{J}$ there exists a positive R_* and, for any $R < R_*$ there exists an analytic* *canonical transformation* $\mathcal{T}_R : B_{s_*}(R/3) \to B_{s_*}(R)$ *which puts the Hamiltonian (2.3) in Birkhoff normal form up to order* R^M*, namely such that*

$$(h_0 + f) \circ \mathcal{T}_R = h_0 + Z + \mathcal{R}$$

with Z *a function that depends only on the actions* I_j*, and* $\mathcal{R}$ *an analytic functional having a Hamiltonian vector field* $X_{\mathcal{R}}$ *which is analytic as a function from* $B_{s_*}(R/3)$ *to* $\mathcal{F}_{s_*}$*. Moreover, for any integer* $s \geq s_*$ *there exists a positive constant* C_s *such that the transformation* $\mathcal{T}_R$ *and the vector field* $X_{\mathcal{R}}$ *are analytic also as maps from* $B_s(R/C_s)$ *to* $\mathcal{F}_s$ *and the following estimates hold*

$$\begin{aligned} \sup_{\|(u,v)\|_{\mathcal{F}_s} \leq R/C_s} \|(u,v) - \mathcal{T}_R(u,v)\|_{\mathcal{F}_s} \leq C_s R^2 \ , \\ \sup_{\|(u,v)\|_{\mathcal{F}_s} \leq R/C_s} \|X_{\mathcal{R}}\|_{\mathcal{F}_s} \leq C_s R^{M-1} \ . \end{aligned} \tag{2.5}$$

Consider now an initial datum $(\bar{u}, \bar{v}) \in \mathcal{F}_s$ for the nonlinear wave equation (0.1) and define the infinite dimensional torus on which all the actions coincide with the actions of $(\bar{u}, \bar{v})$, namely

$$\mathbf{T}_0 := \{(u,v) \in \mathcal{F}_s \ : \ I_j(u,v) = I_j(\bar{u}, \bar{v}) \ , \ \forall j \geq 1\}$$

Then we have the following

Corollary 2.3. *Fix* $M \geq 4$ *and* $m \in \mathcal{J}$*, fix also* $s \geq s_*$*, then there exist positive constants* R_s*,* C_s *such that the following holds true: let* $\zeta(t) \equiv (u(t), v(t))$ *be the solution of the Cauchy problem for equation (0.1) with initial datum* $(\bar{u}, \bar{v}) \in \mathcal{F}_s$ *if*

$$\varepsilon := \|(\bar{u}, \bar{v})\|_{\mathcal{F}_s} \leq R_s$$

then for all times t *with*

$$|t| \leq \frac{1}{C_s \varepsilon^{M-3}} \tag{2.6}$$

the following estimates hold

$$\|\zeta(t)\|_{\mathcal{F}_s} \leq 2C_s\varepsilon \ , \quad \frac{|I_j(t) - I_j(0)|}{\varepsilon^2} \leq \frac{C_s}{j^{2s}}\varepsilon \tag{2.7}$$

moreover, for all $s' < s - 1/2$*,* $s' \geq 2$*, there exists a constant* $C_{s,s'}$ *such that, up to the times (2.6) one has*

$$d_{s'}(\zeta(t), \mathbf{T}_0) \leq C_{s,s'} \varepsilon^{5/4} \tag{2.8}$$

where $d_{s'}(.,.)$ *denotes the distance in* $\mathcal{F}_{s'}$*.*

We point out that the first of the estimates (2.7) holds also for the nonlinear Klein–Gordon equation (corresponding to the case $V = 0$) with periodic boundary conditions (see corollary 6.7 below), and also the second of (2.7) has an anologous in this case.

Remark 2.4. *Corollary 2.3 has to be confronted with the main result of [6]. In [6] the estimates (2.7) where proven for solutions corresponding to initial data in* $\mathcal{F}_{s+s_1}$*, with* ε *defined as the* $\mathcal{F}_{s+s_1}$ *norm of the initial datum, and* s_1 *a large* M *dependent number.*

3. Idea of the proof and discussion

We begin by discussing the renonresonance condition. To this end consider a finite sequence of frequencies $\omega \equiv \{\omega_j\}_{j=1}^n$, and recall that the standard nonresonance condition used in order to prove Birkhoff's theorem in a finite dimensional context is "for any integer $r \geq 3$ there exists a positive δ_r such that $|\omega \cdot \mathbf{k}| \geq \delta_r$, $\forall \mathbf{k} \in \mathbf{Z}^n \backslash \{0\}$ with $|\mathbf{k}| \leq r$". It can be easily seen that in general such a condition is not satisfied in the infinite dimensional case $n = \infty$, so it has to be modified in order to deal with PDEs.

To this end remark that, given $\mathbf{k}$ as above, one has

$$\mathbf{k} \cdot \omega \equiv \pm\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_r}$$

with a suitable choice of the signs and of the indices $i_1, ..., i_r$, and consider the nonresonance condition

$$|\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_r}| \geq \frac{\gamma}{i_1^\tau i_2^\tau ... i_r^\tau} \quad : \tag{3.1}$$

we will prove that, provided τ is sufficiently large (depending on r), the frequencies (2.2) fulfill (3.1) if m is in a set with complement having measure of the order of $\gamma^{1/r}$ (see theorem 6.16). We point out that condition (3.1) is fulfilled with large probability in very general models. For example, this is true if $\omega \in [0,1]^\infty$ with $[0,1]^\infty$ endowed by the product measure.

Condition (3.1) makes clear that very small small–denominators always involve many times large indexes (or many times a large index).

Recall now that small denominators appear in solving the homological equation. More precisely, having introduced variables $\{(\xi_j, \eta_j)\}_{j\geq 1}$ in which the operator of Poisson brackets with h_0 is diagonal, the algorithm of Birkhoff normal form introduces a denominator of the form

$$\omega_{k_1} + \omega_{k_2} + ... + \omega_{k_{n_1}} - \omega_{j_1} - \omega_{j_2} - ... - \omega_{j_{n_2}}$$

in order to eliminate from the nonlinearity a monomial of the form

$$f_{k,j}\ \xi_{k_1}\xi_{k_2}...\xi_{k_{n_1}}\eta_{j_1}\eta_{j_2}...\eta_{j_{n_2}} \tag{3.2}$$

where $f_{k,j} \in \mathbf{C}$ is a complex coefficient. The main point is that, due to the locality of the nonlinearity in (0.1), the vector field of monomial having a nonvanishing coefficient $f_{k,j}$ turns out to be small when at least three of the indexes k_l, j_l are large and the sequences ξ_j, η_j decay fast enough with j (see lemma 5.6 and proposition 5.8 below).

The smallness of monomials involving many variables with large index allows to perform a spatial analogue of the ultraviolet cutoff currently used in KAM theory, namely a momentum cutoff: Precisely we fix a large N and eliminate from the nonlinearity only monomials which contain at most two variables with index larger than N, the others being already small. In order to do that we assume that the small denominators one has to control satisfy a diophantine type condition stronger than (3.1), namely that there exists α such that

$$|\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_r} \pm \omega_j \pm \omega_k| \geq \frac{\gamma}{N^\alpha} \tag{3.3}$$

where the indexes i_l are smaller than N and $k \geq j > N$. This is the condition we will prove to have full measure (see theorem 6.5 below). We point out that a related diophantine type condition was already used in [6] by Bourgain, and that a momentum cutoff where already used in different contexts in [5,11].

We also point out that the properties that allow to prove that monomials involving variables with large index have a small vector field are (i) the regularity of the nonlinearity, (ii) its locality and (iii) the fact that the eigenfunctions of the linearized problem are "well localized with respect to the Fourier basis". Technically we are able to obtain the proof by a trick which consists in embedding the dynamical system (0.1) into a larger one, essentially in the same way as one embeds periodic skewsymmetric functions in the spaces of periodic functions in order to use the Fourier expansions on the exponentials which is simpler to manipulate than the sin expansion.

We conclude by remarking that our approach applies directly also to the case of the nonlinear Klein–Gordon equation with periodic boundary conditions (see corollary 6.7 below), and also to quite general equations in one space dimension. For example it can be used in the case of the nonlinear Schrödinger equation with or without potential, and in particular we expect that it allows to re–obtain the results of [8]. Concerning possible extensions to systems with more than one space dimension we remark that the nonresonance condition (3.1) still holds, and also the smallness property of terms with large index holds for these systems, however equation (3.3) is typical of systems in one space dimension and we have no ideas on the possibility of obtaining some results under weaker nonresonance assumptions.

4. Statement of the abstract result

Denote $\bar{\mathbf{Z}} := \mathbf{Z}\backslash\{0\}$; define the space ℓ^2_s of the complex sequences $\xi \equiv \{\xi_j\}_{j\in\bar{\mathbf{Z}}}$ such that

$$\|\xi\|^2_{\ell^2_s} := \sum_{j\in\bar{\mathbf{Z}}} (1+j^{2s})\,|\xi_j|^2 < \infty \ ,$$

and the symplectic spaces

$$\mathcal{P}_s := \ell^2_s \oplus \ell^2_s \ni \zeta \equiv \left(\{\xi_j\}_{j\in\bar{\mathbf{Z}}}, \{\eta_j\}_{j\in\bar{\mathbf{Z}}}\right) \ ,$$

endowed by the symplectic form $i\sum_{j\in\bar{\mathbf{Z}}} d\eta_j \wedge d\xi_{-j}$. In $\mathcal{P}_s$ we will use the norm $\|\zeta\|^2_s = \|\xi\|^2_{\ell^2_s} + \|\eta\|^2_{\ell^2_s}$, and will denote by $\mathcal{B}_s(R^{(s)})$ the ball of radius $R^{(s)}$ centered at the origin of $\mathcal{P}_s$. For the Hamiltonian vector field of a Hamiltonian function H we will use the notation X_H, so that the Hamilton equations are $\dot{\zeta} = X_H(\zeta)$. Explicitly they are given by

$$\frac{d}{dt}\xi_j = i\frac{\partial H}{\partial \eta_{-j}} \ , \quad \frac{d}{dt}\eta_j = -i\frac{\partial H}{\partial \xi_{-j}} \ .$$

In order to fix ideas one can think of ξ_j as the j-th Fourier coefficient on the basis of the exponentials of a smooth function, and similarly for η_j.

Consider a Hamiltonian of the form

$$H := h_0 + f \tag{4.1}$$

with

$$h_0(\xi,\eta) := \sum_{j\in\bar{\mathbf{Z}}} \omega_j \xi_j \eta_{-j} \ , \tag{4.2}$$

and real ω_j's; the functional f will be assumed to be smooth in a sense that will be specified in a while and to have a zero of order three at the origin.

Definition. *Consider a polynomial $h(\xi,\eta)$ homogeneous of degree n_1 in ξ and of degree n_2 in η, nalmely*

$$h(\xi,\eta) = \sum_{k,j} h_{k,j} \xi_{k_1}...\xi_{k_{n_1}} \eta_{j_1}...\eta_{j_{n_2}} \tag{4.3}$$

where we denoted $k=(k_1,...,k_{n_1})$, $j=(j_1,...,j_{n_2})$. The quantity

$$|h|_\mu := \sum_{l\in\mathbf{Z}} \mathrm{e}^{\mu|l|} \sup_{k_1+...+k_{n_1}+j_1+...+j_{n_2}=l} |h_{k,j}| \tag{4.4}$$

will be called μ–modulus of h.

Definition. *Consider a functional $h=\sum_{n_1,n_2} h^{n_1,n_2}$ with h^{n_1,n_2} homogenous of degrees n_1, n_2 in ξ and η respectively. If the quantity*

$$\langle|h|\rangle_R^\mu := \sum_{n_1,n_2} |h^{n_1,n_2}|_\mu R^{n_1+n_2} \tag{4.5}$$

is finite for some positive R, μ, then we will say that h is of class $\mathcal{M}_R^\mu$ and we will write $h\in\mathcal{M}_R^\mu$.

We will assume that $f\in\mathcal{M}_R^\mu$, and we will prove that in the case of the nonlinear wave equation the perturbation is of class $\mathcal{M}_R^\mu$.

Definition. *We will say that a sequence of frequencies $\{\omega_j\}_{j\in\bar{\mathbf{Z}}}$ is strongly non vanishing if for any $r\geq 3$ there exists a positive γ and a real α such that, for any (large) N, any choice of indexes $i_1,...,i_r,j,k$ with $|i_1|\leq|i_2|\leq...\leq|i_r|\leq N<|j|\leq|k|$ and any choice of $r+1$ signs one has that*

$$\omega_{i_1}\pm\omega_{i_2}\pm...\pm\omega_{i_r}\pm\omega_j\pm\omega_k\neq 0$$

implies

$$|\omega_{i_1}\pm\omega_{i_2}\pm...\pm\omega_{i_r}\pm\omega_j\pm\omega_k| \geq \frac{\gamma}{N^\alpha} \ . \tag{3.3}$$

We will assume that the frequencies are strongly nonvanishing.

The last definition we need is that of Birkhoff normal form:

Definition. *Consider a functional $Z=\sum_{n_1,n_2} Z^{n_1,n_2}$, with Z^{n_1,n_2} homogeneous of degree n_1 in ξ and of degree n_2 in η. Write*

$$Z^{n_1,n_2}=\sum_{k_1,k_2,...,k_{n_1},j_1,j_2,...,j_{n_2}} Z^{n_1,n_2}_{k_1,k_2,...,k_{n_1},j_1,...,j_{n_2}}\xi_{k_1}\xi_{k_2}...\xi_{k_{n_1}}\eta_{j_1}\eta_{j_2}...\eta_{j_{n_2}}\ ;$$

Z will be said to be in Birkhoff normal form if

$$Z^{n_1,n_2}_{k_1,k_2,...,k_{n_1},j_1,...,j_{n_2}}\neq 0 \Longrightarrow \omega_{k_1}+\omega_{k_2}+...+\omega_{k_{n_1}}-\omega_{j_1}-\omega_{j_2}-...-\omega_{j_{n_2}}=0\ .$$

In the application to equation (0.1) we will need to restrict the system to the submanifold of skew–symmetric functions that we will denote by

$$\mathcal{A}:=\left\{(\xi_k,\eta_k)\ :\ \xi_k=\xi_{-k}\ \text{and}\ \eta_k=\eta_{-k}\right\}\ ,$$

and to the manifold

$$\mathcal{R}e:=\left\{(\xi_k,\eta_k)\ :\ \xi^*_k=\eta_{-k}\right\}$$

corresponding to real functions.

We assume that

(1) the frequencies ω_j of the hamiltonian system (4.1) are strongly nonvanishing with some indices α,γ.

(2) there exists $\bar{R}_1>0$ and $\mu>0$ such that the function f is of class $\mathcal{M}^\mu_R$ for any $R\leq\bar{R}_1$ and fulfills the estimate

$$\langle|f|\rangle^\mu_R\leq AR^3$$

with some positive A.

Theorem 4.1. *Under the above assumptions, fix a large M, define $s_*:=4M\alpha(2M)$; then there exist constants C, R_* such that for any positive $R^{(s_*)}$ fulfilling $R^{(s_*)}<R_*$ the following holds true: There exists a canonical transformation $\mathcal{T}:\mathcal{B}_{s_*}(R^{(s_*)}/3)\to\mathcal{B}_{s_*}(R^{(s_*)})$ such that the transformed Hamiltonian has the form*

$$(h_0+f)\circ\mathcal{T}=h_0+Z+\mathcal{R}$$

with Z in Birkhoff normal form, and $\mathcal{R}$ an analytic functional having a Hamiltonian vector field which is analytic as a function from $\mathcal{P}_{s_*}$ to itself; moreover there exists a positive c such that for any $s\geq s_*$, and any $R^{(s)}\leq R_*/c^s$ the transformation $\mathcal{T}$ and the vector field $X_\mathcal{R}$ are analytic also as maps from $\mathcal{B}_s(R^{(s)})$ to $\mathcal{P}_s$ and the following estimate holds

$$\sup_{\|\zeta\|_s\leq R^{(s)}}\|\zeta-\mathcal{T}(\zeta)\|_s\leq C(c^sR^{(s)})^2$$

moreover

$$\sup_{\|\zeta\|_s\leq R^{(s)}}\|X_\mathcal{R}\|_s\leq C\left(c^sR^{(s)}\right)^M\ . \tag{4.6}$$

Finally, if the vector field of f leaves invariant $\mathcal{A}$ and/or $\mathcal{R}e$, then the same is true for $\mathcal{T}$ and for the vector fields of Z and of $\mathcal{R}$.

5. Proof of theorem 4.1.

We split the proof in three parts. In the first we establish some properties of the functions of class $\mathcal{M}^{\mu}_{R}$, in the second we prove that the Lie transform of a function of class $\mathcal{M}^{\mu}_{R}$ is of class $\mathcal{M}^{\mu}_{R'}$ with any $R'<R$, provided it is generated by a function of class $\mathcal{M}^{\mu}_{R}$; in the third subsection we prove the iterative lemma which constitute the main step of the proof, and deduce theorem 4.1.

We will denote

$$\rho_s := \frac{2^{s+1}}{\pi} \ ,$$

$$\Sigma_s := \sqrt{\sum_{j\in\mathbf{Z}}(1+j^{2s})\mathrm{e}^{-2\mu|j|}} \ ,$$

and remark that there exists a positive (μ dependent) constant c such that

$$\Sigma_s < c^s \qquad (5.1) \ .$$

Moreover C will denote a positive (large) constant whose value can change through the paper, and which is independent of the relevant quantities. For analytic functions f, that have an analytic vector field we will write

$$\left\|X_f\right\|_s^{R^{(s)}} := \sup_{\left\|\zeta\right\|_s\leq R^{(s)}} \left\|X_f(\zeta)\right\|_s \ .$$

Having fixed some integer N, we will denote by $(q_j,p_j)_{|j|\leq N}=(\xi_j,\eta_j)_{|j|\leq N}$, the first N canonical variables and by $(Q_j,P_j)_{|j|>N}=(\xi_j,\eta_j)_{|j|>N}$ the remaining ones.

5.1 Properties of the functions of class $\mathcal{M}^{\mu}_{R}$

The main results of this sub–section are proposition 5.3 and 5.8; the hurried reader can go directly to them and leave the rest of the subsection for a subsequent reading.

Given a homogeneous polynomial h of degree n_1 in ξ and of degree n_2 in η, write

$$h(\xi,\eta)=\sum_{k,i} h_{k,i}\xi_{k_1}...\xi_{k_{n_1}}\eta_{i_1}...\eta_{i_{n_2}} \ ; \qquad (5.2)$$

denote by

$$\$i\$:= i_1+i_2+...+i_{n_1}$$

the sum of the components of a multiindex i and define the functional

$$\check{h}(\xi,\eta) := \sum_{k,i} \mathrm{e}^{-\mu|\,\$k\$+\$i\$\,|}\xi_{k_1}...\xi_{k_{n_1}}\eta_{i_1}...\eta_{i_{n_2}} \ , \qquad (5.3)$$

which clearly depends only the values of n_1, n_2, in (5.2). The interest of this functional is that it is very easy to estimate the norm of its vector field, since $\check{h}$ can be written in the form

$$\check{h}(\xi,\eta) = \frac{1}{(2\pi)^{\frac{n-1}{2}}}\int_{-\pi}^{\pi} F(x)\xi(x)^{n_1}\eta(x)^{n_2}dx \tag{5.4}$$

where $n = n_1 + n_2$ and we defined

$$\xi(x) := \frac{1}{\sqrt{2\pi}}\sum_{j\in\bar{\mathbf{Z}}}\xi_j e^{ijx}\ ,\quad \eta(x) := \frac{1}{\sqrt{2\pi}}\sum_{j\in\bar{\mathbf{Z}}}\eta_j e^{ijx}\ ,\quad F(x) := \frac{1}{\sqrt{2\pi}}\sum_{j\in\bar{\mathbf{Z}}} e^{-\mu|j|}e^{ijx}\ . \tag{5.5}$$

Lemma 5.1. *One has*

$$\left\|X_h\right\|_s^R \leq |h|_\mu \left\|X_{\check{h}}\right\|_s^R$$

Proof. We write expliciptely the vector field of h, to this end we use the form (5.2). Consider the η_t component of X_h, it is given (eccept for an irrelevant sign) by

$$i\frac{\partial h}{\partial \xi_{-t}} = n_1 \sum_{k_2,\ldots,k_{n_1},i} h_{-t,k_2,\ldots,k_{n_1},i}\xi_{k_2}\ldots\xi_{k_{n_1}}\eta_{i_1}\ldots\eta_{i_{n_2}}$$

and the norm of the η component of X_h is given by

$$\sum_t \left(1+|t|^{2s}\right)\left| n_1 \sum_{k_2,\ldots,k_{n_1},i} h_{-t,k_2,\ldots,k_{n_1},i}\xi_{k_2}\ldots\xi_{k_{n_1}}\eta_{i_1}\ldots\eta_{i_{n_2}}\right|^2$$

$$\leq \left[\sup_{t,k,i}\left|h_{t,k_2,\ldots,k_{n_1},i}\right| e^{\mu|k_2+\ldots+k_{n_1}+\$i\$+t|}\right]^2 \left[\sum_t \left(1+|t|^{2s}\right)\times\right.$$

$$\left.\left(n_1\sum_{k_2,\ldots,k_{n_1},i} e^{-\mu|k_2+\ldots+k_{n_1}+\$i\$-t|}\left|\xi_{k_2}\ldots\xi_{k_{n_1}}\eta_{i_1}\ldots\eta_{i_{n_2}}\right|\right)^2\right]$$

but the first square bracket is smaller than $|h|_\mu$, while the second one is controlled by the norm of the η component of $X_{\check{h}}$. Indeed, the norm of the η component of $X_{\check{h}}$ is given by

$$\sum_t \left(1+|t|^{2s}\right)\left| n_1 \sum_{k_2,\ldots,k_{n_1},i} e^{-\mu|k_2+\ldots+k_{n_1}+\$i\$-t|}\xi_{k_2}\ldots\xi_{k_{n_1}}\eta_{i_1}\ldots\eta_{i_{n_2}}\right|^2$$

taking real positive ξ,η one obtains the expression in the above square bracket. Similar inequalities clearly hold for the other components of the vector field of h. One thus obtains

$$\left(|h|_\mu \left\|X_{\check{h}}\right\|_s^R\right)^2 = |h|_\mu^2 \sup_{R_1+R_2=R}\left(\left(\left\|X_{\check{h}}^\xi\right\|_s^{R_1}\right)^2 + \left(\left\|X_{\check{h}}^\eta\right\|_s^{R_2}\right)^2\right)$$

$$\geq \left(\sup_{R_1+R_2=R}\left(\left\|X_h^\xi\right\|_s^{R_1}\right)^2 + \left(\left\|X_h^\eta\right\|_s^{R_2}\right)^2\right) = \left(\left\|X_h\right\|_s^R\right)^2$$

which is the thesis. □

Lemma 5.2. *Consider the function $\check{h}$ as above, one has*

$$\|X_{\check{h}}\|_s^R \le n\Sigma_s(\rho_s R)^{n-1}$$

where $n = n_1 + n_2$.

Proof. We begin by considering the ξ component of $X_{\check{h}}$, one has

$$X^{\xi}_{\check{h}} = i\frac{n_2 F\xi^{n_1}\eta^{n_2-1}}{(2\pi)^{\frac{n-1}{2}}} \ ,$$

or more precisely

$$X^{\xi}_{\check{h}}(x) = i\frac{n_2 F(x)[\xi(x)]^{n_1}[\eta(x)]^{n_2-1}}{(2\pi)^{\frac{n-1}{2}}} \ ,$$

from which, using the inequality

$$\left\|\xi^2\right\|_{H^s} \le \sqrt{\frac{2}{\pi}}2^{s+1}\left\|\xi\right\|_{H^s}\left\|\xi\right\|_{H^s} \ , \tag{5.6}$$

and, remarking that $\Sigma_s = \|F\|_{H^s}$, one has $\left\|X^{\xi}_{\check{h}}\right\|_s^R \le n_2\Sigma_s(\rho_s R)^{n-1}$. Considering also the η component one has the thesis. □

Proposition 5.3. *Let $h \in \mathcal{M}^{\mu}_R$; then, for any positive $R^{(s)}$ such that $\rho_s R^{(s)} \le R$ and any $\delta^{(s)} < R^{(s)}$, the Hamiltonian vector field X_h is analytic as a map from $\mathcal{B}_s(R^{(s)})$ to $\mathcal{P}^s$ and fulfills the estimate*

$$\|X_h\|_s^{R^{(s)}-\delta^{(s)}} \le \frac{\Sigma_s}{\rho_s\delta^{(s)}}\langle|h|\rangle^{\mu}_{\rho_s R^{(s)}} \tag{5.7}$$

Proof. First work on homogenous polynomials of degree n. One has $h = \sum_{\substack{n_1,n_2\\ n_1+n_2=n}} h^{n_1,n_2}$, with h^{n_1,n_2} homogeneous in ξ, and in η. One has

$$\begin{aligned}\|X_h\|_s^{R^{(s)}-\delta^{(s)}} &\le \sum_{n_1,n_2}\|X_{h^{n_1,n_2}}\|_s^{R^{(s)}-\delta^{(s)}} \le \sum_{n_1,n_2}|h^{n_1,n_2}|_{\mu}\, n\Sigma_s\left(\rho_s(R^{(s)}-\delta^{(s)})\right)^{n-1}\\ &\le \frac{\Sigma_s}{\rho_s\delta^{(s)}}\sum_{n_1,n_2}|h^{n_1,n_2}|_{\mu}\left[\rho_s R^{(s)}\right]^n = \frac{\Sigma_s}{\rho_s\delta^{(s)}}\langle|h|\rangle^{\mu}_{\rho_s R^{(s)}}\end{aligned}$$

where we used the inequality $n(R - d^{n-1}) \le R^n d^{-1}$. Then the result easily extends to analytic functions. □

Consider now a homogeneous polynomial of degree m_1, m_2, m_3, m_4, in q, p, Q, P respectively, and write it in the form

$$h(q,p,Q,P) = \sum_{k,i,j,l} h_{k,i,j,l}q_{k_1}...q_{k_{m_1}}p_{i_1}...p_{i_{m_2}}Q_{j_1}...Q_{j_{m_3}}P_{l_1}...P_{l_{m_4}} \tag{5.8}$$

Correspondingly we define the the functional

$$\check{\check{h}}(p,q,P,Q) := \frac{1}{(2\pi)^{\frac{n-1}{2}}} \int_{-\pi}^{\pi} F(x) q(x)^{m_1} p(x)^{m_2} Q(x)^{m_3} P(x)^{m_4} dx$$

with $q(x)$, $p(x)$, $Q(x)$, $P(x)$ defined in analogy with (5.5). For its vector field we will prove an estimate stronger than (5.7); this is the heart of the momentum cutoff. Working exactly as in the proof of lemma 5.1 one has

Lemma 5.4. *Let h be homogeneous polynomial of degree m_1, m_2, m_3, m_4 in q, p, Q, P respectively, then one has*

$$\|X_h\|_s^R \leq |h|_\mu \left\| X_{\check{\check{h}}} \right\|_s^R$$

To come to the estimate of the vector field of $\check{\check{h}}$ we need a few lemmas concerning the vector field of the product of the functions $P(x)$ and $Q(x)$.

Lemma 5.5. *Let ϕ be such that*

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \sum_{|k|>N} \phi_k e^{ikx} \ ,$$

then, given positive integers s, j such that $s > j + \frac{1}{2}$, one has

$$\left| \frac{d^j \phi}{dx^j}(x) \right| \leq \sqrt{\frac{2}{\pi}} \frac{1}{N^{s-j-\frac{1}{2}}} \|\phi\|_{H^s} \ .$$

Proof. One has

$$\begin{aligned} \sqrt{2\pi} \left| \frac{d^j \phi}{dx^j}(x) \right| &= \left| \sum_{|k|>N} (ik)^j \phi_k e^{ikx} \right| \leq \sum_{|k|>N} |\phi_k| |k|^j \\ &\leq \left[\sum_{|k|>N} \frac{1}{|k|^{2(s-j)}} \right]^{1/2} \left[\sum_{|k|>N} |k|^{2s} |\phi_k|^2 \right]^{1/2} \\ &\leq \left\{ \frac{2}{[2(s-j)-1] N^{2(s-j)-1}} \right\}^{1/2} \|\phi\|_{H^s} \end{aligned}$$

$\square$

Lemma 5.6. *Let ϕ and ψ be such that*

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \sum_{|k|>N} \psi_k e^{ikx} \ , \quad \phi(x) = \frac{1}{\sqrt{2\pi}} \sum_{|k|>N} \phi_k e^{ikx} \ ,$$

then, for any integer $s \geq 1$, one has

$$\|\psi\phi\|_{H^s} \leq \sqrt{\frac{2}{\pi}} \frac{2^{s+1}}{N^{s-1}} \|\psi\|_{H^s} \|\phi\|_{H^s} \ . \tag{5.9}$$

Proof. The main point consists in estimating the L^2 norm of

$$\frac{d^s}{dx^s}(\psi\phi) = \sum_{j=0}^{s} \begin{pmatrix} s \\ j \end{pmatrix} \psi^{(j)}\phi^{(s-j)}$$

Denote by D_j the $j-th$ addendum at r.h.s. For $1 \leq j \leq s-1$ one has (by lemma 5.5)

$$|D_j(x)| \leq \begin{pmatrix} s \\ j \end{pmatrix} \frac{2}{\pi} \frac{1}{N^{s-1}} \|\phi\|_{H^s} \|\psi\|_{H^s} ,$$

moreover,

$$|D_0(x)| \leq \sqrt{\frac{2}{\pi}} \frac{1}{N^{s-1/2}} \|\psi\|_{H^s} \left|\phi^{(s)}(x)\right| \ , \quad |D_s(x)| \leq \sqrt{\frac{2}{\pi}} \frac{1}{N^{s-1/2}} \|\phi\|_{H^s} \left|\psi^{(s)}(x)\right|$$

from which

$$\|\phi\psi\|_{L^2} + \sum_{j=0}^{s} \|D_j\|_{L^2} \leq 2\sqrt{\frac{2}{\pi}} \frac{1}{N^{s-1}} \|\phi\|_{H^s} \|\psi\|_{H^s} \sum_{j=0}^{s} \begin{pmatrix} s \\ j \end{pmatrix} \ ,$$

using $\sum_{j=0}^{s} \begin{pmatrix} s \\ j \end{pmatrix} = 2^s$ the thesis follows. □

Lemma 5.7. *Let $\check{\check{h}}$ be of degree at least three in P, Q, namely with $m_3 + m_4 \geq 3$, then one has*

$$\left\|X_{\check{\check{h}}}\right\|_s^R \leq n \frac{\Sigma_s}{N^{s-1}} (\rho_s R)^{n-1} \tag{5.10}$$

where $n = m_1 + m_2 + m_3 + m_4$.

Proof. For simplicity we will consider only the case $m_3 \geq 3$, the other cases can be treated exactly in the same way. We proceed as in the proof of lemma 5.2, but consider first the P component of the vector field. Its norm is estimated by

$$\begin{aligned}
(2\pi)^{\frac{n-1}{2}} \left\|X^P_{\check{\check{h}}}\right\|_{H^s} &= m_3 \left\|F\, q^{m_1} p^{m_2} Q^{m_3-1} P^{m_4}\right\|_{H^s} \\
&\leq m_3 \left\|q^{m_1} p^{m_2} Q^{m_3-3} P^{m_4} F\right\|_{H^s} \sqrt{\frac{2}{\pi}} 2^{s+1} \left\|Q^2\right\|_{H^s} \\
\leq m_3 \left\|F q^{m_1} p^{m_2} Q^{m_3-3} P^{m_4}\right\|_{H^s} &\left(\sqrt{\frac{2}{\pi}} 2^{s+1}\right)^2 \frac{1}{N^{s-1}} \|Q\|_{H^s} \|Q\|_{H^s} \\
&\leq m_3 \Sigma_s \frac{1}{N^{s-1}} \left(\sqrt{\frac{2}{\pi}} 2^{s+1} R\right)^{n-1} \ .
\end{aligned}$$

Considering the other components and summing up one has the thesis. □

Finally we have

Proposition 5.8. *Let h be a polynomial of degree less or equal than r and of class $\mathcal{M}^{\mu}_{R}$, assume also that it is at least cubic in (P,Q) namely that it is the sum of polynomials of the form (5.8) each one fulfilling $m_3+m_4\geq 3$, then, for any positive $R^{(s)}$ such that $\rho_s R^{(s)}\leq R$ and any $\delta^{(s)}<R^{(s)}$, one has*

$$\left\|X_h\right\|_s^{R^{(s)}-\delta^{(s)}} \leq \frac{2^{r+1}}{\rho_s\delta^{(s)}}\frac{\Sigma_s}{N^{s-1}}\left\langle|h|\right\rangle^{\mu}_{\rho_s R^{(s)}} \ . \tag{5.11}$$

Proof. First consider a polynomial h^{n_1,n_2} of degree n_1,n_2 in ξ and η respectively. Expand it in Taylor series in P,Q. Such a Taylor series contains at most $2^{n_1+n_2}$ hogeneous polynomials. Denote by h^{n_1,n_2,m_3,m_4} the polynomial of degree m_3,m_4 in Q,P respectively. Then one has

$$\left\langle|h^{n_1,n_2,m_3,m_4}|\right\rangle^{\mu}_{R} \leq \left\langle|h^{n_1,n_2}|\right\rangle^{\mu}_{R} \ ,$$

and thus, by lemma 5.7, working as in the proof of proposition 5.3 one has

$$\left\|X_{h^{n_1,n_2,m_3,m_4}}\right\|_s^{R^{(s)}-\delta^{(s)}} \leq \frac{1}{\rho_s\delta^{(s)}}\frac{\Sigma_s}{N^{s-1}}\left\langle|h^{n_1,n_2,m_3,m_4}|\right\rangle^{\mu}_{\rho_s R^{(s)}} \ . \tag{5.12}$$

Summing over all the different polynomials one gets at most the factor 2^{r+1} and the desired result. □

5.2 Lie transform estimates

In this sub–section we first recall some facts about Lie the transform, and then we estimate the $\mathcal{M}^{\mu}_{R-d}$ norm of the Lie transform of an $\mathcal{M}^{\mu}_{R}$ function. This is given by lemma 5.12 below, which constitutes the main result of this sub–section.

Consider a function χ and the corresponding Hamilton equations $\dot\zeta=X_\chi(\zeta)$, the corresponding flow T^t is a canonical transformation, and $\mathcal{T}:=T^1\equiv T^t\big|_{t=1}$ is called the *Lie transform* generated by χ. Given an analytic function h one can consider the transformed function $h\circ\mathcal{T}$. We recall that one has

$$h\circ\mathcal{T}=\sum_{l=0}^{\infty}h_l \ , \tag{5.13}$$

where the h_l's are defined by

$$h_l=\frac{1}{l}\left\{\chi,h_{l-1}\right\} \ , \quad l\geq 1 \ ; \quad h_0=h \ , \tag{5.14}$$

as is easily seen by iterating the equality

$$\frac{d}{dt}[h \circ T^t] = \{\chi, h\} \circ T^t \ ,$$

and considering the Taylor expansion in t of $h \circ T^t$. In order to estimate the norm of $h \circ \mathcal{T}$ we need first to estimate the Poisson bracket of two $\mathcal{M}^{\mu}_{R}$ functions. We will do it in a few steps.

Lemma 5.9. *Let h be a homogeneous polynomial of degree m and g a homogeneous polynomial of degree n; assume that both h, g are homogenea in ξ and η independently, and that they have finite μ-modulus, then one has*

$$|\{h, g\}|_{\mu} \leq mn|h|_{\mu}\,|g|_{\mu} \ , \tag{5.15}$$

Proof. Write h and g explicitely as follows

$$h(\xi, \eta) = \sum_{i_1,...,i_{n_1},j_1,...,j_{m_1}} h_{i_1,...,i_{n_1},j_1,...,j_{m_1}} \xi_{i_1}...\xi_{i_{n_1}} \eta_{j_1}...\eta_{j_{m_1}}$$

$$g(\xi, \eta) = \sum_{k_1,...,k_{n_2},l_1,...,l_{m_2}} g_{k_1,...,k_{n_2},l_1,...,l_{m_2}} \xi_{k_1}...\xi_{k_{n_2}} \eta_{l_1}...\eta_{l_{m_2}} \ ,$$

One has

$$\varphi := \{h, g\} = i \sum_t \left(\frac{\partial h}{\partial \xi_t} \frac{\partial g}{\partial \eta_{-t}} - \frac{\partial g}{\partial \xi_t} \frac{\partial h}{\partial \eta_{-t}} \right) \ . \tag{5.16}$$

We first consider the case where h depends on ξ only and g depends on η: We have

$$\varphi(\xi, \eta) = \sum_{i_2,...,i_n,j_2,...,j_m} \varphi_{i_2,...,i_n,j_2,...,j_m} \xi_{i_2}, ..., \xi_{i_n} \eta_{j_2}...\eta_{j_m}$$

with

$$\varphi_{i_2,...,i_n,j_2,...,j_m} = mn \sum_{j_1} h_{-j_1,i_2,...,i_n} g_{j_1,j_2,...,j_m} \ .$$

In order to estimate the μ–modulus of φ we write it explicitely as follows:

$$\frac{1}{nm}\,|\varphi|_{\mu} = \sum_l \sup_{i_2,...,i_n,j_2,...,j_{m-1}} \left| \sum_{j_1} g_{j_1,...,j_{m-1},l-\sum_{a=2}^n i_a - \sum_{b=2}^{m-1} j_b} h_{-j_1,i_2,...,i_n} \right| e^{\mu|l|}$$

Chainging the variable in the sum from j_1 to $l_1 := -j_1 + \sum_{a=2}^n i_a$, this becomes

$$\sum_l \sup_{i_2,...,i_n,j_2,...,j_{m-1}} \sum_{l_1} \left| g_{\sum_{a=2}^n i_a - l_1, j_2,...,j_{m-1},l-\sum_{a=2}^n i_a - \sum_{b=2}^{m-1} j_b} h_{l_1 - \sum_{a=2}^n i_a, i_2,...,i_n} \right| e^{\mu|l|}$$

$$\leq \sum_l \sum_{l_1} \sup_{i_2,..,i_n,j_2,..,j_{m-1}} \left| g_{\sum_{a=2}^n i_a - l_1, j_2,...,j_{m-1},l-\sum_{a=2}^n i_a - \sum_{b=2}^{m-1} j_b} h_{l_1 - \sum_{a=2}^n i_a, i_2,...,i_n} \right| e^{\mu|l|}$$

Changing the variable of summation l to $l_2 = l - l_1$ one can estimate the above quantity by

$$\sum_{l_2}\sum_{l_1} \sup_{i_2,\ldots,i_n,j_2,\ldots,j_{m-1}} \left| g_{\sum_{a=2}^n i_a - l_1, j_2,\ldots,j_{m-1}, l_2+l_1-\sum_{a=2}^n i_a - \sum_{b=2}^{m-1} j_b} e^{\mu|l_2|} \right. \\ \left. \times\, h_{l_1 - \sum_{a=2}^n i_a, i_2,\ldots,i_n} \right| e^{\mu|l_1|}$$

$$\leq \left\{ \sum_{l_2} \sup_{i_2,\ldots,i_n,j_2,\ldots,j_{m-1},l_1} \left| g_{\sum_{a=2}^n i_a - l_1, j_2,\ldots,j_{m-1}, l_1+l_2-\sum_{a=2}^n i_a - \sum_{b=2}^{m-1} j_b} \right| e^{\mu|l_2|} \right\} \\ \times \left\{ \sum_{l_1} \sup_{i_2,\ldots,i_n} \left| h_{l_1 - \sum_{a=2}^n i_a, i_2,\ldots,i_n} \right| e^{\mu|l_1|} \right\}$$

but the first supremum is taken over all the combination of indices such that their sum is l_2, and therefore the first curly bracket is just the μ–modulus of g, the second is the μ–modulus of h. So the thesis follows in the present case.

The general case follows using

$$|\{h, g\}|_\mu \leq (n_1 m_2 + n_2 m_1)|h|_\mu\, |g|_\mu \leq (n_1 + m_1)(m_2 + n_2)|h|_\mu\, |g|_\mu\ ,$$

which is the thesis. $\square$

Remark 5.10. *Equation (5.15) holds also in the case polynomials which are not homogenea in ξ, η, to verify this point just use the definition of the μ–modulus.*

Lemma 5.11. *Let $h \in \mathcal{M}_R^\mu$ and $g \in \mathcal{M}_{R-d}^\mu$, then for any positive $d' < R - d$ one has $\{h, g\} \in \mathcal{M}_{R-d-d'}^\mu$*

$$\langle|\{h, g\}|\rangle_{R-d-d'}^\mu \leq \frac{1}{d'(d + d')} \langle|h|\rangle_R^\mu\, \langle|g|\rangle_{R-d}^\mu \tag{5.17}$$

Proof. Write $h = \sum_j h_j$ and $g = \sum_k g_k$ with h_j homogeneous of degree j and similarly for g, we have

$$\{h, g\} = \sum_{j,k} \{h_j, g_k\}\ .$$

Now each term of the series is estimated by

$$\begin{aligned} \langle|\{h_k, g_k\}|\rangle_{R-d-d'}^\mu &= |\{h_j, g_k\}|_\mu (R - d - d')^{j+k-2} \\ &\leq |h_j|_\mu\, |g_k|_\mu jk(R - d - d')^{j+k-2} \\ &= |h_j|_\mu\, |g_k|_\mu j(R - d - d')^{j-1} k(R - d - d')^{k-1} \\ \leq |h_j|_\mu\, |g_k|_\mu R^j \frac{1}{d + d'} \frac{1}{d'} (R - d)^k &= \frac{1}{d'(d + d')} \langle|h_j|\rangle_R^\mu\, \langle|g_k|\rangle_{R-d}^\mu\ . \end{aligned} \tag{5.18}$$

$\square$

We estimate now the terms of the series (5.13), (5.14) defining the Lie transform.

Lemma 5.12. *Let $h \in \mathcal{M}^{\mu}_{R}$ and $\chi \in \mathcal{M}^{\mu}_{R}$ be two analytic functions, denote by h_n the functions defined recoursively by (5.14); then, for any positive $d < R$, one has $h_n \in \mathcal{M}^{\mu}_{R-d}$, and the following estimate holds*

$$\langle|h_n|\rangle^{\mu}_{R-d} \leq \langle|h|\rangle^{\mu}_{R} \left(\frac{e^2}{d^2}\langle|\chi|\rangle^{\mu}_{R}\right)^n \ . \tag{5.19}$$

Proof. Fix n, and denote $\tilde{\delta} := d/n$, we look for a sequence $C^{(n)}_l$ such that

$$\langle|h_l|\rangle^{\mu}_{R-\tilde{\delta}l} \leq C^{(n)}_l \ , \quad \forall l \leq n \ .$$

By (5.17) this sequence can be defined by

$$C^{(n)}_0 = \langle|h|\rangle^{\mu}_{R} \ , \quad C^{(n)}_l = \frac{1}{l}\frac{1}{l\tilde{\delta}\tilde{\delta}}C^{(n)}_{l-1}\langle|\chi|\rangle^{\mu}_{R} = \frac{n^2}{l^2\delta^2}C^{(n)}_{l-1}\langle|\chi|\rangle^{\mu}_{R} \ .$$

So one has

$$C^{(n)}_n = \frac{1}{(n!)^2}\left(\frac{n^2\langle|\chi|\rangle^{\mu}_{R}}{\delta^2}\right)^n \langle|g|\rangle^{\mu}_{R} \ .$$

Using the inequality $n^n < n!e^n$, which is easily verified by writing the iterative definition of $n^n/n!$ one has the thesis. $\square$

Remark 5.13. *Let χ be an analytic function with Hamiltonian vector field which is analytic as a map from $\mathcal{B}_s(R^{(s)})$ to $\mathcal{P}_s$, fix $\delta^{(s)} < R^{(s)}$. Assume $\|X_\chi\|^{R^{(s)}}_s < \delta^{(s)}$ and consider the time t flow T^t of X_χ. Then, for $|t| \leq 1$, one has*

$$\sup_{\|\zeta\|_s \leq R^{(s)}-\delta^{(s)}} \left\|T^t(\zeta) - \zeta\right\|_s \leq \|\chi\|^{R^{(s)}}_s \ . \tag{5.20}$$

Lemma 5.14. *Consider χ as above and let $h : \mathcal{B}_s\left(R^{(s)}\right) \to \mathbf{C}$ be an analytic function with vector field analytic in $\mathcal{B}_s\left(R^{(s)}\right)$, fix $0 < \delta^{(s)} < R^{(s)}$ assume $\|X_\chi\|^{R^{(s)}}_s \leq \delta^{(s)}/3$, then, for $|t| \leq 1$, one has*

$$\|X_{h\circ T^t}\|^{R^{(s)}-\delta^{(s)}}_s \leq \left(1 + \frac{3}{\delta^{(s)}}\|X_\chi\|^{R^{(s)}}_s\right)\|X_h\|^{R^{(s)}}_s$$

For the proof see [2] proof of lemma 8.2.

5.3 Iterative lemma

In the statement of the forthcoming iterative lemma we will use the following notations: For any positive R, define $\delta := R/2r$ and $R_k := R - k\delta$. Moreover, for any positive $R^{(s)}$ define $\delta^{(s)} := R^{(s)}/2r$ and $R_k^{(s)} := R^{(s)} - k\delta^{(s)}$.

Proposition 5.15. Iterative Lemma. *Fix $r \geq 4$ and consider the Hamiltonian (4.1); define R_* by*

$$\frac{1}{R_*} := \frac{16\mathrm{e}^2 A N^\alpha r^2}{\gamma} \ ,$$

and assume

$$R < \bar{R}_1 \ , \quad \frac{R}{R_*} \leq \frac{1}{2} \ ; \tag{5.21}$$

then, for any $k \leq r-4$ there exists a canonical transformation $\mathcal{T}^{(k)}$ which puts (4.1) in the form

$$H^{(k)} := H \circ \mathcal{T}^{(k)} = h_0 + Z^{(k)} + f^{(k)} + \mathcal{R}_N^{(k)} + \mathcal{R}_T^{(k)} \ , \tag{5.22}$$

with $Z^{(k)}$ in Birkhoff normal form; moreover

1) the following estimates hold

$$\begin{aligned} \left\langle \left| Z^{(k)} \right| \right\rangle_{R_k}^{\mu} &\leq AR^3 \sum_{l=0}^{k-1} \left(\frac{R}{R_*} \right)^l \\ \left\langle \left| f^{(k)} \right| \right\rangle_{R_k}^{\mu} &\leq AR^3 \left(\frac{R}{R_*} \right)^k \end{aligned} \tag{5.23}$$

2) for any $s \geq 1$ define

$$R_*^{(s)} := \frac{R_*}{\rho_s}$$

then, for any $R^{(s)}$ such that

$$\frac{R^{(s)}}{R_*^{(s)}} \leq \min \left\{ \frac{1}{2}, \frac{2\mathrm{e}^2}{\rho_s \Sigma_s} \right\} \tag{5.24}$$

the transformation $\mathcal{T}^{(k)}$ is analytic as a map

$$\mathcal{T}^{(k)} : \mathcal{B}_s\left(R_{k+1}^{(s)}\right) \to \mathcal{B}_s\left(R_1^{(s)}\right) \ , \tag{5.25}$$

and the following estimates hold

$$\sup_{\|\zeta\|_s \leq R_{k+1}^{(s)}} \left\| \zeta - \mathcal{T}^{(k)}(\zeta) \right\|_s \leq R^{(s)} \frac{\rho_s \Sigma_s}{16\mathrm{e}^2 r} \sum_{l=1}^{k} \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^l \tag{5.26}$$

$$\begin{aligned} \left\| X_{\mathcal{R}_N^{(k)}} \right\|_s^{R_{k+1}^{(s)}} \leq \frac{2^{r+1}}{N^{s-1}} 2rA\Sigma_s (\rho_s R^{(s)})^2 &\left[\prod_{l=0}^{k-1} \left(1 + \frac{3\rho_s \Sigma_s}{16\mathrm{e}^2} \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^{l+1} \right) \right] \\ \times \sum_{l=0}^{k-1} \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^l & \end{aligned} \tag{5.27}$$

$$\left\| X_{\mathcal{R}_T^{(k)}} \right\|_s^{R_{k+1}^{(s)}} \leq \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^{r-1} \Sigma_s 2^{r-2} r A R_*^2 \left[\prod_{l=0}^{k-1} \left(1 + \frac{3\rho_s \Sigma_s}{16\mathrm{e}^2} \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^{l+1} \right) \right] \times \sum_{l=0}^{k-1} \frac{1}{2^l} \; . \tag{5.28}$$

furthermore if the vector fields of h_0 and of f leave invariant $\mathcal{A}$ and/or $\mathcal{R}e$, then the same is true for $\mathcal{T}^{(k)}$ and for the Hamiltonian vector fields of all the terms of (5.22),

Remark 5.16. *In particular $X_{\mathcal{R}_N^{(k)}}$ is of order $[R^{(s)}]^2/N^{s-1}$ and $X_{\mathcal{R}_T^{(k)}}$ is of order $(R^{(s)}/R_*^{(s)})^{r-1}$.*

Proof. We proceed by induction. First we split $f^{(k)}$ (f in the case $k=0$) into an effective part and a remainder. Write

$$f^{(k)} = f_0^{(k)} + f_T^{(k)}$$

where $f_0^{(k)}$ is the the Taylor polynomial of $f^{(k)}$ truncated at order $r-1$ and $f_T^{(k)}$ is the remainder of the Taylor series (which begins with terms of order r). Since $f_0^{(k)}$ is a truncation of $f^{(k)}$one has

$$\left\langle \left| f_0^{(k)} \right| \right\rangle_{R_k}^{\mu} \leq \left\langle \left| f^{(k)} \right| \right\rangle_{R_k}^{\mu} \; .$$

Consider $f_0^{(k)}$, and its Taylor series *in* (P,Q) *only*, we write

$$f_0^{(k)} = \tilde{f}^{(k)} + f_N^{(k)}$$

where $\tilde{f}^{(k)}$ is the truncation of such a series at second order (it contains at most terms quadratic in (P,Q)) and $f_N^{(k)}$ is the remainder of the expansion. One has

$$\left\langle \left| f_N^{(k)} \right| \right\rangle_{R_k}^{\mu} \leq \left\langle \left| f^{(k)} \right| \right\rangle_{R_k}^{\mu} \; , \quad \left\langle \left| \tilde{f}^{(k)} \right| \right\rangle_{R_k}^{\mu} \leq \left\langle \left| f^{(k)} \right| \right\rangle_{R_k}^{\mu} \; .$$

Rewrite

$$\tilde{H}^{(k)} := h_0 + Z^{(k)} + \tilde{f}^{(k)} + \tilde{\mathcal{R}}_N^{(k)} + \tilde{\mathcal{R}}_T^{(k)}$$

where

$$\tilde{\mathcal{R}}_N^{(k)} := \mathcal{R}_N^{(k)} + f_N^{(k)} \; , \quad \tilde{\mathcal{R}}_T^{(k)} := \mathcal{R}_T^{(k)} + f_T^{(k)} \; .$$

Consider the Lie transform $\mathcal{T}_k$ generated by a function χ_k, and use it to tranform the main part of the Hamiltonian, namely $h_0 + Z^{(k)} + \tilde{f}^{(k)}$; by formulae (5.13), (5.14) one has

$$\left[h_0 + Z^{(k)} + \tilde{f}^{(k)} \right] \circ \mathcal{T}_k = h_0 + Z^{(k)} + \left[\{\chi_k, h_0\} + \tilde{f}^{(k)} \right] + \left\{ \sum_{l=1}^{\infty} Z_l^{(k)} + \sum_{l=1}^{\infty} \tilde{f}_l^{(k)} + \sum_{l=2}^{\infty} h_{0l} \right\} \; , \tag{5.29}$$

where

$$Z_l^{(k)} := \frac{1}{l}\left\{\chi, Z_{l-1}^{(k)}\right\} \ , \ l \geq 1 \ ; \quad Z_0^{(k)} = Z^{(k)} \ ,$$

and similarly for $\tilde{f}_l^{(k)}$ and h_{0l}. If χ_k fulfills the equation

$$\{\chi_k, h_0\} + \tilde{f}^{(k)} = Z_k \ , \tag{5.30}$$

with Z_k in normal form, then the order of the non nonormalized part has increased by one.

We determine a function χ_k fulfilling (5.30). To this end expand $\tilde{f}^{(k)}$ in Taylor series:

$$\tilde{f}^{(k)}(\xi, \eta) = \sum_{|l|+|j|\leq r} f_{lj}^{(k)} \xi^j \eta^l \ ,$$

where $l = \{l_i\}_{i\in\mathbf{Z}}$ and $j = \{j_i\}_{i\in\mathbf{Z}}$ are sequences of nonnegative integers, and we used the vector notation $\xi^j = \prod_{i\in\mathbf{Z}} \xi_i^{j_i}$. Denote by $\mathcal{RE} := \{(l, j) \ : \ \omega\cdot(j-l) = 0\}$, and define

$$Z_k(\xi, \eta) := \sum_{(l,j)\in\mathcal{RE}} f_{lj}^{(k)} \xi^l \eta^j \ , \quad \chi_k(\xi, \eta) := \sum_{(l,j)\notin\mathcal{RE}} \frac{f_{lj}^{(k)}}{i\omega\cdot(j-l)} \xi^j \eta^l \ ,$$

then Z_k is in normal form. The norms of these functions are estimates by

$$\langle|Z_k|\rangle_R^\mu \leq \left\langle\left|\tilde{f}^{(k)}\right|\right\rangle_R^\mu \ , \quad \langle|\chi_k|\rangle_R^\mu \leq \frac{N^\alpha}{\gamma}\left\langle\left|\tilde{f}^{(k)}\right|\right\rangle_R^\mu \ ,$$

as one can easily see by remarking that Z_k is a truncation of $\tilde{f}^{(k)}$ and that the denominators in the definition of χ are larger than γ/N^α. Then, using the inductive hypothesis, one has

$$\langle|Z_k|\rangle_{R_k}^\mu \leq AR^3\left(\frac{R}{R_*}\right)^k \ , \quad \langle|\chi_k|\rangle_{R_k}^\mu \leq \frac{AN^\alpha}{\gamma}R^3\left(\frac{R}{R_*}\right)^k . \tag{5.31}$$

Define $\mathcal{T}^{(k+1)} := \mathcal{T}^{(k)} \circ \mathcal{T}_k$ and remark that, by the standard theory of systems with symmetry, if $\mathcal{T}^{(k)}$ and the Hamiltonian vector field of the different parts of $\tilde{H}^{(k)}$ leave invariant the manifold $\mathcal{A}$ and/or $\mathcal{R}e$, then the same is true for $\mathcal{T}^{(k+1)}$.

Then define $Z^{(k+1)} := Z^{(k)} + Z_k$, and $f^{(k+1)}$ as the curly bracket of (5.29). For $Z^{(k+1)}$ clearly (5.23) holds with $k+1$ in place of k. We now estimate $f^{(k+1)}$. To this end let us denote

$$\epsilon := \frac{\mathrm{e}^2}{\delta^2}\langle|\chi|\rangle_{R_k}^\mu \ ,$$

and notice that, by (5.31), and the definition of R_*, we have

$$\epsilon \leq \frac{1}{4}\left(\frac{R}{R_*}\right)^{k+1} < 1/2 \ , \quad \forall k$$

So, using the estimate (5.19), from lemma 5.12, we have

$$\left\langle\left|\sum_{l=1}^{\infty} Z_l^{(k)}\right|\right\rangle_{R_{k+1}}^{\mu} \leq \sum_{l=1}^{\infty}\epsilon^l \left\langle\left|Z^{(k)}\right|\right\rangle_{R_k}^{\mu} = \frac{\epsilon}{1-\epsilon}\left\langle\left|Z^{(k)}\right|\right\rangle_{R_k}^{\mu} \\ \leq 4\epsilon A R^{1+1}$$

and similarly

$$\left\langle\left|\sum_{l=1}^{\infty} \tilde{f}_l^{(k)}\right|\right\rangle_{R_{k+1}}^{\mu} \leq 2AR^2\epsilon\left(\frac{R}{R_*}\right)^k .$$

Finally, even if h_0 is not of class $\mathcal{M}_{R_k}^{\mu}$, all the terms of the sequence h_{0l} generated by h_0 turn out to be of class $\mathcal{M}_{R_{k+1}}^{\mu}$. Indeed, remarking that

$$h_{01} = \{\chi_k, h_0\} = \tilde{f}^{(k)} - Z_k \in \mathcal{M}_{R_k}^{\mu}$$

and proceeding as in the proof of lemma 5.12 one gets

$$\langle|h_{0l}|\rangle_{R_{k+1}}^{\mu} \leq \epsilon^{r-1}\left\langle\left|\tilde{f}^{(k)}\right|\right\rangle_{R_k}^{\mu} \ , \quad l\geq 1 \ ,$$

and therefore

$$\left\langle\left|\sum_{l=2}^{\infty} h_{0k}\right|\right\rangle_{R_{k+1}}^{\mu} \leq 2AR^{1+1}\epsilon\left(\frac{R}{R_*}\right)^k ;$$

Summing up the different contributions to $f^{(k+1)}$ and using the definition of ϵ, we obtain the second of (5.23) with $k+1$ in place of k.

We come to item 2). First we define

$$\mathcal{R}_N^{(k+1)} := \tilde{\mathcal{R}}_N^{(k)} \circ \mathcal{T}_k \ , \quad \mathcal{R}_T^{(k+1)} := \tilde{\mathcal{R}}_T^{(k)} \circ \mathcal{T}_k$$

and remark that, by proposition 5.3, one has

$$\|X_{\chi_k}\|_s^{R_{k+1}^{(s)}} \leq \frac{\Sigma_s}{\rho_s\delta^{(s)}}\frac{AN^\alpha}{\gamma}(\rho_s R^{(s)})^3\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^k = \frac{R^{(s)}\rho_s\Sigma_s}{16\mathrm{e}^2 r}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{k+1} \tag{5.32}$$

where we used the definition of R_*; The above expression, by lemma 5.13 and by (5.24), ensures that

$$\mathcal{T}_k : \mathcal{B}_s\left(R_{k+2}^{(s)}\right) \to \mathcal{B}_s\left(R_{k+1}^{(s)}\right) \ ,$$

which together with the inductive assumptions ensures the validity of the estimate (5.26) with $k+1$ in place of k.

Estimate of $\mathcal{R}_N^{(k+1)}$: To begin with remark that one has

$$\left\|X_{\tilde{f}_N^{(k)}}\right\|_s^{R_{k+1}^{(s)}} \le \frac{A\Sigma_s(\rho_s R^{(s)})^3}{\rho_s\delta^{(s)}}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^k \frac{2^{r+1}}{N^{s-1}}$$
$$= 2rA\Sigma_s\left(\rho_s R^{(s)}\right)^2\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^k \frac{2^{r+1}}{N^{s-1}} ,$$

and, from lemma 5.14 one has

$$\left\|X_{f_N^{(k)}\circ\mathcal{T}_k}\right\|_s^{R_{k+2}^{(s)}} \le \left(1+\frac{3\rho_s\Sigma_s}{16e^2}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{k+1}\right)2rA\Sigma_s\left(\rho_s R^{(s)}\right)^2\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^k\frac{2^{r+1}}{N^{s-1}} ,$$

and also

$$\left\|X_{\mathcal{R}_N^{(k)}\circ\mathcal{T}_k}\right\|_s^{R_{k+2}^{(s)}} \le \left(1+\frac{3\rho_s\Sigma_s}{16e^2}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{k+1}\right)\frac{2^{r+1}}{N^{s-1}}2rA\Sigma_s$$
$$\times(\rho_s R^{(s)})^2\left[\prod_{l=0}^{k-1}\left(1+\frac{3\rho_s\Sigma_s}{16e^2}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{l+1}\right)\right]\sum_{l=0}^{k-1}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^l$$

which shows that the estimate (5.27) holds with $k+1$ in place of k.

Estimate of (5.28): First we estimate the $\mathcal{M}_{R_k}^{\mu}$ norm of $f_T^{(k)}$. To this end remark that $\left\langle\left|f^{(k)}\right|\right\rangle_{R_k}^{\mu}$ is a function of R which is analytic and has all the Taylor coefficients that are positive. Moreover, the quantity $\left\langle\left|f_T^{(k)}\right|\right\rangle_{R_k}^{\mu}$ is the sum of the terms of order higher than r in such a series. From this one easily sees that

$$\left\langle\left|f_T^{(k)}\right|\right\rangle_{R_k}^{\mu} \le \left(\frac{2R}{R_*}\right)^r\left\langle\left|f^{(k)}\right|\right\rangle_{R_{*k}/2}^{\mu} \le 2^{r-3}AR_*^3\left(\frac{R}{R_*}\right)^r\frac{1}{2^k} . \tag{5.33}$$

Then, by proposition 5.3 one has

$$\left\|X_{f_T^{(k)}}\right\|_s^{R_{k+1}^{(s)}} \le \Sigma_s\frac{2r}{\rho_s R^{(s)}}2^{r-3}AR_*^3\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^r\frac{1}{2^k}$$
$$= \Sigma_s 2^{r-2}rAR_*^2\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{r-1}\frac{1}{2^k} .$$

Finally, proceeding as in the estimate of $\mathcal{R}_N^{(k)}$ one obtains the thesis. □

Corollary 5.17. *Consider the Hamiltonian (4.1); There exists a constant C such that for any $s\ge 1$ and any $R^{(s)}$ such that*

$$R^{(s)} < (C2^sN^\alpha\Sigma_s^\alpha)^{-1}$$

the following holds true: There exists a canonical transformation $\mathcal{T}_r : \mathcal{B}_s(R^{(s)}/3) \to \mathcal{B}_s(R^{(s)})$ *fulfilling*

$$\|\zeta - \mathcal{T}_r(\zeta)\|_s \le C 2^s \Sigma_s (R^{(s)} 2^s N^\alpha)^2 R^{(s)}$$

such that the transformed Hamiltonian has the form

$$H \circ \mathcal{T}_r = h_0 + Z + \mathcal{R}_T + \mathcal{R}_N$$

with Z *in normal form, and the following estimates hold*

$$\begin{aligned} \|X_{\mathcal{R}_N}\|_s^{R^{(s)}/3} &\le \frac{1}{N^{s-1}} C\Sigma_s \left(2^s R^{(s)}\right)^2 \\ \|X_{\mathcal{R}_T}\|_s^{R^{(s)}/3} &\le C\Sigma_s \left(R^{(s)} 2^s N^\alpha\right)^{r-1} \end{aligned} \tag{5.34}$$

Proof. Apply proposition 5.15 with $k = k_r = (r-3)$, then estimate the vector field of $f^{(k_r)}$ by using (5.7), and define $\mathcal{R}_N := \mathcal{R}_N^{(k_r)}$ and $\mathcal{R}_T := \mathcal{R}_T^{(k_r)} + f^{(k_r)}$. Inserting the definition of $R_*^{(s)}$ and inserting all irrelevant constants in C we get the result. $\square$

We are now ready for the

Proof of Theorem 4.1. In corollary 5.17 we have N and s at our disposal. We begin by choosing $N := R^{-1/2\alpha}$ so that the bracket in the second of (5.34) becomes of order $[R^{(s)}]^{1/2}$. Inserting in (5.34) we get

$$\|X_{\mathcal{R}_N} + X_{\mathcal{R}_T}\|_s^{R^{(s)}/3} \le C\Sigma_s (R^{(s)})^{\frac{1}{2\alpha}(s-1)} + C\Sigma_s \left(4^s C R^{(s)}\right)^{\frac{r-1}{2}} .$$

Finally choosing $s \ge s_* := 4\alpha M$ with $M = (r-1)/2$ one has that the first addendum is smaller than the second. To obtain the statement use also eq. (5.1). $\square$

6. Application to the nonlinear wave equation

We split this section in two subections: in the first we will show how to apply theorem 4.1 in order to obtain theorem 2.2, and in the second we will show that the frequencies of the system we study are strongly nonvanishing for most values of the parameter m.

6.1 Application of theorem 4.1 to equation (0.1)

Consider the eigenfunctions φ_j of the Sturm–Liouville problem (2.1) and their Fourier expansion

$$\varphi_j(x) = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbf{Z}} \varphi_k^j e^{ikx} ,$$

it is known [11,9] that one has $|\varphi_k^j| \le C\mathrm{e}^{-\sigma||k|-j|}$ with positive $\sigma < \sigma_1$ and a suitable C. Remark also that $\varphi_0^j = 0$. We want to introduce the analogue for our problem of the Fourier basis of the exponentials. So, we define

$$\psi_j(x) := \frac{1}{\sqrt{2\pi}} \sum_{k>0} \varphi_k^j \mathrm{e}^{ikx} \ , \quad \psi_{-j}(x) := \frac{1}{\sqrt{2\pi}} \sum_{k<0} \varphi_k^j \mathrm{e}^{ikx} \ , \quad j > 0 \ ;$$

considering the Fourier expansion of the ψ_j's, one has

$$\psi_j(x) = \frac{1}{\sqrt{2\pi}} \sum_{k\in\bar{\mathbf{Z}}} \psi_k^j \mathrm{e}^{ikx} \ , \quad |\psi_k^j| \le C\mathrm{e}^{-\sigma|k-j|} \tag{6.1}$$

which is the property we need.

We fix the phase space to be $\mathcal{F}_s^p := \ell_s^2 \oplus \ell_{s-1}^2 \ni (\hat{q},\hat{p})$ endowed by the symplectic form $i\sum_{k\in\bar{\mathbf{Z}}} dp_k \wedge dq_{-k}$; for any point $(\hat{q},\hat{p}) \in \mathcal{F}_s^p$ we define the functions

$$\hat{u}(x) := \sum_{k\in\bar{\mathbf{Z}}} \hat{q}_k \psi_k(x) \ , \quad \hat{v}(x) := \sum_{k\in\bar{\mathbf{Z}}} \hat{p}_k \psi_k(x) \ ; \tag{6.2}$$

their skew–symmetric part will play the role of the uncknown of equation (0.1).

Remark 6.1. *The norm of $\mathcal{F}_s^p$ is defined in terms of the coefficients of the expansion of $(\hat{u},\hat{v})$ on the basis of the eigenfunction of a Sturm Liouville problem with analytic potential, so it is equivalent to the standard Sobolev norm of the corresponding functions. In particular (6.2) establishes an isomorphism of $\mathcal{F}_s^p$ with the Sobolev space $H^s(\mathbb{R}/2\pi\mathbf{Z}) \oplus H^{s-1}(\mathbb{R}/2\pi\mathbf{Z})$.*

Consider the Hamiltonian system

$$\sum_{k\in\bar{\mathbf{Z}}} \frac{\hat{p}_k\hat{p}_{-k} + \omega_k^2 \hat{q}_k\hat{q}_{-k}}{2} + \int_{-\pi}^{\pi} G(x,\hat{u}(x))dx \ , \tag{6.3}$$

with $\omega_{-k} := \omega_k \equiv \sqrt{\mu_k + m}$, $(k \ge 1)$ and $\hat{u}$ which has to be thought as a function of $\hat{q}$.

Remark 6.2. *In general the functions ψ_j are not eigenfunctions of the Sturm Liouville problem with potential V and periodic boundary conditions, therefore the dynamical system (6.3) does not coincide with the system obtained by considering (0.1) with periodic boundary conditions on $[-\pi,\pi]$.*

Remark 6.3. *A remarkable exception is the Klein Gordon equation (corresponding to $V = 0$) where the functions ψ_j are just the imaginary exponentials which are also eigenfunctions of the linerized problem with periodic boundary conditions. It follows that in this case the Hamiltonian (6.3) is the restriction of the Hamiltonian of the nonlinear Klein–Gordon equation with periodic boundary conditions in $[-\pi,\pi]$ to the space of the functions with zero average. To obtain the true nonlinear Klein Gordon equation one has to add the degree of freedom corresponding to the average u. This can be easily done still fitting the framework of the previous section.*

Define the subspace of skewsymmetric functions by

$$\mathcal{A} := \{(\hat{p}_k, \hat{q}_k) \ : \ \hat{q}_k = \hat{q}_{-k} \text{ and } \hat{p}_k = \hat{p}_{-k}\}$$

then we have the following

Proposition 6.4. *The space $\mathcal{A}$ is invariant for the dynamics of the system (6.3); The dynamics of (6.3) in $\mathcal{A}$ coincides with the dynamics of the nonlinear wave equation (0.1).*

Proof. Consider first the vector field of h_0. It is clear that it leaves invariant $\mathcal{A}$ and that here its dynamics coincides with the dynamics of the linear part of (0.1) by identifying $\hat{q}_k$ $(k \geq 1)$ with u_k and $\hat{p}_k$ $(k \geq 1)$ with v_k.

Consider the nonlinear part and remark first that, since

$$\int_{-\pi}^{\pi} \psi_j(x)\psi_k(x)dx = \delta_{j,-k} \ , \tag{6.4}$$

the symplectic form can be written in terms of the variables $\hat{u}, \hat{v}$ as

$$\int_{-\pi}^{\pi} d\hat{v}(x) \wedge d\hat{u}(x) dx \ . \tag{6.5}$$

Expand the nonlinearity f of (6.3) in power series in u, namely write $f = \sum_{n\geq 3} f^n$, with

$$f^n(\hat{u}) = \int_{-\pi}^{\pi} G^n(x)[\hat{u}(x)]^n dx \ , \tag{6.6}$$

and consider the Hamiltonian vector field of f^n. By (6.5) it is given by

$$\frac{d}{dt}\hat{v} = nG^n \hat{u}^{n-1} \ , \quad \frac{d}{dt}\hat{u} = 0 \ ; \tag{6.7}$$

by the skew symmetry property of g it leaves invariant $\mathcal{A}$. On this manifold the vector field (6.7)coincides with the one of the single terms of the expansion of the nonlinearity of (0.1) by identifying (as natural) u and v with the restriction of $\hat{u}$ and $\hat{v}$ to $[0, \pi]$. $\square$

Introduce now complex variables ξ, η by

$$\xi_j := \frac{1}{\sqrt{2}}\left(\hat{q}_j\sqrt{\omega_j} - i\frac{\hat{p}_j}{\sqrt{\omega_j}}\right) \ , \quad \eta_j := \frac{1}{\sqrt{2}}\left(\hat{q}_j\sqrt{\omega_j} + i\frac{\hat{p}_j}{\sqrt{\omega_j}}\right) \ , \tag{6.8}$$

then clearly the quadratic part of the hamiltonian takes the form (4.2).

Fix a potential V as in sect. 2, then, concerning the frequencies, we have the following theorem that will be proved in the next subsection.

Theorem 6.5. *Having fixed m_1 and m_2 as above, there exists a subset $\mathcal{J}$ of $[m_1, m_2]$ with measure $m_2 - m_1$ such that for any $m \in \mathcal{J}$ the frequencies $\omega_j := \sqrt{\mu_j + m}$ $(j \geq 1)$ are strongly non vanishing. The index $\alpha = \alpha(r)$ can be chosen equal to $16r^5$. The same is true if one adds the frequencies $\omega_{-j} := \omega_j$, $j \geq 1$.*

Concerning the nonlinearity we have the following

Proposition 6.6. *Under the above assumptions there exist positive μ and R such that the nonlinearity of (6.3) is of class $\mathcal{M}^{\mu}_{R}$.*

Proof. Define first the functions

$$\bar{\xi}(x) := \sum_k \frac{\xi_k}{\sqrt{\omega_k}}\psi_k(x) \ , \quad \bar{\eta}(x) := \sum_k \frac{\eta_k}{\sqrt{\omega_k}}\psi_k(x) \ ,$$

and, to start with consider a homogeneous polynomial

$$f(\xi,\eta) := \int_{-\pi}^{\pi} G(x)[\bar{\xi}(x)]^{n_1}[\bar{\eta}(x)]^{n_2}dx \ , \tag{6.9}$$

with G periodic and analytic in a strip of width κ around the real axis; the expression (6.9) coincides with

$$\sum_{l,i} \frac{\xi_{l_1}...\xi_{l_{n_1}}\eta_{i_1}...\eta_{i_{n_2}}}{\sqrt{\omega_{l_1}...\omega_{l_{n_1}}\omega_{i_1}...\omega_{i_{n_2}}}}\int_{-\pi}^{\pi} G(x)\psi_{l_1}(x)...\psi_{l_{n_1}}(x)\psi_{i_1}(x)...\psi_{i_{n_2}}(x)dx \ , \tag{6.10}$$

we estimate the integral. Denote $k=(k_1...k_n):=(l,i)=(l_1,...,l_{n_1},i_1,...,i_{n_2})$, then such an integral is given by

$$\left|\sum_j \psi^{k_1}_{j_1}...\psi^{k_n}_{j_n}\int_{-\pi}^{\pi}G(x)\mathrm{e}^{i(j_1+...+j_n)x}dx\right| \leq \sum_j C^n\mathrm{e}^{-\sigma(|k_1-j_1|+...+|k_n-j_n|)}|G|_\kappa \mathrm{e}^{-\kappa|j_1+...+j_n|} \ ,$$

where $|G|_\kappa$ is the supremum of G in the strip $|\mathrm{Im}\,x|<\kappa$; here we used the estimate (6.1) of ψ^k_j, the fact that the last integral is the Fourier coefficient of G and the standard decay estimate of Fourier coefficients of analytic functions. To fix ideas assume that $\kappa>\sigma$, then, denoting by $f_k\equiv f_{k_1,...,k_n}$ the integral in (6.10), and choosing a positive $\delta<\sigma$ we have

$$\begin{aligned}\frac{|f_k|}{|G|_\kappa} &\leq \sum_j C^n\mathrm{e}^{-\sigma(|k_1-j_1|+...+|k_n-j_n|)}\mathrm{e}^{-\kappa|j_1+...+j_n|}\\ &\leq \sum_j C^n\mathrm{e}^{-\delta[|k_1-j_1|+...+|k_n-j_n|]}\mathrm{e}^{-(\sigma-\delta)[|k_1-j_1|+...+|k_n-j_n|]}\mathrm{e}^{-(\sigma-\delta)|j_1+...+j_n|}\\ &\leq C^n\mathrm{e}^{-(\sigma-\delta)|k_1+...+k_n|}\sum_j \mathrm{e}^{-\delta|k_1-j_1|}\mathrm{e}^{-\delta|k_2-j_2|}...\mathrm{e}^{-\delta|k_n-j_n|} \leq C^n\mathrm{e}^{-(\sigma-\delta)|k_1+...k_n|}\end{aligned}$$

where we redefined the constant C and used the triangular inequality:

$$|k_1+...+k_n| \leq |k_1-j_1+...+k_n-j_n|+|j_1+...+j_n| \leq |k_1-j_1|+...+|k_n-j_n|+|j_1+...+j_n| \ .$$

Take now $\mu=\sigma-2\delta$, then one has

$$|f|_\mu \leq |G|_\kappa\sum_l \mathrm{e}^{\mu|l|}c^n \sup_{k_1+...+k_n=l}\frac{\mathrm{e}^{-(\sigma-\delta)|k_1+...+k_n|}}{\sqrt{\omega_{k_1}...\omega_{k_n}}} = c^n|G|_\kappa\sum_l \mathrm{e}^{-\delta|l|}$$

where we used the fact that ω_k is strictly positive for all k's. So, one has that, for $\mu < \min\{\sigma, \kappa\}$, (6.9) has finite μ–modulus smaller than $C|G|_\kappa c^n$ where C, c are positive constants that depend on G only through κ. To obtain the statement just remark that by analyticity there exist constants such that $|G^n|_\kappa \leq Cc^n$, and that G^n is the sum of at most 2^n polynomials of the form (6.9). □

We have thus verifyed that the system (6.3) fulfills the assumptions of theorem 4.1 that therefore applies. Theorem 2.2 immediately follows as a corollary in virtue of proposition 6.4 and of the remark that, since on $\mathcal{A}$ $\xi_k = \xi_{-k}$ and $\eta_k = \eta_{-k}$, the restriction of the normal form to this manifold depends only on the actions I_k.

Proof of Corollary 2.3. The estimates (2.7) can be obtained just estimating the time derivatives of the functions $\|\zeta\|^2_{\mathcal{F}_s}$ and $j^{2s}I_j(\zeta)$, which are analytic. The estimate (2.8) is obtained by remarking that the distance $d_{s'}(\zeta, \mathbf{T}_0)$ is estimated by

$$[d_{s'}(\zeta, \mathbf{T}_0)]^2 \leq \left(\sum_{j\geq 1} j^{-2(s-s')}\right)^{1/2} \left[\sum_{j\geq 1} j^{4s}(I_j(\zeta) - I_j(\bar{u}, \bar{v}))^2\right]^{1/2}$$

and by estimating the time derivative of the function in square brackets which is analytic in $\mathcal{F}_s$. □

A further corollary concerns the case of the nonlinear Klein Gordon equation with periodic boundary conditions. To state it denote by $J_k := I_k + I_{-k}$ the sum of actions corresponding to modes with the same frequency.

Corollary 6.7. *Consider equation (0.1) with periodic boundary conditions on $[-\pi, \pi]$ in the particular case $V = 0$, and fix m in such a way that the frequencies are strongly nonvanishing. Fix $M \geq 4$ and $s \geq s_*$, then there exist positive constants R_s, C_s such that the following holds true: let $\zeta(t) \equiv (u(t), v(t))$ be the solution of the Cauchy problem for equation (0.1) with initial datum $(u_0, v_0) \in H^s(\mathbb{R}/2\pi\mathbf{Z}) \times H^{s-1}(\mathbb{R}/2\pi\mathbf{Z})$ if*

$$R := \|(u, v)\|_{H^s \times H^{s-1}} \leq R_s$$

then for all times t with

$$|t| \leq \frac{1}{C_s R^{M-3}} \tag{6.11}$$

the following estimates hold

$$\|\zeta(t)\|_{H^s \times H^{s-1}} \leq 2C_s R \ , \quad \frac{|J_j(t) - J_j(0)|}{R^2} \leq \frac{C_s}{j^{2s}} R \tag{6.12}$$

6.2 Measure estimates

Since $\omega_{-j} = \omega_j$ we will consider only the case of positive j's. We first introduce some notations.

Given integers $0 < i_1, ..., i_r \leq N$, and a choice σ of r signs, define the vector $\mathbf{k} := \sum_{l=1}^{r} \pm e_{i_l}$ where $e_{i_l} \in \mathbf{Z}^{\infty}$ is the vector with al components equal to zero but the i_l–th which is equal to 1, so that

$$\mathbf{k} \cdot \omega = \pm \omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_r}$$

We will consider only choices of σ *such that* $\mathbf{k} \neq 0$. Then, for fixed τ, γ_1 we define

$$\mathcal{R}_i^{\sigma}(\gamma_1, \tau) := \left\{ m \in \mathcal{I} \ : \ |\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_r}| < \frac{\gamma_1}{i_1^{\tau} i_2^{\tau} ... i_r^{\tau}} \right\} \tag{6.13}$$

Given $j > N$ define

$$\mathcal{R}_{ij}^{\sigma}(\gamma_1, \tau) := \left\{ m \in \mathcal{I} \ : \ \left|\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_{r-1}} \pm \omega_j\right| < \frac{\gamma_1}{i_1^{\tau} i_2^{\tau} ... i_{r-1}^{\tau}} \right\}$$

and finally, given also $k \geq j$ define

$$\mathcal{R}_{ijk}^{\sigma}(\gamma_1, \tau) := \left\{ m \in \mathcal{I} \ : \ \left|\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_{r-2}} \pm \omega_j \pm \omega_k\right| < \frac{\gamma_1}{i_1^{\tau} i_2^{\tau} ... i_{r-2}^{\tau}} \right\}$$

we assume that when $j = k$ the sign in front of ω_j agrees with the sign in front of ω_k. Obviously $\mathcal{R}_{ij}^{\sigma}$ and $\mathcal{R}_{ijk}^{\sigma}$ do not depend on i_r and on i_{r-1}, i_r respectively; this is not important for the following.

We begin by proving the following estimate

Lemma 6.8. *For any $K \leq r$, let $i_1 < ... < i_K$ be K different positive indeces; then one has*

$$\begin{vmatrix} \omega_{i_1} & \omega_{i_2} & \cdots & \omega_{i_K} \\ \frac{d\omega_{i_1}}{dm} & \frac{d\omega_{i_2}}{dm} & \cdots & \frac{d\omega_{i_K}}{dm} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{d^{K-1}\omega_{i_1}}{dm^{K-1}} & \frac{d^{K-1}\omega_{i_2}}{dm^{K-1}} & \cdots & \frac{d^{K-1}\omega_{i_K}}{dm^{K-1}} \end{vmatrix} \geq \frac{C}{i_1^{2K-1} i_2^{2K-1} ... i_K^{2K-1}} \ . \tag{6.14}$$

Proof. First remark that, by explicit computation one has

$$\frac{d^n \omega_i}{dm^n} = \frac{(2n-3)!}{2^{n-2}(n-2)!2^n} \frac{(-1)^{n+1}}{(\mu_i + m)^{n-\frac{1}{2}}} \ . \tag{6.15}$$

Substituting (6.15) in the l.h.s. of (6.14) we get the determinant to be estimated. To obtain the estimate factorize from the $i-th$ column the term $(\mu_i + m)^{1/2}$, and from the $j-th$ row the term $\frac{(2j-3)!}{2^{j-2}(j-2)!2^j}$. Forgetting the inessential powers of -1, we obtain that the determinant to be estimated is given by

$$\left[\prod_{l=1}^{K} \omega_{i_l}\right] \left[\prod_{n=1}^{K-1} \frac{(2n-3)!}{2^{n-2}(n-2)!2^n}\right] \times \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{i_1} & x_{i_2} & x_{i_3} & \cdots & x_{i_K} \\ x_{i_1}^2 & x_{i_2}^2 & x_{i_3}^2 & \cdots & x_{i_K}^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{i_1}^{K-1} & x_{i_2}^{K-1} & x_{i_3}^{K-1} & \cdots & x_{i_K}^{K-1} \end{vmatrix} \tag{6.16}$$

where we denoted by $x_i := (\mu_i + m)^{-1}$. The last determinant is a Vandermond determinant whose value is given by

$$\prod_{l<n\leq K} (x_{i_l} - x_{i_n}) \qquad (6.17)\ .$$

Now one has

$$|x_{i_l} - x_{i_n}| = \left| \frac{1}{\mu_{i_l}+m} - \frac{1}{\mu_{i_n}+m} \right| = \frac{|\mu_{i_n} - \mu_{i_l}|}{(\mu_{i_l}+m)(\mu_{i_n}+m)} \geq C x_{i_l} x_{i_n}$$

with a suitable C. So, (6.17) is estimated by

$$\prod_{l=2}^{K}\prod_{n=1}^{l-1} C x_{i_l} x_{i_n} = C^{\sum_{l=2}^{K}(l-1)} \prod_{l=2}^{K}\left(x_{i_l}^{l-1} \prod_{n=1}^{l-1} x_{i_n} \right) = C \prod_{l=1}^{K} x_{i_l}^{K}\ ,$$

from which, using the asymptotics of the frequencies, the thesis immediately follows. □

The above result will be used in view of the following

Lemma 6.9. *(Proposition of Appendix B in [4]) Let $u^{(1)}, ..., u^{(K)}$ be K independent vectors with $\left\|u^{(i)}\right\|_{\ell^1} \leq 1$. Let $w \in \mathbb{R}^K$ be an arbitrary vector, then there exist $i \in [1, ..., K]$, such that*

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

where $\det(u^{(i)})$ is the determinant of the matrix formed by the components of the vectors $u^{(i)}$.

For the proof see [4].

Corollary 6.10. *Let $w \in \mathbb{R}^\infty$ be a vector with only K components different from zero, namely those with index $i_1, ..., i_K$; assume $K \leq r$. Then, for any $m \in \mathcal{I}$ there exists, an index $i \in [0, ..., K-1]$ such that*

$$\left| w \cdot \frac{d^i\omega}{dm^i}(m) \right| \geq C \frac{\|w\|_{\ell^1}}{\prod_{l=1}^{K} i_l^{\beta}}$$

where ω is the frequency vector, and $\beta = 2r$.

Proof. Consider the vectors

$$u^{(i)} := \begin{cases} \frac{1}{\left\|\frac{d^i\omega}{dm^i}\right\|}\frac{d^i\omega}{dm^i} & \text{if} \quad \left\|\frac{d^i\omega}{dm^i}\right\| > 1 \\ \frac{d^i\omega}{dm^i} & \text{if} \quad \left\|\frac{d^i\omega}{dm^i}\right\| \leq 1 \end{cases}$$

and apply lemma 6.9. We thus get that there exists i such that

$$\left| w \cdot \frac{d^i\omega}{dm^i} \right| \geq C \frac{\|w\|_{\ell^1}}{\prod_{l=0}^{K} \max\left\{1, \left\|\frac{d^l\omega}{dm^l}\right\|_{\ell^1}\right\}} \prod_{l=1}^{K} \frac{1}{i_l^{\tilde\beta}}$$

with $\tilde{\beta} = 2r - 1$ and ω denotes here the vector with all components equal to zero except those with indexes $i_1, ..., i_K$ which are put equal to $\omega_{i_1} ... \omega_{i_K}$. Then remark that the ℓ^1 norm of $\frac{d^i \omega}{dm^i}$ is bounded by a constant (independent of w) for all $i \geq 2$, while for $i = 1$ we have that it is given by the sum of at most r components different from zero. Remark also that each component of $\frac{d\omega}{dm}$ is bounded uniformly with respect to the choice of the indexes, and therefore also for $i = 1$ the above norm is bounded by a constant independent of w. Thus we have

$$\left| w \cdot \frac{d^i \omega}{dm^i} \right| \geq C \frac{\|w\|_{\ell^1}}{\|\omega\|_{\ell^1}} \prod_{l=1}^{K} \frac{1}{i_l^{\tilde{\beta}}}$$

Now, as one can easily check, $\|\omega\|$ is bounded by

$$C \prod_{l=1}^{K} \omega_{i_l}$$

from which, using the asymptotic of the frequencies we get the thesis. $\square$

Lemma 6.11. *(Lemma 8.4 of [3]) Let $g : \mathcal{I} \to \mathbb{R}$ be r times differentiable, and assume that*

(1) $\forall m \in \mathcal{I}$ there exists $i \leq r - 1$ such that $g^{(i)}(m) > d$

(2) there exists A such that $|g^{(i)}(m)| \leq A$ $\forall m \in \mathcal{I}$, and $\forall i$ with $1 \leq i \leq r$.

Define

$$\mathcal{I}_h := \{ m \in \mathcal{I} \ : \ |g(m)| \leq h \} \ ,$$

then

$$\frac{|\mathcal{I}_h|}{|\mathcal{I}|} \leq \frac{A}{d} 2(2 + 3 + ... + r + 2d^{-1}) h^{1/r}$$

For the proof see [3] and [20].

By combining lemma 6.11 and lemma 6.10 we get the following

Lemma 6.12. *For any choice of $i_1 \leq i_2 \leq ... \leq i_r$ and any acceptable choice σ of the signs, one has*

$$|\mathcal{R}_i^{\sigma}(\gamma_1, \tau)| \leq |\mathcal{I}| C \frac{{\gamma_1}^{1/r}}{\prod_{l=1}^{r} i_l^{\delta}} \tag{6.18}$$

with $\delta = -2\beta + \tau / r$. Moreover, for any choice also of $j > N$ one has

$$\left| \mathcal{R}_{ij}^{\sigma}(\gamma_1, \tau) \right| \leq |\mathcal{I}| C \frac{{\gamma_1}^{1/r} j^{\beta}}{\prod_{l=1}^{r-1} i_l^{\delta}} \tag{6.19}$$

and, for any choice of $k \geq j$ one has

$$\left| \mathcal{R}_{ijk}^{\sigma}(\gamma_1, \tau) \right| \leq |\mathcal{I}| C \frac{{\gamma_1}^{1/r} j^{\beta} k^{\beta}}{\prod_{l=1}^{r-2} i_l^{\delta}} \tag{6.20}$$

Lemma 6.13. *Assume $i_1 \leq i_2 ... \leq i_{r-1} \leq N$, then there exists C_1 such that*

$$\mathcal{R}^{\sigma}_{ij} = \emptyset$$

when $j \geq C_1 N$

Proof. Defining $\bar{\mathbf{k}}$ as $\sum_{l=1}^{r-1} \pm e_{i_l}$ we have $|\bar{\mathbf{k}}| \leq r-1$, and therefore, for large enough j we have,

$$\left|\bar{\mathbf{k}} \cdot \omega \pm \omega_j\right| \geq \omega_j - (r-1)\omega_N$$

which is automatically larger than $(r-1)\omega_N$ and also than $\gamma_1 / i_1^{\tau} ... i_{r-1}^{\tau}$ if

$$\omega_j \geq 2(r-1)\omega_N$$

which is implied by the hypothesis, with a suitable choice of C_1. $\square$

Lemma 6.14. *Assume $i_1 \leq i_2 ... \leq i_{r-1} \leq N$, then one has*

$$\left|\bigcup_{j \geq N} \mathcal{R}^{\sigma}_{ij}(\tau, \gamma_1)\right| \leq \frac{C|\mathcal{I}|{\gamma_1}^{1/r} N^{\beta+1}}{\prod_{l=1}^{r-1} i_l^{\delta}} \ . \tag{6.21}$$

Proof. We have

$$\left|\bigcup_{j > N} \mathcal{R}^{\sigma}_{ij}(\tau, \gamma_1)\right| = \left|\bigcup_{C_1 N > j \geq N} \mathcal{R}^{\sigma}_{ij}(\tau, \gamma_1)\right| \leq \sum_{j=N+1}^{C_1 N} \left|\mathcal{R}^{\sigma}_{ij}\right|$$

and using lemma 6.12 this gives (6.21). $\square$

The estimate of the union over j, k of the measure of the sets $\mathcal{R}^{\sigma}_{ijk}$ is more difficult. It follows the proof of lemma 8 of [18] and lemma 3.1 of [12].

Lemma 6.15. *Assume $i_1 \leq i_2 ... \leq i_{r-2} \leq N$. One has*

$$\left|\bigcup_{\substack{j \geq N \\ k > j}} \mathcal{R}^{\sigma}_{ijk}(\tau, \gamma_1)\right| \leq C \frac{{\gamma_1}^{\theta/r^2} N^{\frac{\beta\theta}{r}+1}}{\prod_{l=1}^{r-2} i_l^{\frac{\tau\theta}{r^2} - 2\beta}}$$

with

$$\theta := \frac{1}{2\beta + 1 + 1/r} \ .$$

Proof. Clearly the difficulties arise when the sign in front of ω_k is opposite to the sign in front of ω_j. So, we denote

$$\tilde{\mathcal{R}}_{i,j,k} := \left\{ m \in \mathcal{I} \ : \ \left|\omega_{i_1} \pm \omega_{i_2} \pm ... \pm \omega_{i_{r-2}} \pm (\omega_j - \omega_k)\right| < \frac{\gamma_1}{i_1^{\tau} i_2^{\tau} ... i_{r-2}^{\tau}} \right\}$$

Denote $n = k - j$, one has

$$\omega_k - \omega_j = n + r_{jk}$$

with

$$|r_{jk}| \leq \frac{c}{j} \ .$$

Now, one can easily verify that, defining

$$Q_{ij}^n := \left\{ m \in \mathcal{I} \ : \ |\mathbf{k} \cdot \omega + n| \leq \frac{\gamma_1}{i_1^\tau i_2^\tau ... i_{r-2}^\tau} + \frac{c}{j} \right\}$$

one has $\tilde{\mathcal{R}}_{i,j,j+n} \subset Q_{ij}^n$ (for details see [3]), remarking also that, for $j \geq j_0$ one has $Q_{ij}^n \subset Q_{ij_0}^n$, we obtain

$$\left| \bigcup_j \tilde{\mathcal{R}}_{i,j,j+n} \right| \leq \left| Q_{ij_0}^n \cup \left(\bigcup_{j<j_0} \mathcal{R}_{i,j,j+n} \right) \right|$$

$$\leq C \left(\frac{\gamma_1}{i_1^\tau ... i_{r-2}^\tau} + \frac{c}{j_0} \right)^{1/r} \prod_{l=1}^{r-2} i_l^{2\beta} + \sum_{j<j_0} \frac{{\gamma_1}^{1/r} j^\beta (j+n)^\beta C}{\prod_{l=1}^{r-2} i_l^\delta} \tag{6.22}$$

$$\leq C \left(\frac{\gamma_1}{i_1^\tau ... i_{r-2}^\tau} + \frac{c}{j_0} \right)^{1/r} \prod_{l=1}^{r-2} i_l^{2\beta} + \frac{N^\beta {\gamma_1}^{1/r} j_0^{2\beta+1} C}{\prod_{l=1}^{r-2} i_l^\delta} \ ,$$

where we used the fact that we always have $n < CN$ otherwise, by the argument of lemma 6.13, the set $\tilde{\mathcal{R}}_{i,j,j+n}$ is empty, and lemma 6.11 with $g = \bar{\mathbf{k}} \cdot \omega + n$ and $\bar{\mathbf{k}} := \sum_{l=1}^{r-2} \pm e_j$. We now choose j_0 so that both terms in the previous formula have the same order of magnitude. Thus we put

$$j_0 = \left(\frac{\prod_{l=1}^{r-2} i_l^{\delta+2\beta}}{N^\beta {\gamma_1}^{1/r}} \right)^\theta$$

inserting in (6.22) and summing over $n \leq CN$ we get the thesis. □

Theorem 6.16. *Provided $\delta > 1$ there exists C such that*

$$\left| \bigcup_{i_1=1}^\infty \bigcup_{i_2=1}^\infty ... \bigcup_{i_r=1}^\infty \bigcup_\sigma \mathcal{R}_i^\sigma \right| \leq C {\gamma_1}^{1/r} |\mathcal{I}|$$

(see (6.13) for the definition of $\mathcal{R}_i^\sigma$).

Proof. By (6.18) one has

$$\left| \bigcup_{i_1=0}^\infty \bigcup_{i_2=1}^\infty ... \bigcup_{i_r=1}^\infty \bigcup_\sigma \mathcal{R}_i^\sigma \right| \leq \sum_{i_1,...,i_r,\sigma} |\mathcal{R}_i^\sigma| \leq C {\gamma_1}^{1/r} |\mathcal{I}| 2^r \left(\sum_{i=0}^\infty \frac{1}{i^\delta} \right)^r \ ,$$

from which, remarking that the series is convergent one has the thesis. □

Remark 6.17. *Here we did not assume that the indexes are smaller than N; on the contrary we considered them as varying from zero to infinity thus this theorem ensures that the condition (3.1) is fulfilled for almost all masses.*

Remark 6.18. *The condition $\delta > 1$ explicitly reads $\tau > 2r^2 + r$.*

Analogously one has

Theorem 6.19. *Assume*

$$\frac{\tau\theta}{r^2} - 2\beta > 1 \ , \tag{6.23}$$

then

$$\left| \bigcup_{i_1=0}^{N} \bigcup_{i_2=0}^{N} \dots \bigcup_{i_{r-1}=0}^{N} \bigcup_{\sigma} \bigcup_{j\geq N+1} \mathcal{R}_{ij}^{\sigma} \right| \leq C|\mathcal{I}|{\gamma_1}^{1/r} N^{\beta+1}$$

and moreover

$$\left| \bigcup_{i_1=0}^{N} \bigcup_{i_2=0}^{N} \dots \bigcup_{i_{r-1}=0}^{N} \bigcup_{\sigma} \bigcup_{j\geq N+1} \bigcup_{k\geq N+1} \mathcal{R}_{ijk}^{\sigma} \right| \leq C|\mathcal{I}|{\gamma_1}^{\theta/r^2} N^{\frac{\beta\theta}{r}+1} \ .$$

Remark 6.20. *Eq. (6.23) is implied by $\tau > Cr^4$.*

Proof of theorem 6.5. Just take $\gamma_1 = \gamma/N^{4r^3}$ and apply theorem 6.19. □

Remark 6.21. *Remark that in the definition of strongly nonvanishing frequencies appear the quantities in the definition of the sets $\mathcal{R}_{ijk}^{\sigma}$ with $r+2$ instead of r. Actualy the quantities entering in the perturbative construction are those we estimated in this section. If one wants to fit in the definition of strongly nonvanishing frequencies one should just substitute in the estimates of this section $r+2$ to r.*

Remark 6.22. *The theory of this subsection directly applies also to the case of frequencies of the form*

$$\omega_j = \sqrt{\mu_j + m}$$

with $\mu_j \sim j^d$ and $d > 1$.

References

[1] D. Bambusi, N.N. Nekhoroshev: *A property of exponential stability in nonlinear wave equation near the fundamental linear mode.* Physica D **122**, 73-104 (1998).

[2] D. Bambusi: *Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations.* Math. Z. **130**, 345–387, (1999).

20.03.02

[3] D. Bambusi: *On long time stability in Hamiltonian perturbations of nonresonant linear PDE's,* Nonlinearity **12**, 823-850, (1999).

[4] G. Benettin, L. Galgani, A. Giorgilli: *A Proof of Nekhoroshev's Theorem for the Stability Times in Nearly Integrable Hamiltonian Systems* Cel. Mech. **37**, 1-25, (1985).

[5] G. Benettin, J. Fröhlich, A. Giorgilli: *A Nekhoroshev-type Theorem for Hamiltonian Systems with Infinitely Many Degrees of Freedom.* Commun. Math. Phys. **119**, 95-108 (1988).

[6] J. Bourgain: *Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equation* GAFA **6** 201–230 (1995).

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

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

[9] J. Bourgain: Nonlinear Schrödinger equations. In "Hyperbolic equations and frequency interactions"; L. Caffarelli, E. Weinan editors. IAS/Park City mathematics series 5. American Mathematical Society (Providence, Rhode Island 1999).

[10] L. Chierchia, J. You: *KAM tori for 1D nonlinear wave equations with periodic boundary conditions.* Comm. Math. Phys. **211**, 497–525 (2000).

[11] W. Craig, C.E. Wayne: *Newton's Method and Periodic Solutions of Nonlinear Wave Equations* Comm. Pure and Appl. Math. **46** 1409-1501 (1993).

[12] S.B. Kuksin: Nearly Integrable Infinite-Dimensional Hamiltonian Systems, Lect. Notes Math. 1556 (Springer 1994)

[13] S.B. Kuksin, J. Pöschel: *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation* Ann. of Math. **142** 149–179 (1995).

[14] A. Morbidelli, A. Giorgilli: *Superexponential stability of KAM tori.* J. Statist. Phys. **78**, 1607–1617 (1995).

[15] N.N. Nekhoroshev: *Behaviour of Hamiltonian systems close to integrable.* Funct. Anal. and Appl. **5** 338–339 (1971).

[16] N.N. Nekhoroshev: *Exponential estimate of the stability time of near integrable Hamiltonian systems.* Russ. Math. Surveys. **32** (6) 1-65 (1977).

[17] L. Niederman: *Nonlinear Stability around an Elliptic Equilibrium Point in a Hamiltonian System* Nonlinearity **11**, 1465–1479, (1998).

[18] J. Pöschel: *A KAM–Theorem for some Partial Differential Equations* Ann. Scuola Norm. Sup. Pisa Cl. Sci. **23**, 119–148 (1996).

[19] J. Pöschel: *Quasi–periodic solutions for a nonlinear wave equation.* Comment. Math. Helv. **71**, 269–296, (1996).

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