A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus.

Erwan Faou, Benoît Grébert

A Nekhoroshev type theorem for the nonlinear Schrödinger equation on the d-dimensional torus.

Erwan Faou [1] and Benoît Grébert [2]

[1] INRIA & ENS Cachan Bretagne,
Avenue Robert Schumann F-35170 Bruz, France.
email: Erwan.Faou@inria.fr

[2] Laboratoire de Mathématiques Jean Leray, Université de Nantes,
2, rue de la Houssinière F-44322 Nantes cedex 3, France.
email: benoit.grebert@univ-nantes.fr

March 25, 2010

Abstract

We prove a Nekhoroshev type theorem for the nonlinear Schrödinger equation

$$iu_t = -\Delta u + V \star u + \partial_{\bar{u}} g(u, \bar{u}) \, , \quad x \in \mathbb{T}^d,$$

where V is a typical smooth potential and g is analytic in both variables. More precisely we prove that if the initial datum is analytic in a strip of width $\rho > 0$ with a bound on this strip equals to ε then, if ε is small enough, the solution of the nonlinear Schrödinger equation above remains analytic in a strip of width $\rho/2$ and bounded on this strip by $C\varepsilon$ during very long time of order $\varepsilon^{-\alpha|\ln \varepsilon|^\beta}$ for some constants $C > 0$, $\alpha > 0$ and $\beta < 1$.

MSC numbers: 37K55, 35B40, 35Q55.

Keywords: Nekhoroshev theorem. Nonlinear Schrödinger equation. Normal forms.

Contents

1 Introduction and statements

We consider the nonlinear Schrödinger equation

$$iu_t = -\Delta u + V \star u + \partial_{\bar{u}} g(u, \bar{u}) \, , \quad x \in \mathbb{T}^d, \quad t \in \mathbb{R}, \tag{1.1}$$

where V is a smooth convolution potential and g is an analytic function on a neighborhood of the origin in $\mathbb{C}^2$ which has a zero of order at least 3 at the origin and satisfies $g(z, \bar{z}) \in \mathbb{R}$. In more standard models, the convolution term is replaced by a multiplicative potential. The use of a convolution potential makes easier the analysis of the resonances.
For instance when $g(u, \bar{u}) = \frac{a}{p+1}|u|^{2p+2}$ with $a \in \mathbb{R}$ and $p \in \mathbb{N}$, we recover the standard NLS equation $iu_t = -\Delta u + V \star u + a|u|^{2p}u$. We notice that (1.1) is a Hamiltonian system associated with the Hamiltonian function

$$H(u, \bar{u}) = \int_{\mathbb{T}^d} \left(|\nabla u|^2 + (V \star u)\bar{u} + g(u, \bar{u}) \right) \mathrm{d}x.$$

and the symplectic structure inherent to the complex structure, $i\mathrm{d}u \wedge \mathrm{d}\bar{u}$.

This equation has been considered with Hamiltonian tools in two recent works. In the first one (see [BG03] and also [BG06] and [Bou96] for related results) Bambusi & Grébert prove a Birkhoff normal form theorem adapted to this equation and obtain dynamical consequences on the long time behavior of the solutions with small initial Cauchy data in Sobolev spaces. More precisely they prove that if the Sobolev norm of index s of the initial datum u_0 is sufficiently small (of order ε) then the Sobolev norm of the solution is bounded by 2ε during very long time (of order ε^{-r} with r arbitrary). In the second one (see [EK]) Eliasson & Kuksin obtain a KAM theorem adapted to this equation. In particular they prove that, in a neighborhood of $u = 0$, many of the invariant finite dimensional tori of the linear part of the equation are preserved by small Hamiltonian perturbations. In other words, (1.1) has many quasi-periodic solutions. In both cases non resonances conditions (not exactly the same) have to be imposed on the frequencies of the linear part and thus on the potential V.

Both results are related to the stability of the zero solution which is an elliptic equilibrium of the linear equation. The first establishes the stability for polynomials times with respect to the size of the (small) initial datum while the second proves the stability for all time of certain solutions. In the present work we extend the technic of normal form and we establish the stability for times of order $\varepsilon^{-\alpha|\ln \varepsilon|^\beta}$ for some constants $\alpha > 0$ and $\beta < 1$, ε being the size of the

initial datum in an analytic space.

We now state precisely our result. We assume that V belongs to the following space ($m > d/2$, $R > 0$)

$$\mathcal{W}_m = \{V(x) = \sum_{a\in\mathbb{Z}^d} v_a e^{ia\cdot x} \mid v'_a := v_a(1+|a|)^m/R \in [-1/2, 1/2] \text{ for any } a \in \mathbb{Z}^d\} \quad (1.2)$$

that we endow with the product probability measure. Here, for $a = (a_1, \dots, a_d) \in \mathbb{Z}^d$, $|a|^2 = a_1^2 + \cdots + a_d^2$.

For $\rho > 0$, we denote by $\mathcal{A}_\rho \equiv \mathcal{A}_\rho(\mathbb{T}^d; \mathbb{C})$ the space of functions ϕ that are analytic on the complex neighborhood of d-dimensional torus $\mathbb{T}^d$ given by $I_\rho = \{x + iy \mid x \in \mathbb{T}^d,\ y \in \mathbb{R}^d \text{ and } |y| < \rho\}$ and continuous on the closure of this strip. We then denote by $|\cdot|_\rho$ the usual norm on $\mathcal{A}_\rho$

$$|\phi|_\rho = \sup_{z\in I_\rho} |\phi(z)|.$$

We note that $(\mathcal{A}_\rho, |\cdot|_\rho)$ is a Banach space.
Our main result is a Nekhoroshev type theorem:

Theorem 1.1 *There exists a subset $\mathcal{V} \subset \mathcal{W}_m$ of full measure, such that for $V \in \mathcal{V}$, $\beta < 1$ and $\rho > 0$, the following holds: there exist $C > 0$ and $\varepsilon_0 > 0$ such that if*

$$u_0 \in \mathcal{A}_{2\rho} \quad \text{and} \quad |u_0|_{2\rho} = \varepsilon \le \varepsilon_0$$

then the solution of (1.1) with initial datum u_0 exists for times $|t| \le \varepsilon^{-\alpha(\ln\varepsilon)^\beta}$ and satisfies

$$|u(t)|_{\rho/2} \le C\varepsilon \quad \text{for} \quad |t| \le \varepsilon^{-\sigma_\rho |\ln\varepsilon|^\beta}, \quad (1.3)$$

with $\sigma_\rho = \min\{\frac{1}{8}, \frac{\rho}{2}\}$.
Furthermore, writing $u(t) = \sum_{k\in\mathbb{Z}^d} \xi_k(t) e^{ik\cdot x}$, we have

$$\sum_{k\in\mathbb{Z}^d} e^{\rho|k|} \big||\xi_k(t)| - |\xi_k(0)|\big| \le \varepsilon^{3/2} \quad \text{for} \quad |t| \le \varepsilon^{-\sigma_\rho |\ln\varepsilon|^\beta}. \quad (1.4)$$

Estimate (1.4) asserts that there is almost no variation of the actions[1] and in particular no possibility of weak turbulence, i.e. exchanges between low Fourier modes and high Fourier modes. This kind of turbulence may induce the growth of the Sobolev norm $\sum(1+|k|^s)^2|\xi_k|^2$ ($s > 1$) of the solution as recently proved in [CKSTT09].
In finite dimension n, the standard Nekhoroshev result [Nek77] controls the dynamic over times of order $\exp\left(\frac{-\alpha}{\varepsilon^{1/(\tau+1)}}\right)$ for some $\alpha > 0$ and $\tau > n+1$ (see for instance [BGG85, GG85, Pös93]) which is of course much better than $\varepsilon^{-\alpha|\ln\varepsilon|^\beta} = e^{-\alpha|\ln\varepsilon|^{(1+\beta)}}$. Nevertheless this standard result does not extend to the infinite dimensional context. Actually, when $n \to \infty$, that $\varepsilon^{-1/(\tau+1)}$ can be transformed in $|\ln\varepsilon|^{(1+\beta)}$ is a good news!
The only previous work in the direction of Nekhoroshev estimates for PDEs was obtained by

[1] Here the actions are the modulus of the Fourier coefficients to the square, $I_k = |\xi_k|^2$.

Bambusi in [Bam99]. He also worked in spaces of analytic functions in a strip and for times of order $e^{-\alpha|\ln\varepsilon|^{1+\beta}}$, nevertheless the control of the solution was not obtained uniformly in a strip but in a complicated way involving the Fourier coefficients of the solution.

We now focus on the three main differences with the previous works on normal forms:

- we crucially use the zero momentum condition: in the Fourier space, the nonlinear term contains only monomials $z_{j_1}\cdots z_{j_k}$ with $j_1+\cdots+j_k=0$ (cf. Definition 2.4). This property allows to control the largest index by the others.
- we use ℓ^1-type norms to control the Fourier coefficients and the vector fields instead of ℓ^2-type norms as usual. Of course this choice does not allow to work in Hilbert spaces and makes obligatory a slight lost of regularity each time the estimates are transposed from the Fourier space to the initial space of analytic functions. But it turns out that this choice makes much more simpler the estimates on the vector fields (cf. Proposition 2.5 below and [FG10] for a similar framework in the context of numerical analysis).
- we notice that the vector field of a monomial,$z_{j_1}\cdots z_{j_k}$ containing at least three Fourier modes z_ℓ with large indices ℓ induces a flow whose dynamics is under control during very long time in the sense that the dynamic almost excludes exchanges between high Fourier modes and low Fourier modes (see Proposition 2.10). In [Bam03] or [BG06], such terms were neglected since the vector field of a monomial containing at least three Fourier modes with large indices is small in *Sobolev norm* (but not in analytic norm) and thus will almost keep invariant all the modes. This more subtle analysis for monomials was still used in [FGP10].

Finally we notice that our method could be generalized by considering not only zero momentum monomials but also monomials with finite or exponentially decreasing momentum. This would certainly allow to consider a nonlinear Schrödinger equation with a multiplicative potential V and nonlinearities depending periodically on x:

$$iu_t=-\Delta u+Vu+\partial_{\bar u}g(x,u,\bar u)\,,\quad x\in\mathbb{T}^d.$$

Nevertheless this generalization would generate a lot of technicalities and we prefer to focus here on the simplicity of the arguments.

2 Setting and Hypothesis

2.1 Hamiltonian formalism

The equation (1.1) is a semi linear PDE locally well posed in the Sobolev space $H^2(\mathbb{T}^d)$ (see for instance [Caz03]). Let u be a (local) solution of (1.1) and consider $(\xi,\eta)=(\xi_a,\eta_a)_{a\in\mathbb{Z}^d}$ the Fourier coefficients of u, $\bar u$ respectively, i.e.

$$u(x)=\sum_{a\in\mathbb{Z}^d}\xi_a e^{ia\cdot x}\quad\text{and}\quad\bar u(x)=\sum_{a\in\mathbb{Z}^d}\eta_a e^{-ia\cdot x}. \tag{2.1}$$

A standard calculus shows that u is solution in $H^2(\mathbb{T}^d)$ of (1.1) if and only if (ξ,η) is a solution in[2] $\ell^2_2\times\ell^2_2$ of the system

$$\left\{\begin{array}{lll}\dot\xi_a &= -i\omega_a\xi_a - i\frac{\partial P}{\partial\eta_a}, & a\in\mathbb{Z}^d,\\ \dot\eta_a &= i\omega_a\eta_a - i\frac{\partial P}{\partial\xi_a}, & a\in\mathbb{Z}^d,\end{array}\right. \tag{2.2}$$

where the linear frequencies are given by $\omega_a=|a|^2+v_a$ where as in (1.2), $V=\sum v_a e^{ia\cdot x}$, and the nonlinear part is given by

$$P(\xi,\eta)=\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} g(\sum\xi_a e^{ia\cdot x},\sum\eta_a e^{-ia\cdot x})\,\mathrm{d}x. \tag{2.3}$$

This system is reinterpreted in a Hamiltonian context endowing the set of couples $(\xi_a,\eta_a)\in\mathbb{C}^{\mathbb{Z}^d}\times\mathbb{C}^{\mathbb{Z}^d}$ with the symplectic structure

$$i\sum_{a\in\mathbb{Z}^d}\mathrm{d}\xi_a\wedge\mathrm{d}\eta_a. \tag{2.4}$$

We define the set $\mathcal{Z}=\mathbb{Z}^d\times\{\pm1\}$. For $j=(a,\delta)\in\mathcal{Z}$, we define $|j|=|a|$ and we denote by $\overline{j}$ the index $(a,-\delta)$.
We identify a couple $(\xi,\eta)\in\mathbb{C}^{\mathbb{Z}^d}\times\mathbb{C}^{\mathbb{Z}^d}$ with $(z_j)_{j\in\mathcal{Z}}\in\mathbb{C}^{\mathcal{Z}}$ via the formula

$$j=(a,\delta)\in\mathcal{Z}\Longrightarrow\left\{\begin{array}{lllll} z_j &=& \xi_a & \text{if} & \delta=1,\\ z_j &=& \eta_a & \text{if} & \delta=-1.\end{array}\right. \tag{2.5}$$

By a slight abuse of notation, we often write $z=(\xi,\eta)$ to denote such an element.
For a given $\rho>0$, we consider the Banach space $\mathcal{L}_\rho$ made of elements $z\in\mathbb{C}^{\mathcal{Z}}$ such that

$$\|z\|_\rho:=\sum_{j\in\mathcal{Z}}e^{\rho|j|}|z_j|<\infty,$$

and equipped with the symplectic form (2.4). We say that $z\in\mathcal{L}_\rho$ is *real* when $z_{\overline{j}}=\overline{z_j}$ for any $j\in\mathcal{Z}$. In this case, we write $z=(\xi,\bar\xi)$ for some $\xi\in\mathbb{C}^{\mathbb{Z}^d}$. In this situation, we can associate with z the function u defined by (2.1).
The next lemma shows the relation with the space $\mathcal{A}_\rho$ defined above:

Lemma 2.1 *Let u be a complex valued function analytic on a neighborhood of $\mathbb{T}^d$, and let $(z_j)_{j\in\mathcal{Z}}$ be the sequence of its Fourier coefficients defined by (2.1) and (2.5). Then for all $\mu<\rho$, we have*

$$\textit{if}\quad u\in\mathcal{A}_\rho\quad\textit{then}\quad z\in\mathcal{L}_\mu\quad\textit{and}\quad\|z\|_\mu\le c_{\rho,\mu}|u|_\rho\,; \tag{2.6}$$

$$\textit{if}\quad z\in\mathcal{L}_\rho\quad\textit{then}\quad u\in\mathcal{A}_\mu\quad\textit{and}\quad|u|_\mu\le c_{\rho,\mu}\|z\|_\rho\,, \tag{2.7}$$

where $c_{\rho,\mu}$ is a constant depending on ρ and μ and the dimension d.

[2] As usual, $\ell^2_2=\{(\xi_a)_{a\in\mathbb{Z}^d}\mid\sum(1+|a|^2)|\xi_a|^2<+\infty\}$.

Proof. Assume that $u \in \mathcal{A}_\rho$. Then by Cauchy formula, we have for all $j \in \mathcal{Z}$, $|z_j| \leq |u|_\rho e^{-\rho|j|}$. Hence for $\mu < \rho$, we have

$$\|z\|_\mu \leq |\phi|_\rho \sum_{j\in\mathcal{Z}} e^{(\mu-\rho)|j|} \leq |\phi|_\rho \Big(2\sum_{n\in\mathbb{Z}} e^{\frac{(\mu-\rho)}{\sqrt{d}}|n|}\Big)^d \leq \left(\frac{2}{1-e^{\frac{(\mu-\rho)}{\sqrt{d}}}}\right)^d |u|_\rho.$$

Conversely, assume that $z \in \mathcal{L}_\rho$. Then $|\xi_a| \leq \|z\|_\mu e^{-\rho|a|}$ for all $a \in \mathbb{Z}^d$, and thus by (2.1), we get for all $x \in \mathbb{T}^d$ and $y \in \mathbb{R}^d$ with $|y| \leq \mu$,

$$|u(x+iy)| \leq \sum_{a\in\mathbb{Z}^d} |\xi_a| e^{|ay|} \leq \|z\|_\rho \sum_{a\in\mathbb{Z}^d} e^{-(\rho-\mu)|a|} \leq \left(\frac{2}{1-e^{\frac{(\mu-\rho)}{\sqrt{d}}}}\right)^d \|z\|_\rho .$$

Hence u is bounded on the strip I_μ. ∎

For a function F of $\mathcal{C}^1(\mathcal{L}_\rho, \mathbb{C})$, we define its Hamiltonian vector field by $X_F = J\nabla F$ where J is the symplectic operator on $\mathcal{L}_\rho$ induced by the symplectic form (2.4), $\nabla F(z) = \left(\frac{\partial F}{\partial z_j}\right)_{j\in\mathcal{Z}}$ and where by definition we set for $j = (a,\delta) \in \mathbb{Z}^d \times \{\pm 1\}$,

$$\frac{\partial F}{\partial z_j} = \begin{cases} \dfrac{\partial F}{\partial \xi_a} & \text{if} \quad \delta = 1, \\ \dfrac{\partial F}{\partial \eta_a} & \text{if} \quad \delta = -1. \end{cases}$$

For two functions F and G, the Poisson Bracket is (formally) defined as

$$\{F,G\} = \nabla F^T J \nabla G = i \sum_{a\in\mathbb{Z}^d} \frac{\partial F}{\partial \eta_a}\frac{\partial G}{\partial \xi_a} - \frac{\partial F}{\partial \xi_a}\frac{\partial G}{\partial \eta_a}. \tag{2.8}$$

We say that a Hamiltonian function H is *real* if $H(z)$ is real for all real z.

Definition 2.2 *For a given $\rho > 0$, we denote by $\mathcal{H}_\rho$ the space of real Hamiltonians P satisfying*

$$P \in \mathcal{C}^1(\mathcal{L}_\rho, \mathbb{C}), \quad \textit{and} \quad X_P \in \mathcal{C}^1(\mathcal{L}_\rho, \mathcal{L}_\rho).$$

Notice that for F and G in $\mathcal{H}_\rho$ the formula (2.8) is well defined. With a given Hamiltonian function $H \in \mathcal{H}_\rho$, we associate the Hamiltonian system

$$\dot{z} = X_H(z) = J\nabla H(z)$$

which also reads

$$\dot{\xi}_a - -i\frac{\partial H}{\partial \eta_a} \quad \text{and} \quad \dot{\eta}_a = i\frac{\partial H}{\partial \xi_a}, \quad a \in \mathbb{Z}^d. \tag{2.9}$$

We define the local flow $\Phi_H^t(z)$ associated with the previous system (for an interval of times $t \geq 0$ depending a priori on the initial condition z). Note that if $z = (\xi, \bar{\xi})$ and if H is real, the flow $(\xi^t, \eta^t) = \Phi_H^t(z)$ is also real, $\xi^t = \bar{\eta}^t$ for all t. Further, choosing the Hamiltonian given by

$$H(\xi,\eta) = \sum_{a\in\mathbb{Z}^d} \omega_a \xi_a \eta_a + P(\xi,\eta),$$

P being given by (2.3), we recover the system (2.2), i.e. the expression of the NLS equation (1.1) in Fourier modes.

Remark 2.3 *The quadratic Hamiltonian $H_0 = \sum_{a\in\mathbb{Z}^d} \omega_a \xi_a \eta_a$ corresponding to the linear part of (1.1) does not belong to $\mathcal{H}_\rho$. Nevertheless it generates a flow which maps $\mathcal{L}_\rho$ into $\mathcal{L}_\rho$ explicitly given for all time t and for all indices a by $\xi_a(t) = e^{-i\omega_a t}\xi_k(0)$, $\eta_a(t) = e^{i\omega_a t}\eta_k(0)$. On the contrary, we will see that, in our setting, the nonlinearity P belongs to $\mathcal{H}_\rho$.*

2.2 Space of polynomials

In this subsection we define a class of polynomials on $\mathbb{C}^{\mathcal{Z}}$.
We first need more notations concerning multi-indices: let $\ell \geq 2$ and $\boldsymbol{j} = (j_1, \dots, j_\ell) \in \mathcal{Z}^\ell$ with $j_i = (a_i, \delta_i)$, we define

- the monomial associated with $\boldsymbol{j}$:

$$z_{\boldsymbol{j}} = z_{j_1} \cdots z_{j_\ell},$$

- the momentum of $\boldsymbol{j}$:

$$\mathcal{M}(\boldsymbol{j}) = a_1\delta_1 + \cdots + a_\ell\delta_\ell, \tag{2.10}$$

- the divisor associated with $\boldsymbol{j}$:

$$\Omega(\boldsymbol{j}) = \delta_1\omega_{a_1} + \cdots + \delta_\ell\omega_{a_\ell} \tag{2.11}$$

  where, for $a \in \mathbb{Z}^d$, $\omega_a = |a|^2 + v_a$ are the frequencies of the linear part of (1.1).

We then define the set of indices with **zero momentum** by

$$\mathcal{I}_\ell = \{\boldsymbol{j} = (j_1, \dots, j_\ell) \in \mathcal{Z}^\ell, \quad \text{with} \quad \mathcal{M}(\boldsymbol{j}) = 0\}. \tag{2.12}$$

On the other hand, we say that $\boldsymbol{j} = (j_1, \dots, j_\ell) \in \mathcal{Z}^\ell$ is **resonant**, and we write $\boldsymbol{j} \in \mathcal{N}_r$, if ℓ is even and $\boldsymbol{j} = \boldsymbol{i} \cup \bar{\boldsymbol{i}}$ for some choice of $\boldsymbol{i} \in \mathcal{Z}^{\ell/2}$. In particular, if $\boldsymbol{j}$ is resonant then its associated divisor vanishes, $\Omega(\boldsymbol{j}) = 0$, and its associated monomials depends only on the actions:

$$z_{\boldsymbol{j}} = z_{j_1} \cdots z_{j_r} = \xi_{a_1}\eta_{a_1} \cdots \xi_{a_{\ell/2}}\eta_{a_{\ell/2}} = I_{a_1} \cdots I_{a_{\ell/2}},$$

where for all $a \in \mathbb{Z}^d$, $I_a(z) = \xi_a\eta_a$ denotes the action associated with the index a.
Finally we note that if z is real, then $I_a(z) = |\xi_a|^2$ and we remark that for odd r the resonant set $\mathcal{N}_r$ is the empty set.

Definition 2.4 *Let $k \geq 2$, a (formal) polynomial $P(z) = \sum a_{\boldsymbol{j}} z_{\boldsymbol{j}}$ belongs to $\mathcal{P}_k$ if P is real, of degree k, have a zero of order at least 2 in $z = 0$, and if*

- *P contains only monomials having zero momentum, i.e. such that $\mathcal{M}(\boldsymbol{j}) = 0$ when $a_{\boldsymbol{j}} \neq 0$ and thus P reads*

$$P(z) = \sum_{\ell=2}^{k} \sum_{\boldsymbol{j}\in\mathcal{I}_\ell} a_{\boldsymbol{j}} z_{\boldsymbol{j}} \tag{2.13}$$

  with the relation $a_{\bar{\boldsymbol{j}}} = \bar{a}_{\boldsymbol{j}}$.

- *The coefficients $a_{\boldsymbol{j}}$ are bounded, i.e.* $\forall\, \ell = 2, \dots, k$, $\sup_{\boldsymbol{j}\in\mathcal{I}_\ell} |a_{\boldsymbol{j}}| < +\infty$.

We endow $\mathcal{P}_k$ with the norm

$$\|P\| = \sum_{\ell=2}^{k} \sup_{\boldsymbol{j}\in\mathcal{I}_\ell} |a_{\boldsymbol{j}}|. \tag{2.14}$$

The zero momentum assumption in Definition 2.4 is crucial to obtain the following Proposition:

Proposition 2.5 *Let $k \geq 2$ and $\rho > 0$. We have $\mathcal{P}_k \subset \mathcal{H}_\rho$, and for P a homogeneous polynomial of degree k in $\mathcal{P}_k$, we have the estimates*

$$|P(z)| \leq \|P\|\, \|z\|_\rho^k \tag{2.15}$$

and

$$\forall\, z \in \mathcal{L}_\rho, \quad \|X_P(z)\|_\rho \leq 2k\|P\|\, \|z\|_\rho^{k-1}. \tag{2.16}$$

Eventually, for $P \in \mathcal{P}_k$ and $Q \in \mathcal{P}_\ell$, then $\{P, Q\} \in \mathcal{P}_{k+\ell-2}$ and we have the estimate

$$\|\{P, Q\}\| \leq 2k\ell\|P\|\, \|Q\|. \tag{2.17}$$

Proof. Let

$$P(z) = \sum_{\boldsymbol{j}\in\mathcal{I}_k} a_{\boldsymbol{j}} z_{\boldsymbol{j}},$$

we have

$$|P(z)| \leq \|P\| \sum_{\boldsymbol{j}\in\mathcal{Z}^k} |z_{j_1}| \cdots |z_{j_k}| \leq \|P\|\, \|z\|_{\ell^1}^k \leq \|P\|\, \|z\|_\rho^k$$

and the first inequality (2.15) is proved.
To prove the second estimate, let us take $\ell \in \mathcal{Z}$ and calculate using the zero momentum condition,

$$\left|\frac{\partial P}{\partial z_\ell}\right| \leq k\|P\| \sum_{\substack{\boldsymbol{j}\in\mathcal{Z}^{k-1} \\ \mathcal{M}(\boldsymbol{j})=-\mathcal{M}(\ell)}} |z_{j_1} \cdots z_{j_{k-1}}|.$$

Therefore

$$\|X_P(z)\|_\rho = \sum_{\ell\in\mathcal{Z}} e^{\rho|\ell|} \left|\frac{\partial P}{\partial z_\ell}\right| \leq k\|P\| \sum_{\ell\in\mathcal{Z}} \sum_{\substack{\boldsymbol{j}\in\mathcal{Z}^{k-1} \\ \mathcal{M}(\boldsymbol{j})=-\mathcal{M}(\ell)}} e^{\rho|\ell|} |z_{j_1} \cdots z_{j_{k-1}}|.$$

But if $\mathcal{M}(\boldsymbol{j}) = -\mathcal{M}(\ell)$,

$$e^{\rho|\ell|} \leq \exp\left(\rho(|j_1| + \cdots + |j_{k-1}|)\right) \leq \prod_{n=1,\dots,k-1} e^{\rho|j_n|}.$$

Hence, after summing in ℓ we get[3]

$$\|X_P(z)\|_\rho \leq 2k\|P\| \sum_{\boldsymbol{j}\in\mathcal{Z}^{k-1}} e^{\rho|j_1|}|z_{j_1}| \cdots e^{\rho|j_{k-1}|}|z_{j_{k-1}}| \leq 2k\|P\|\, \|z\|_\rho^{k-1}$$

[3]Take care that $\mathcal{M}(a, \delta) = \mathcal{M}(-a, -\delta)$ whence the coefficient 2.

which yields (2.16).
Assume now that P and Q are homogeneous polynomials of degrees k and ℓ respectively and with coefficients $a_{\boldsymbol{k}}$, $\boldsymbol{k} \in \mathcal{I}_k$ and $b_{\boldsymbol{\ell}}$, $\boldsymbol{\ell} \in \mathcal{I}_\ell$. It is clear that $\{P, Q\}$ is a monomial of degree $k + \ell - 2$ satisfying the zero momentum condition. Furthermore writing

$$\{P, Q\}(z) = \sum_{\boldsymbol{j} \in \mathcal{I}_{k+\ell-2}} c_{\boldsymbol{j}} z_{\boldsymbol{j}},$$

$c_{\boldsymbol{j}}$ expresses as a sum of coefficients $a_{\boldsymbol{k}} b_{\boldsymbol{\ell}}$ for which there exists an $a \in \mathbb{Z}^d$ and $\epsilon \in \{\pm 1\}$ such that

$$(a, \epsilon) \subset \boldsymbol{k} \in \mathcal{I}_k \quad \text{and} \quad (a, -\epsilon) \subset \boldsymbol{\ell} \in \mathcal{I}_\ell,$$

and such that if for instance $(a, \epsilon) = k_1$ and $(a, -\epsilon) = \ell_1$, we necessarily have $(k_2, \dots, k_k, \ell_2, \dots, \ell_\ell) = \boldsymbol{j}$. Hence for a given $\boldsymbol{j}$, the zero momentum condition on $\boldsymbol{k}$ and on $\boldsymbol{\ell}$ determines the value of ϵa which in turn determines two possible value of (ϵ, a).
This proves (2.17) for monomials. The extension to polynomials follows from the definition of the norm (2.14).
The last assertion, as well as the fact that the Poisson bracket of two real Hamiltonian is real, immediately follow from the definitions. ∎

2.3 Nonlinearity

The nonlinearity g in (1.1) is assumed to be complex analytic in a neighborhood of $\{0, 0\}$ in $\mathbb{C}^2$. So there exist positive constants M and R_0 such that the Taylor expansion

$$g(v_1, v_2) = \sum_{k_1, k_2 \geq 0} \frac{1}{k_1! k_2!} \partial_{k_1} \partial_{k_2} g(0, 0) v_1^k v_2^\ell$$

is uniformly convergent and bounded by M on the ball $|v_1| + |v_2| \leq 2R_0$ of $\mathbb{C}^2$. Hence, formula (2.3) defines an analytic function on the ball $\|z\|_\rho \leq R_0$ of $\mathcal{L}_\rho$ and we have

$$P(z) = \sum_{k \geq 0} P_k(z)$$

where, for all $k \geq 0$, P_k is a homogeneous polynomial defined by

$$P_k = \sum_{k_1 + k_2 = k} \sum_{(\boldsymbol{a}, \boldsymbol{b}) \in (\mathbb{Z}^d)^{k_1} \times (\mathbb{Z}^d)^{k_1}} p_{\boldsymbol{a}, \boldsymbol{b}} \xi_{a_1} \cdots \xi_{a_{k_1}} \eta_{b_1} \cdots \eta_{b_{k_2}}$$

with

$$p_{\boldsymbol{a}, \boldsymbol{b}} = \frac{1}{k_1! k_2!} \partial_{k_1} \partial_{k_2} g(0, 0) \int_{\mathbb{T}^d} e^{i \mathcal{M}(\boldsymbol{a}, \boldsymbol{b}) \cdot x} \, \mathrm{d}x,$$

and $\mathcal{M}(\boldsymbol{a}, \boldsymbol{b}) = a_1 + \cdots + a_{k_1} - b_1 - \cdots - b_{k_2}$ is the moment of $\xi_{a_1} \cdots \xi_{a_{k_1}} \eta_{b_1} \cdots \eta_{b_{k_2}}$. Therefore it is clear that P_k satisfies the zero momentum condition and thus $P_k \in \mathcal{P}_k$ for all $k \geq 0$. Furthermore $\|P_k\| \leq M R_0^{-k}$.

2.4 Non resonance condition

In order to control the divisors (2.11), we need to impose a non resonance condition on the linear frequencies ω_a, $a \in \mathbb{Z}^d$.
For $r \geq 3$ and $\boldsymbol{j} = (j_1, \dots, j_r) \in \mathcal{Z}^r$, we define $\mu(\boldsymbol{j})$ as the third largest integer amongst $|j_1|, \cdots, |j_r|$ and we recall that $\boldsymbol{j} \in \mathcal{Z}^r$ is said resonant if r is even and $\boldsymbol{j} = \boldsymbol{i} \cup \bar{\boldsymbol{i}}$ for some $\boldsymbol{i} \in \mathcal{Z}^{r/2}$.

Hypothesis 2.6 *There exist $\gamma > 0$, $\nu > 0$ and $c_0 > 0$ such that for all $r \geq 3$ and all $\boldsymbol{j} \in \mathcal{Z}^r$ non resonant, we have*

$$|\Omega(\boldsymbol{j})| \geq \frac{\gamma c_0^r}{\mu(\boldsymbol{j})^{\nu r}}. \tag{2.18}$$

Recall that for $V = \sum_{a \in \mathbb{Z}^d} v_a e^{ia\cdot x}$ in the space $\mathcal{W}_m$ defined in (1.2), the frequencies read

$$\omega_a = |a|^2 + v_a = |a|^2 + \frac{R v'_a}{(1+|a|)^m}, \quad a \in \mathbb{Z}^d.$$

In Appendix we prove

Proposition 2.7 *Fix $\gamma > 0$ small enough and $m > d/2$. There exist positive constants c_0 and ν depending only on m, R and d, and a set $F_\gamma \subset \mathcal{W}_m$ whose measure is larger than $1 - 4\gamma^{1/7}$ such that if $V \in F_\gamma$ then (2.18) holds true for all non resonant $\boldsymbol{j} \in \mathcal{Z}^r$ and all $r \geq 3$.*

Thus Hypothesis 2.6 is satisfied for all $V \in \mathcal{V}$ where

$$\mathcal{V} = \cup_{\gamma > 0} F_\gamma \tag{2.19}$$

is a subset of full measure in $\mathcal{W}_m$.

2.5 Normal forms

We fix an index $N \geq 1$. For a fixed integer $k \geq 3$, we set

$$\mathcal{J}_k(N) = \{ \boldsymbol{j} \in \mathcal{I}_k \mid \mu(\boldsymbol{j}) > N \}.$$

Definition 2.8 *Let N be an integer. We say that a polynomial $Z \in \mathcal{P}_k$ is in N-normal form if it can be written*

$$Z = \sum_{\ell=3}^{k} \sum_{\boldsymbol{j} \in \mathcal{N}_\ell \cup \mathcal{J}_\ell(N)} a_{\boldsymbol{j}} z_{\boldsymbol{j}}$$

In other words, Z contains either monomials depending only of the actions or monomials whose indices $\boldsymbol{j}$ satisfies $\mu(\boldsymbol{j}) > N$, i.e. monomials involving at least three modes with index greater than N.

We now motivate the introduction of such normal form. First, we recall the

Lemma 2.9 *let $f : \mathbb{R} \to \mathbb{R}_+$ a continuous function, and $y : \mathbb{R} \to \mathbb{R}_+$ a differentiable function satisfying the inequality*

$$\forall\, t \in \mathbb{R}, \quad \frac{\mathrm{d}}{\mathrm{d}t} y(t) \leq 2f(t)\sqrt{y(t)}.$$

Then we have the estimate

$$\forall\, t \in \mathbb{R}, \quad \sqrt{y(t)} \leq \sqrt{y(0)} + \int_0^t f(s)\,\mathrm{d}s.$$

Proof. Let $\epsilon > 0$ and define $y_\epsilon = y + \epsilon$ which is a non negative function whose square root is derivable. We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\sqrt{y_\epsilon(t)} \leq 2f(t)\frac{\sqrt{y(t)}}{\sqrt{y_\epsilon(t)}} \leq 2f(t)$$

and thus

$$\sqrt{y_\varepsilon(t)} \leq \sqrt{y_\epsilon(0)} + \int_0^t f(s)\,\mathrm{d}s.$$

The claim is obtained when $\epsilon \to 0$. ∎

For a given number N and $z \in \mathcal{L}_\rho$ we define

$$\mathsf{R}_\rho^N(z) = \sum_{|j|>N} e^{\rho|j|}|z_j|.$$

Notice that if $z \in \mathcal{L}_{\rho+\mu}$ then

$$\mathsf{R}_\rho^N(z) \leq e^{-\mu N}\|z\|_{\rho+\mu}. \tag{2.20}$$

Proposition 2.10 *Let $N \in \mathbb{N}$ and $k \geq 3$. Let Z a homogeneous polynomial of degree k in N-normal form. Let $z(t)$ be a real solution of the flow associated with the Hamiltonian $H_0 + Z$. Then we have*

$$\mathsf{R}_\rho^N(t) \leq \mathsf{R}_\rho^N(0) + 4k^3\|Z\| \int_0^t \mathsf{R}_\rho^N(s)^2\|z(s)\|_\rho^{k-3}\,\mathrm{d}s \tag{2.21}$$

and

$$\|z(t)\|_\rho \leq \|z(0)\|_\rho + 4k^3\|Z\| \int_0^t \mathsf{R}_\rho^N(s)^2\|z(s)\|_\rho^{k-3}\,\mathrm{d}s \tag{2.22}$$

Proof. Let $a \in \mathbb{Z}^d$ be fixed, and let $I_a(t) = \xi_a(t)\eta_a(t)$ the actions associated with the solution of the Hamiltonian system induced by $H_0 + Z$. We have using (2.17) and $H_0 = H_0(I)$,

$$|e^{2\rho|a|}\dot{I}_a| = |e^{2\rho|a|}\{I_a, Z\}| \leq 2k\|Z\|\,|e^{\rho|a|}\sqrt{I_a}|\,\Big(\sum_{\substack{\mathcal{M}(\boldsymbol{j})=\pm a \\ \text{2 indices}>N}} e^{\rho|a|}|z_{j_1}\cdots z_{j_{k-1}}|\Big)$$

Using the previous Lemma, we get

$$e^{\rho|a|}\sqrt{I_a(t)} \leq e^{\rho|a|}\sqrt{I_a(0)} + 2k\|Z\| \int_0^t \Big(\sum_{\substack{\mathcal{M}(\boldsymbol{j})=\pm a \\ \text{2 indices}>N}} e^{\rho|j_1|}|z_{j_1}|\cdots e^{\rho|j_{k-1}|}|z_{j_{k-1}}|\Big)\mathrm{d}s. \tag{2.23}$$

Ordering the multi-indices in such way $|j_1|$ and $|j_2|$ are the largest, and using the fact that $z(t)$ is real (and thus $|z_j| = \sqrt{I_a}$ for $j = (a, \pm 1) \in \mathcal{Z}$), we obtain after summation in $|a| > N$

$$\mathsf{R}_\rho^N(z(t)) \le \mathsf{R}_\rho^N(z(0)) + 4k^3\|Z\| \int_0^t \Big(\sum_{\substack{|j_1|,|j_2|\ge N \\ j_3,\cdots,j_k\in\mathcal{Z}}} e^{\rho|j_1|}|z_{j_1}|\cdots e^{\rho|j_{k-1}|}|z_{j_{k-1}}|\Big)\mathrm{d}s$$

$$\le \mathsf{R}_\rho^N(0) + 4k^3\|Z\| \int_0^t \mathsf{R}_\rho^N(s)^2\|z(s)\|_\rho^{k-3}\,\mathrm{d}s.$$

In the same way we obtain (2.22). ∎

Remark 2.11 *These estimates will be crucially used in the final bootstrap argument. In particular, along the solution associated with a Hamiltonian in N-normal formal and initial datum $\|z_0\|_{2\rho} = \varepsilon$. Then as $\mathsf{R}_\rho^N(z_0) = \mathcal{O}(\varepsilon e^{-\rho N})$, Eqns. (2.21)-(2.22) guarantee that $\mathsf{R}_\rho^N(z(t))$ remains of order $\mathcal{O}(\varepsilon e^{-\rho N})$ and the norm of $z(t)$ remains of order ε over exponentially long time $t = \mathcal{O}(e^{\rho N})$.*

The next result is an easy consequence of the non resonance condition and the definition of the normal forms:

Proposition 2.12 *Assume that the non resonance condition (2.18) is satisfied, and let N be fixed. Let Q be a homegenous polynomial of degree k. Then the homological equation*

$$\{\chi, H_0\} - Z = Q \tag{2.24}$$

admits a polynomial solution (χ, Z) homogeneous of degree k, such that Z is in N-normal form, and such that

$$\|Z\| \le \|Q\| \quad \text{and} \quad \|\chi\| \le \frac{N^{\nu k}}{\gamma c_0^k}\|Q\| \tag{2.25}$$

Proof. Assume that $Q = \sum_{\boldsymbol{j}\in\mathcal{I}_k} Q_{\boldsymbol{j}} z_{\boldsymbol{j}}$ and search $Z = \sum_{\boldsymbol{j}\in\mathcal{I}_k} Z_{\boldsymbol{j}} z_{\boldsymbol{j}}$ and $\chi = \sum_{\boldsymbol{j}\in\mathcal{I}_k} \chi_{\boldsymbol{j}} z_{\boldsymbol{j}}$ such that (2.24) be satisfied. Then the equation (2.24) can be written in term of polynomial coefficients

$$i\Omega(\boldsymbol{j})\chi_{\boldsymbol{j}} - Z_{\boldsymbol{j}} = Q_{\boldsymbol{j}}, \quad j \in \mathcal{I}_k,$$

where $\Omega(\boldsymbol{j})$ is defined in (2.11). We then define

$$\begin{array}{llll} Z_{\boldsymbol{j}} = Q_{\boldsymbol{j}} & \text{and} & \chi_{\boldsymbol{j}} = 0 & \text{if} \quad \boldsymbol{j} \notin \mathcal{N}_k \text{ or } \mu(\boldsymbol{j}) \le N, \\ Z_{\boldsymbol{j}} = 0 & \text{and} & \chi_{\boldsymbol{j}} = \frac{Q_{\boldsymbol{j}}}{i\Omega(\boldsymbol{j})} & \text{if} \quad \boldsymbol{j} \in \mathcal{N}_k \text{ and } \mu(\boldsymbol{j}) > N. \end{array}$$

In view of (2.18), this yields (2.25). ∎

3 Proof of the main Theorem

3.1 Recursive equation

We aim at constructing a canonical transformation τ such that in the new variables, the Hamiltonian $H_0 + P$ is under normal form modulo a small remainder term. Using Lie transforms to

generate τ, the problem can be written: Find polynomials $\chi = \sum_{k=3}^{r} \chi_k$ and $Z = \sum_{k=3}^{r} Z_k$ under normal form and a smooth Hamiltonian R satisfying $\partial^\alpha R(0) = 0$ for all $\alpha \in \mathbb{N}^{\mathcal{Z}}$ with $|\alpha| \geq r$, such that

$$(H_0 + P) \circ \Phi_\chi^1 = H_0 + Z + R. \tag{3.1}$$

Then the exponential estimate will by obtained by optimizing the choice of r and N.
We recall that for χ and K to Hamiltonian, we have for all $k \geq 0$

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}(K \circ \Phi_\chi^t) = \{\chi, \{\cdots\{\chi, K\}\cdot\}(\Phi_\chi^t) = (\mathrm{ad}_\chi^k K)(\Phi_\chi^t),$$

where $\mathrm{ad}_\chi K = \{\chi, K\}$. On the other hand, if K, L are homogeneous polynomials of degree respectively k and ℓ then $\{K, L\}$ is a homogeneous polynomial of degree $k+l-2$. Therefore, we obtain by using the Taylor formula

$$(H_0 + P) \circ \Phi_\chi^1 - (H_0 + P) = \sum_{k=0}^{r-3} \frac{1}{(k+1)!} \mathrm{ad}_\chi^k(\{\chi, H_0 + P\}) + \mathcal{O}_r \tag{3.2}$$

where $\mathcal{O}_r$ stands for any smooth function R satisfying $\partial^\alpha R(0) = 0$ for all $\alpha \in \mathbb{N}^{\mathcal{Z}}$ with $|\alpha| \geq r$. Now we know that for $\zeta \in \mathbb{C}$, the following relation holds:

$$\left(\sum_{k=0}^{r-3} \frac{B_k}{k!}\zeta^k\right)\left(\sum_{k=0}^{r-3} \frac{1}{(k+1)!}\zeta^k\right) = 1 + O(|\zeta|^{r-2})$$

where B_k are the Bernoulli numbers defined by the expansion of the generating function $\frac{z}{e^z-1}$. Therefore, defining the two differential operators

$$A_r = \sum_{k=0}^{r-3} \frac{1}{(k+1)!} \mathrm{ad}_\chi^k \quad \text{and} \quad B_r = \sum_{k=0}^{r-3} \frac{B_k}{k!} \mathrm{ad}_\chi^k,$$

we get

$$B_r A_r = \mathrm{Id} + C_r$$

where C_r is a differential operator satisfying

$$C_r \mathcal{O}_3 = \mathcal{O}_r.$$

Applying B_r to the two sides of equation (3.2), we obtain

$$\{\chi, H_0 + P\} = B_r(Z - P) + \mathcal{O}_r.$$

Plugging the decompositions in homogeneous polynomials of χ, Z and P in the last equation and equating the terms of same degree, we obtain after a straightforward calculus, the following recursive equations

$$\{\chi_m, H_0\} - Z_m = Q_m, \quad m = 3, \cdots, r, \tag{3.3}$$

where

$$
\begin{aligned}
Q_m = -P_m + \sum_{k=3}^{m-1} \{P_{m+2-k}, \chi_k\} \\
+ \sum_{k=1}^{m-3} \frac{B_k}{k!} \sum_{\substack{\ell_1+\cdots\ell_{k+1}=m+2k \\ 3\le \ell_i \le m-k}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} (Z_{\ell_{k+1}} - P_{\ell_{k+1}}).
\end{aligned}
\tag{3.4}
$$

Notice that in the last sum, $\ell_i \le m-k$ as a consequence of $3 \le \ell_i$ and $\ell_1 + \cdots \ell_{k+1} = m+2k$. Once these recursive equations solved, we define the remainder term as $R = (H_0 + P) \circ \Phi^1_\chi - H_0 - Z$. By construction, R is analytic on a neighborhood of the origin in $\mathcal{L}_\rho$ and $R = \mathcal{O}_r$. As a consequence, by the Taylor formula,

$$
\begin{aligned}
R = \sum_{m\ge r+1} \sum_{k=1}^{m-3} \frac{1}{k!} \sum_{\substack{\ell_1+\cdots\ell_k=m+2k \\ 3\le \ell_i \le r}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} H_0 \\
+ \sum_{m\ge r+1} \sum_{k=0}^{m-3} \frac{1}{k!} \sum_{\substack{\ell_1+\cdots\ell_{k+1}=m+2k \\ 3\le \ell_1+\cdots\ell_k \le r \\ 3\le \ell_{k+1}}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} P_{\ell_{k+1}}.
\end{aligned}
\tag{3.5}
$$

Lemma 3.1 *Assume that the non resonance condition (2.18) is fulfilled. Let r and N be fixed. For $m = 3, \cdots, r$, there exist homogeneous polynomials χ_m and Z_m of degree m, with Z_m in $N-$normal form, solutions of the recursive equation (3.3) and satisfying*

$$
\|\chi_m\| + \|Z_m\| \le (CmN^\nu)^{m^2} \tag{3.6}
$$

where the constant C does not depend on r or N.

Proof. We define χ_m and Z_m by induction using Proposition 2.12. Note that (3.6) is clearly satisfied for $m = 3$, provided C is big enough. Estimate (2.25), together with (2.17) and the estimate on the Bernoulli numbers, $|B_k| \le k!\, c^k$ for some $c > 0$, yields for all $m \ge 3$,

$$
\begin{aligned}
\gamma c_0^m N^{-\nu m} \|\chi_m\| + \|Z_m\| \le \|P_m\| + 2 \sum_{k=3}^{m-1} k(m+2-k) \|P_{m+2-k}\| \, \|\chi_k\| \\
+ 2 \sum_{k=1}^{m-3} (Cm)^k \sum_{\substack{\ell_1+\cdots\ell_{k+1}=m+2k \\ 3\le \ell_i \le m-k}} \ell_1 \|\chi_{\ell_1}\| \cdots \ell_k \|\chi_{\ell_k}\| \, \|Z_{\ell_{k+1}} - P_{\ell_{k+1}}\| \, .
\end{aligned}
$$

for some constant C. We set $\beta_m = m(\|\chi_m\| + \|Z_m\|)$. Using $\|P_m\| \leq MR_0^{-m}$ (see end of subsection 2.4), we obtain

$$\beta_m \leq \beta_m^{(1)} + \beta_m^{(2)} \quad \text{where}$$

$$\beta_m^{(1)} = (CN^\nu)^m m^3 \sum_{k=3}^{m-1} \beta_k \quad \text{and}$$

$$\beta_m^{(2)} = N^{\nu m}(Cm)^{m-1} \sum_{k=1}^{m-3} \sum_{\substack{\ell_1+\cdots\ell_{k+1}=m+2k \\ 3\leq \ell_i \leq m-k}} \beta_{\ell_1}\cdots\beta_{\ell_k}(\beta_{\ell_{k+1}} + \|P_{\ell_{k+1}}\|)$$

where C depends on M, R_0, γ and c_0. It remains to prove by recurrence that $\beta_m \leq (CmN^\nu)^{m^2}$, $m \geq 3$. Again this is true for $m = 3$ adapting C if necessary. Thus assume that $\beta_j \leq (CjN^\nu)^{j^2}$ $j = 3, \ldots, m-1$, we then get for

$$\beta_m^{(1)} \leq (CN^\nu)^m m^4 (CmN^\nu)^{(m-1)^2} \leq (CmN^\nu)^{m^2-m+1} \leq \frac{1}{2}(CmN^\nu)^{m^2}$$

as soon as $m \geq 4$, and provided $C > 2$. On the other hand, since $\|P_m\| \leq MR_0^{-m}$, we can assume that $\|P_{\ell_{k+1}}\| \leq \beta_{\ell_{k+1}}$ and we get

$$\beta_m^{(2)} \leq N^{\nu m}(Cm)^{m-1} \sum_{k=1}^{m-3} \sum_{\substack{\ell_1+\cdots\ell_{k+1}=m+2k \\ 3\leq \ell_i \leq m-k}} (CN^\nu(m-k))^{\ell_1^2+\cdots+\ell_{k+1}^2}.$$

Notice that the maximum of $\ell_1^2+\cdots+\ell_{k+1}^2$ when $\ell_1+\cdots\ell_{k+1} = m+2k$ and $3 \leq \ell_i \leq m-k$ is obtained for $\ell_1 = \cdots = \ell_k = 3$ and $\ell_{k+1} = m-k$ and its value is $(m-k)^2+9k$. Furthermore the cardinal of $\{\ell_1 + \cdots \ell_{k+1} = m+2k,\ 3 \leq \ell_i \leq m-k\}$ is smaller than m^{k+1}, hence we obtain

$$\beta_m^{(2)} \leq \max_{k=1,\cdots,m-3} N^{\nu m}(Cm)^{m-1} Cm^{k+2}(CN^\nu(m-k))^{(m-k)^2+9k} \leq \frac{1}{2}(CmN^\nu)^{m^2}$$

for all $m \geq 4$ and adapting again C if necessary. ∎

3.2 Normal form result

For a number R_0 , we set $B_\rho(R_0) = \{z \in \mathcal{L}_\rho \,|\, \|z\|_\rho < R_0\}$.

Theorem 3.2 *Assume that P is analytic on a ball $B_\rho(R_0)$ for some $R_0 > 0$ and $\rho > 0$. Assume that the non resonance condition (2.18) is satisfied, and let $\beta < 1$ and $M > 1$ be fixed. Then there exist constants $\varepsilon_0 > 0$ and $\sigma > 0$ such that for all $\varepsilon < \varepsilon_0$, there exists: a polynomial χ, a polynomial Z in $|\ln\varepsilon|^{1+\beta}$ normal form, and a Hamiltonian R analytic on $B_\rho(M\varepsilon)$, such that*

$$(H_0 + P) \circ \Phi_\chi^1 = H_0 + Z + R. \tag{3.7}$$

Furthermore, for all $z \in B_\rho(M\varepsilon)$,

$$\|X_Z(z)\|_\rho + \|X_\chi(z)\|_\rho \leq 2\varepsilon^{3/2}, \quad \textit{and} \quad \|X_R(z)\|_\rho \leq \varepsilon\, e^{-\frac{1}{4}|\ln\varepsilon|^{1+\beta}}. \tag{3.8}$$

Proof. Using Lemma 3.1, for all N and r, we can construct polynomial Hamiltonians

$$\chi(z) = \sum_{k=3}^{r} \chi_k(z) \quad \text{and} \quad Z(z) = \sum_{k=3}^{r} Z_k(z),$$

with Z in N-normal form, such that (3.7) holds with $R = \mathcal{O}_r$. Now for fixed $\varepsilon > 0$, we choose

$$N \equiv N(\varepsilon) = |\ln \varepsilon|^{1+\beta} \quad \text{and} \quad r \equiv r(\varepsilon) = |\ln \varepsilon|^{\beta}.$$

This choice is motivated by the necessity of a balance between Z and R in (3.7): The error induced by Z is controlled as in Remark 2.11, while the error induced by R is controlled by Lemma 3.1. By (3.6), we have

$$\begin{aligned} \|\chi_k\| &\leq (CkN^\nu)^{k^2} \leq \exp(k(\nu k(1+\beta)\ln|\ln\varepsilon| + k\ln Ck)) \\ &\leq \exp(k(\nu r(1+\beta)\ln|\ln\varepsilon| + r\ln Cr)) \\ &\leq \exp(k|\ln\varepsilon|(\nu|\ln\varepsilon|^{\beta-1}(1+\beta)\ln|\ln\varepsilon| + |\ln\varepsilon|^{\beta-1}\ln C|\ln\varepsilon|^{\beta})) \\ &\leq \varepsilon^{-k/8}, \end{aligned} \tag{3.9}$$

as $\beta < 1$, and for $\varepsilon \leq \varepsilon_0$ sufficiently small. Therefore using Proposition 2.5, we obtain for $z \in B_\rho(M\varepsilon)$

$$|\chi_k(z)| \leq \varepsilon^{-k/8}(M\varepsilon)^k \leq M^k\varepsilon^{7k/8}$$

and thus

$$|\chi(z)| \leq \sum_{k\geq 3} M^k\varepsilon^{7k/8} \leq \varepsilon^{3/2}$$

for ε small enough. Similarly, we have for all $k \leq r$,

$$\|X_{\chi_k}(z)\|_\rho \leq 2k\varepsilon^{-k/8}(M\varepsilon)^{k-1} \leq 2kM^{k-1}\varepsilon^{7k/8-1}$$

and

$$\|X_{\chi}(z)\|_\rho \leq \sum_{k\geq 3} 2kM^{k-1}\varepsilon^{7k/8-1} \leq C\varepsilon^{-1}\varepsilon^{\frac{21}{8}} \leq \varepsilon^{3/2}$$

for ε small enough. Similar bounds clearly hold for $Z = \sum_{k=3}^{r} Z_k$, which shows the first estimate in (3.8).
On the other hand, using $\mathrm{ad}_{\chi_{\ell_k}} H_0 = Z_{\ell_k} + Q_{\ell_k}$ (see (3.3)), then using Lemma 3.1 and the definition of Q_m (see (3.4)), we get $\|\mathrm{ad}_{\chi_{\ell_k}} H_0\| \leq (CkN^\nu)^{\ell_k{}^2} \leq \varepsilon^{-\ell_k/8}$, where the last inequality proceeds as in (3.9). Thus, using (3.5), (3.9) and $\|P_{\ell_{k+1}}\| \leq MR_0^{-\ell_{k+1}}$ we obtain by Proposition 2.5 that for $z \in B_\rho(M\varepsilon)$

$$\|X_R(z)\|_\rho \leq \sum_{m\geq r+1} \sum_{k=0}^{m-3} m(Cr)^{3m}\varepsilon^{-\frac{m+2k}{8}}\varepsilon^{m-1} \leq \sum_{m\geq r+1} m^2(Cr)^{3m}\varepsilon^{m/2} \leq (Cr)^{3r}\varepsilon^{r/2}.$$

Therefore, since $r = |\ln\varepsilon|^\beta$, we get $\|X_R(z)\|_\rho \leq \varepsilon\, e^{-\frac{1}{4}|\ln\varepsilon|^{1+\beta}}$ for $z \in B_\rho(M\varepsilon)$ and ε small enough. ∎

3.3 Bootstrap argument

We are now in position to prove the main theorem of Section 1 which is actually a consequence of Theorem 3.2.

Let $u_0 \in \mathcal{A}_{2\rho}$ with $|u_0|_{2\rho} = \varepsilon$ and denotes by $z(0)$ the corresponding sequence of its Fourier coefficients which belongs, by Lemma 2.1, to in $\mathcal{L}_{\frac{3}{2}\rho}$ with $||z(0)||_{\frac{3}{2}\rho} \leq \frac{c_\rho}{4}\varepsilon$ with $c_\rho = \frac{2^{d+2}}{(1-e^{-\rho/2\sqrt{d}})^d}$. Let $z(t)$ be the local solution in $\mathcal{L}_\rho$ of the Hamiltonian system associated with $H = H_0 + P$.

Let χ, Z and R given by Theorem 3.2 with $M = c_\rho$ and let $y(t) = \Phi^1_\chi(z(t))$. We recall that since $\chi(z) = O(\|z\|^3)$, the transformation Φ^1_χ is close to the identity, $\Phi^1_\chi(z) = z + O(\|z\|^2)$ and thus, for ε small enough, we have $\|y(0)\|_{\frac{3}{2}\rho} \leq \frac{c_\rho}{2}\varepsilon$. In particular, as noticed in (2.20), $R^\rho_N(y(0)) \leq \frac{c_\rho}{2}\varepsilon\, e^{-\frac{\rho}{2}N} \leq \frac{c_\rho}{2}\varepsilon\, e^{-\sigma N}$ where $\sigma = \sigma_\rho \leq \frac{\rho}{2}$.

Let T_ε be maximum of time T such that $R^\rho_N(y(t)) \leq c_\rho\varepsilon\, e^{-\sigma N}$ and $\|y(t)\|_\rho \leq c_\rho\varepsilon$ for all $|t| \leq T$. By construction,

$$y(t) = y(0) + \int_0^t X_{H_0+Z}(y(s))ds + \int_0^t X_R(y(s))ds$$

so using (2.21) for the first flow and (3.8) for the second one, we get for $|t| < T_\varepsilon$,

$$\begin{aligned} R^\rho_N(y(t)) &\leq \frac{1}{2}c_\rho\varepsilon\, e^{-\sigma N} + 4|t|\sum_{k=3}^r \|Z_k\|\, k^3(c_\rho\varepsilon)^{k-1}e^{-2\sigma N} + |t|\varepsilon\, e^{-\frac{1}{4}|\ln\varepsilon|^{1+\beta}} \\ &\leq \left(\frac{1}{2} + 4|t|\sum_{k=3}^r \|Z_k\|\, k^3(c_\rho\varepsilon)^{k-2}e^{-\sigma N} + |t|\varepsilon\, e^{-\frac{1}{8}|\ln\varepsilon|^{1+\beta}}\right) c_\rho\varepsilon\, e^{-\sigma N} \end{aligned} \tag{3.10}$$

where in the last inequality we used $\sigma = \min\{\frac{1}{8}, \frac{\rho}{2}\}$ and $N = \ln|\varepsilon|^{1+\beta}$.

Using Lemma 3.1, we then verify

$$R^\rho_N(y(t)) \leq \left(\frac{1}{2} + C|t|\varepsilon\, e^{-\sigma N}\right) c_\rho\varepsilon\, e^{-\sigma N}$$

and thus, for ε small enough,

$$R^\rho_N(y(t)) \leq c_\rho\varepsilon\, e^{-\sigma N} \quad \text{for all} \quad |t| \leq \min\{T_\varepsilon, e^{\sigma N}\}. \tag{3.11}$$

Similarly we obtain

$$\|y(t)\|_\rho \leq c_\rho\varepsilon \quad \text{for all} \quad |t| \leq \min\{T_\varepsilon, e^{\sigma N}\}. \tag{3.12}$$

In view of the definition of T_ε, (3.11) and (3.12) imply $T_\varepsilon \geq e^{\sigma N}$. In particular $\|z(t)\|_\rho \leq 2c_\rho\varepsilon$ for $|t| \leq e^{\sigma N} = \varepsilon^{-\sigma|\ln\varepsilon|^\beta}$ and using (2.7), we finally obtain (1.3) with $C = \frac{2^{2d+5}}{(1-e^{-\rho/2\sqrt{d}})^{2d}}$.

Estimate (1.4) is an other consequence of the normal form result and Proposition 2.10. Actually we use that the Fourier coefficients of $u(t)$ are given by $z(t)$ which is ε^2-close to $y(t)$ which in turns is almost invariant: in view of (2.23) and as in (3.10), we have

$$\sum_{j\in\mathbb{Z}} e^{\rho|j|}\big||y_j(t)| - |y_j(0)|\big| \leq \left(4|t|\sum_{k=3}^r \|Z_k\|\, k^3(c_\rho\varepsilon)^{k-1}e^{-2\sigma N} + |t|\varepsilon\, e^{-\frac{1}{4}|\ln\varepsilon|^{1+\beta}}\right)$$

from which we deduce

$$\sum_{j\in\mathbb{Z}} e^{\rho|j|}\big||y_j(t)| - |y_j(0)|\big| \leq |t|\, e^{-\sigma N}$$

and then (1.4).

A Proof of the non resonance hypothesis

Instead of proving Proposition 2.7, we prove a slightly more general result. For a multi-index $\boldsymbol{j} \in \mathcal{Z}^r$ we define

$$N(\boldsymbol{j}) = \prod_{k=1}^{r} (1 + |j_k|).$$

Proposition A.1 *Fix $\gamma > 0$ small enough and $m > d/2$. There exist positive constants C and ν depending only on m, R and d, and a set $F_\gamma \subset \mathcal{W}_m$ whose measure is larger than $1 - 4\gamma$ such that if $V \in F_\gamma$ then for any $r \geq 1$*

$$|\Omega(\boldsymbol{j}) + \varepsilon_1 \omega_{\ell_1} + \varepsilon_2 \omega_{\ell_2}| \geq \frac{C^r \gamma^7}{N(\boldsymbol{j})^\alpha} \tag{A.1}$$

for any $\boldsymbol{j} \in \mathcal{Z}^r$, for any indexes $l_1,\ l_2 \in \mathbb{Z}^d$, and for any $\varepsilon_1, \varepsilon_2 \in \{0, 1, -1\}$ such that $(\boldsymbol{j}, (\ell_1, \varepsilon_1), (\ell_2, \varepsilon_2)) \notin \mathcal{N}_r$ is non resonant.

In order to prove proposition 2.7, we first prove that $\Omega(\boldsymbol{j})$ cannot accumulate on $\mathbb{Z}$. Precisely we have

Lemma A.2 *Fix $\gamma > 0$ and $m > d/2$. There exist $0 < C < 1$ depending only on m, R and d and a set $F'_\gamma \subset \mathcal{W}_m$ whose measure is larger than $1 - 4\gamma$ such that if $V \in F'_\gamma$ then for any $r \geq 1$*

$$|\Omega(\boldsymbol{j}) - b| \geq \frac{C^r \gamma}{N(\boldsymbol{j})^{m+d+3}} \tag{A.2}$$

for any non resonant $\boldsymbol{j} \in \mathbb{Z}^r$ and for any $b \in \mathbb{Z}$.

Proof. Let $(\alpha_1, \ldots, \alpha_r) \neq 0$ in $\mathbb{Z}^r$, $M > 0$ and $c \in \mathbb{R}$. By induction we can prove that the Lesbegue measure of

$$\{x \in [-M, M]^r \mid |\sum_{i=1}^{r} \alpha_i x_i + c| < \eta\}$$

is smaller than $(2M)^{r-1} 2\eta$. Hence given $\boldsymbol{j} = (a_i, \delta_i)_{i=1}^r \in \mathcal{Z}^r$, and $b \in \mathbb{Z}$, the Lesbegue measure of

$$\mathcal{X}_\eta := \left\{ x \in [-1/2, 1/2]^r \ : \ \left| \sum_{i=1}^{r} \delta_i (|a_i|^2 + x_i) - b \right| < \eta \right\}$$

is smaller than 2η. Now consider the set

$$\{V \in \mathcal{W}_m \mid |\Omega(\boldsymbol{j}) - b| < \eta\} = \left\{ V \in \mathcal{W}_m \mid \left| \sum_{i=1}^{r} \delta_i (|a_i|^2 + \frac{v_{a_i} R}{(1 + |a_i|)^m}) - b \right| < \eta \right\}, \tag{A.3}$$

it is contained in the set of the V's such that $(R v_{a_i}/(1 + |a_i|)^m)_{i=1}^r \in \mathcal{X}_\eta$. Hence the measure of (A.3) is smaller than $2R^{-r} N(\boldsymbol{j})^m \eta$. To conclude the proof we have to sum over all the $\boldsymbol{j}$'s and all the b's. Now for a given $\boldsymbol{j}$, remark that if $|\Omega(\boldsymbol{j}) - b| \geq \eta$ with $\eta \leq 1$ then $|b| \leq 2N(\boldsymbol{j})^2$.

So that to guarantee (A.2) for all possible choices of $\boldsymbol{j}$, b and r, it suffices to remove from $\mathcal{W}_m$ a set of measure

$$4\gamma \sum_{\boldsymbol{j}\in\mathcal{Z}^r} \frac{C^r}{R^r N(\boldsymbol{j})^{m+3+d}} N(\boldsymbol{j})^{m+2} \leq 4\gamma \left[\frac{2C}{R}\sum_{\ell\in\mathbb{Z}^d}\frac{1}{(1+|\ell|)^{d+1}}\right]^r.$$

Choosing $C \leq \frac{1}{2}R\big(\sum_{\ell\in\mathbb{Z}^d}\frac{1}{(1+|\ell|)^{d+1}}\big)^{-1}$ proves the result. ∎

Proof of proposition A.1. First of all, for $\varepsilon_1 = \varepsilon_2 = 0$, (A.1) is a direct consequence of lemma A.2 choosing $\nu \geq m+d+3$, $\gamma \leq 1$ and $F_\gamma = F'_\gamma$.
When $\varepsilon_1 = \pm 1$ and $\varepsilon_2 = 0$, (A.1) reads

$$|\Omega(\boldsymbol{j}) \pm \omega_{\ell_1}| \geq \frac{C^r\gamma}{N(\boldsymbol{j})^\nu}. \tag{A.4}$$

Notice that $|\Omega(\boldsymbol{j})| \leq N(\boldsymbol{j})^2$ and thus, if $|\ell_1| \geq 2N(\boldsymbol{j})$, (A.4) is always true. When $|\ell_1| \leq 2N(\boldsymbol{j})$, using that $N(\boldsymbol{j},\ell) = N(\boldsymbol{j})(1+|\ell_1|)$, we get applying lemma A.2 with $b = 0$,

$$|\Omega(\boldsymbol{j}) + \varepsilon_1\omega_{\ell_1}| = |\Omega(\boldsymbol{j},(\ell_1,\varepsilon_1))| \geq \frac{C^{r+1}\gamma}{N(\boldsymbol{j})^{m+d+3}(3N(\boldsymbol{j}))^{m+d+3}} \geq \frac{\tilde{C}^r\gamma}{N(\boldsymbol{j})^\nu}$$

with $\nu = 2(m+d+3)$ and $\tilde{C} = \frac{C^2}{3^{m+d+3}}$. In the same way we prove (A.1) when $\varepsilon_1\varepsilon_2 = 1$ with the same choice of ν. So it remains to establish an estimate of the form

$$|\Omega(\delta,\boldsymbol{j}) + \omega_{\ell_1} - \omega_{\ell_2}| \geq \frac{\tilde{C}^r\gamma^4}{N(\boldsymbol{j})^\nu}. \tag{A.5}$$

Assuming $|\ell_1| \leq |\ell_2|$, we have

$$|\omega_{\ell_1} - \omega_{\ell_2} - \ell_1^2 + \ell_2^2| \leq \left|\frac{R|v_{\ell_1}|}{(1+|\ell_1|)^m} - \frac{R|v_{\ell_2}|}{(1+|\ell_2|)^m}\right| \leq \frac{R}{(1+|\ell_1|)^m}.$$

Therefore if $(1+|\ell_1|)^m \geq \frac{2R}{C^r\gamma}N(\boldsymbol{j})^{m+d+3}$, we obtain (A.5) directly from lemma A.2 applied with $b = \ell_1^2 - \ell_2^2$ and choosing $\nu = m+d+3$, $\tilde{C} = C/2$ and $F_\gamma = F'_\gamma$.
Finally assume $(1+|\ell_1|)^m \leq \frac{2R}{C^r\gamma}N(\boldsymbol{j})^{m+d+3}$, taking into acount $|\Omega(\boldsymbol{j})| \leq N(\boldsymbol{j})^2$, (A.5) is satisfied when $\ell_2^2 - \ell_1^2 \leq 2N(\boldsymbol{j})^2$. So it remains to consider the case when

$$1+|\ell_1| \leq 1+|\ell_2| \leq \left[\left(\frac{2R}{C^r\gamma}N(\boldsymbol{j})^{m+d+3}\right)^{2/m} + 2N(\boldsymbol{j})^2\right]^{1/2} \leq \left(\frac{3R}{C^r\gamma}\right)^{\frac{1}{m}} N(\boldsymbol{j})^{\frac{m+d+3}{m}}.$$

Again we use lemma A.2 to conclude

$$\begin{aligned}|\Omega(\boldsymbol{j}) + \omega_{\ell_1} - \omega_{\ell_2}| &\geq \frac{C^{r+2}\gamma}{[N(\boldsymbol{j})(1+|\ell_1|)(1+|\ell_2|)]^{m+d+3}} \\ &\geq \frac{C^{r+2}\gamma\left(\frac{C^r\gamma}{3R}\right)^{\frac{m+d+3}{m}}}{N(\boldsymbol{j})^{m+d+3}N(\boldsymbol{j})^{2\frac{(m+d+3)^2}{m}}} \geq \frac{\tilde{C}^r\gamma^{4+3/m}}{N(\boldsymbol{j})^\nu}\end{aligned}$$

with $\nu = m+d+3+(m+d+3)^2/m$ and $\tilde{C} = \frac{C^{(4m+d+3)/m}}{3R}$. ∎

References

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

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

[BG03] D. Bambusi and B. Grebert, *Forme normale pour NLS en dimension quelconque*, Compt. Rendu. Acad. Sciences Paris **337** (2003), 409–414.

[BG06] D. Bambusi and B. Grébert, *Birkhoff normal form for PDE's with tame modulus*. Duke Math. J. 135 no. 3 (2006), 507—567.

[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** (1985), 1–25.

[Bou96] J. Bourgain, *Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations.* Geom. Funct. Anal. 6 (1996) 201–230.

[Caz03] T. Cazenave, *Semilinear Schrödinger equations*. Courant Lecture Notes in Mathematics, 10. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003.

[CKSTT09] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao, *Weakly turbulent solutions for the cubic defocusing nonlinear Schrödinger equation*. To appear in Inventiones Math.

[EK] H. L. Eliasson, S. B. Kuksin, *KAM for non-linear Schroedinger equation*, To appear in Annals of Math.

[FG10] E. Faou and B. Grébert, *Hamiltonian interpolation of splitting approximations for nonlinear PDEs.* arXiv:0912.2882v1

[FGP10] E. Faou, B. Grébert and E. Paturel, *Birkhoff normal form for splitting methods applied to semi linear Hamiltonian PDEs. Part II: Abstract splitting.* Numer. Math. 114 (2010) 459–490.

[GG85] L. Galgani and A. Giorgilli, *Rigourous estimates for the series expansions of Hamiltonian perturbation theory*, Celestial Mech. **37** (1985), 95–112.

[Gre07] B. Grébert, *Birkhoff normal form and Hamiltonian PDEs.* Séminaires et Congrès 15 (2007), 1–46

[Nek77] N. N. Nekhoroshev, *Exponential estimate of the stability of near integrable Hamiltonian systems*, Russ. Math. Surveys **32** (1977), no. 6, 1–65.

[Pös93] J. Pöschel, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, Math. Z. **213** (1993), 187–216.