NEKHOROSHEV THEOREM FOR SMALL AMPLITUDE SOLUTIONS IN NONLINEAR SCHRÖDINGER EQUATIONS

Dario BAMBUSI
Dipartimento di Matematica dell'Università,
Via Saldini 50, 20133 MILANO, Italy. *

Summary. We prove a Nekhoroshev type result [26,27] for the nonlinear Schrödinger equation

$$iu_t = -u_{xx} - mu - u\varphi(|u|^2) \ , \tag{0.1}$$

with vanishing or periodic boundary conditions on $[0, \pi]$; here $m \in \mathbb{R}$ is a parameter and $\varphi : \mathbb{R} \to \mathbb{R}$ is a function analytic in a neighborhood of the origin and such that $\varphi(0) = 0$, $\varphi'(0) \neq 0$. More precisely, we consider the Cauchy problem for (0.1) with initial data which extend to analytic integer functions of finite order, and prove that all the actions of the linearized system are approximate constants of motion up to times growing faster than any negative power of the size of the initial datum. The proof is obtained by a method which applies to Hamiltonian perturbations of linear PDE's with the following properties: (i) the linear dynamics is periodic (ii) there exists a finite order Birkhoff normal form which is integrable and quasi convex as a function of the action variables. Eq. (0.1) satisfies (i) and (ii) when restricted to a level surface of $\|u\|_{L^2}$, which is an integral of motion. The main technical tool used in the proof is a normal form lemma for systems with symmetry which is also proved here.

1. Introduction and Statement

In this paper a Nekhoroshev type result [26,27] for some nonlinear PDE's will be proved. To be definite we will concentrate on the nonlinear Schrödinger equation (0.1) (NLS) with vanishing boundary conditions on $[0, \pi]$ (the class of dynamical systems to which our technique applies will be specified later). It will be proved that for small amplitude smooth initial data all the actions of the linearized system are approximate constant of motions up to times growing faster than any negative power of the norm of the initial datum. This result will be obtained as a corollary of a theorem giving freezing of all the actions close (in H^1 norm) to small amplitude *finite dimensional* invariant tori of the linearized system.

* `e-mail`: BAMBUSI@MAT.UNIMI.IT

To come to a precise statement consider the space Fourier expansion of $u(x,t)$, namely

$$u(x,t)=\sum_{k\geq 1}u_k(t)\sqrt{\frac{2}{\pi}}\sin kx\ , \tag{1.1}$$

and define the action variables Z_k of the linearized system by

$$Z_k:=|u_k|^2/2\ ;$$

our first result is concerned with the Cauchy problem for (0.1) with H^1 initial datum; in the statement of the corresponding theorem we will denote by u^0 the initial datum, by u^0_k its k–th Fourier component, and by ε its H^1 norm, namely

$$\varepsilon:=\left\|u^0\right\|_{H^1}:=\left(\frac{1}{2}\int_0^\pi|u^0_x(x)|^2dx\right)^{1/2}\ .$$

Moreover we will denote by $\mathcal{J}$ the set of indexes labeling the Fourier modes in which the majority of the energy is stored (and we shall take into account the dependence of all the constants on $\mathcal{J}$).

Theorem 1.1. *Assume that the function φ is analytic in a neighborhood of the origin, that $\varphi(0)=0$ and $\varphi'(0)\neq 0$; then there exist positive constants $\varepsilon_*, C_1, C_2, C_3$ (dependent only on φ) with the following property: fix a finite set of indices $\mathcal{J}\subset\mathbf{N}\backslash\{0\}$, denote by $n<\infty$ its cardinality and*

$$\mathcal{C}(\mathcal{J}):=\sum_{i\in\mathcal{J}}i^2\ . \tag{1.2}$$

If the following conditions are satisfied:

$$\varepsilon<\left(\frac{\varepsilon_*}{\mathcal{C}(\mathcal{J})^5}\right)^n\ ,\quad \sum_{k\notin\mathcal{J}}k^2\frac{|u^0_k|^2}{2}\leq C_1\mathcal{C}(\mathcal{J})^2\varepsilon^3\ , \tag{1.3}$$

then, along the solutions of (0.1) with vanishing boundary conditions, one has

$$\sum_{k\in\mathcal{J}}k^2|Z_k(t)-Z_k(0)|+\sum_{k\notin\mathcal{J}}k^2|Z_k(t)|\leq C_2\mathcal{C}(\mathcal{J})^3\varepsilon^{2+1/n}\ , \tag{1.4}$$

for all times t satisfying

$$|t|\leq C_3\exp\left(\frac{\varepsilon_*}{\mathcal{C}(\mathcal{J})^5\varepsilon^{1/n}}\right)\ . \tag{1.5}$$

It is possible to rephrase theorem 1.1 as a stability result for finite dimensional tori. To this end fix a set of indices $\mathcal{J}$ with cardinality n, and a vector $\omega:=\{\omega_k\}_{k\in\mathcal{J}}$ with positive entries, which moreover satisfy the normalization condition $\sum_{k\in\mathcal{J}} k^2\omega_k = 1$. Given $\epsilon > 0$ consider the n–dimensional torus

$$\mathbf{T}^{\omega}_{\mathcal{J},\epsilon} := \left\{ u \in H^1_0([0,\pi],\mathbf{C}) \ : \ \frac{|u_k|^2}{2} = \epsilon\omega_k \ , \ k \in \mathcal{J} \ ; \ |u_k| = 0 \ , \ \forall k \notin \mathcal{J} \right\} \ .$$

Theorem 1.2. *There exist positive constants C_4, C_5 (depending on φ only), such that, provided $2\epsilon < (\varepsilon_*/\mathcal{C}(\mathcal{J})^5)^n$, solutions of (0.1) corresponding to initial data u^0 satisfying*

$$d_{H^1}(u^0, \mathbf{T}^{\omega}_{\mathcal{J},\epsilon}) \le C_4 \mathcal{C}(\mathcal{J})\varepsilon^{3/2} \ ,$$

(where $d_{H^1}(.,.)$ is the distance induced by the H^1 norm) remain close to $\mathbf{T}^{\omega}_{\mathcal{J},\epsilon}$, namely satisfy

$$d_{H^1}(u(t), \mathbf{T}^{\omega}_{\mathcal{J},\epsilon}) \le C_5 \mathcal{C}(\mathcal{J})^3 \varepsilon^{1+1/n} \ , \tag{1.6}$$

for exponentially long times, namely for the times satisfying (1.5).

Remark 1.3. *Eq. (1.4) implies*

$$\sum_{k\ge 1} k^2 |Z_k(t) - Z_k(0)| \le C_6 \mathcal{C}(\mathcal{J})^3 \varepsilon^{2+1/n} \ , \tag{1.7}$$

which however is considerably weaker than (1.4), since it yields control of the distance of $u(t)$ from $\mathbf{T}^{\omega}_{\mathcal{J},\epsilon}$ only in L^2 norm.

In the case of periodic boundary conditions (which can also be dealt with by our technique) theorem 1.1 could be reformulated in terms of finite gap solutions and action variables of the Zakhararov–Shabat nonlinear Schrödinger equation instead than invariant tori and action variables of the linearized system. This can be done just exploiting analyticity properties of action variables of the Zakhararov–Shabat equation [5,19].

The second of (1.3) is automatically satisfied for smooth initial data. Precisely, fix $\alpha > 1$, $\beta > 0$, and define the space $\mathcal{A}_{\alpha,\beta}$ of the functions u such that

$$\|u\|^2_{\alpha,\beta} := \sum_{k\ge 1} e^{\beta|k|^\alpha} \frac{|u_k|^2}{2} < \infty \ . \tag{1.8}$$

Having fixed a positive K we consider the Cauchy problem for eq. (0.1) with an $\mathcal{A}_{\alpha,\beta}$ initial datum satisfying

$$\left\|u^0\right\|_{\alpha,\beta} < K \ ;$$

then we have the following

Corollary 1.4. *There exist positive constants C_7, C_8, C_9, C_{10}, depending on α, β and φ only, and C_{11} depending also on K, with the following property: define*

$$\eta(\varepsilon) := \exp\left[-C_7 \left|\ln\varepsilon\right|^{(\alpha-1)/\alpha} \right] \ , \quad n_\varepsilon := C_{10} \left|\ln\varepsilon\right|^{1/\alpha} \ ;$$

if $\varepsilon := \left\|u^0\right\|_{H^1} < C_{11}$ *then one has*

$$\sum_{1\leq k\leq n_\varepsilon} k^2|Z_k(t)-Z_k(0)| + \sum_{k>n_\varepsilon} k^2|Z_k(t)| \leq C_8\varepsilon^2\eta(\varepsilon) \ , \tag{1.9}$$

for all times t *satisfying*

$$|t| \leq C_3 \exp\left(C_9/\eta(\varepsilon)\right) \tag{1.10}$$

Remark 1.5. *The union over* α *and* β *of the spaces* $\mathcal{A}_{\alpha,\beta}$ *coincides with the space of analytic skew symmetric* 2π*–periodic integer functions of finite order. In the appendix we will recall the main facts on integer functions of finite order and discuss more in detail this point.*

Remark 1.6. *The function* $\eta(\varepsilon)$ *tends to zero as* $\varepsilon \to 0$ *with a rate slower than any power of* ϵ*, while the function*

$$\exp\left(C_9/\eta(\varepsilon)\right)$$

tends to infinity faster than any negative power of ε*.*

Remark 1.7. *In the above corollary 1.4 the control of the size of the variation of the actions (cf. eq. (1.9)) and of the times over which it holds (cf. eq. (1.10)) depend only on the* H^1 *norm of the initial datum, while the threshold of applicability of the corollary (i.e. the constant* C_{11}*) depends also on the* $\mathcal{A}_{\alpha,\beta}$ *norm of* u^0 *(i.e. on the constant* K*). Obviously the* H^1 *norm can be controlled by the* $\mathcal{A}_{\alpha,\beta}$ *norm, and therefore, in corollary 1.4, one could substitute everywhere* $\varepsilon_{\alpha,\beta} := \left\|u^0\right\|_{\alpha,\beta}$ *to* ε*, but this would give rise to worst estimates.*

Remark 1.8. *An analog of corollary 1.4 could be proved for a slightly more general class of initial data containing also integer functions of infinite order. Such a class is obtained by fixing a monotonically increasing function* g *with the property that for any* $\beta > 0$ *one has*

$$\lim_{N\to\infty} \frac{e^{g(N)}}{N^\beta} = +\infty \ ,$$

and defining the space $\mathcal{A}_g$ *composed by the functions on* $[0,\pi]$ *such that*

$$\|u\|_g^2 := \sum_{k\geq 1} e^{kg(k)}\frac{|u_k|^2}{2} < \infty \ . \tag{1.11}$$

Corresponding to initial data in $\mathcal{A}_g$*, with small* H^1 *norm, one has again freezing of all the actions for times longer than any negative power of* $\varepsilon = \|u\|_{H^1}$*.*

The proof of the above results is based on a technique that applies to more general dynamical systems than NLS; indeed one can deal with Hamiltonian perturbations of linear systems with the following properties: (i) the linear dynamics is periodic with a period independent of the initial datum; (ii) there exists a finite order Birkhoff normal form which is integrable and quasi convex as a function of the action variables. Actually, one can deal also with systems that do not directly fulfill the above requirements, provided there exist some integrals of motion with the property that the restriction of the system to the surface of constant value of the integrals satisfy (i) and (ii). This is exactly the situation of NLS.

The rest of the paper is organized as follows: Sect. 2 contains the scheme of the proof of theorem 1.1; in sect. 3 we compare our result with some previous results on the dynamics of NLS and with some recent Nekhoroshev type results for PDE's. Sect. 4 contains the statement of an abstract normal form theorem for systems with approximate symmetry which is used in the proof of theorem 1.1 (we think it could have some independent interest). In sect. 5 we apply (twice) such a normal form theorem to NLS, and in sect. 6 we use the so obtained normal form to prove theorem 1.1. Sect. 7 contains the proof of the abstract normal form theorem of sect. 4, and sect. 8 contains some technical lemmas.

Acknowledgement I thank Antonio Giorgilli for many useful discussions. In particular the idea of performing two or more subsequent normalization in order to study systems with many degrees of freedom was suggested by him several years ago. I thank also Jürgen Pöschel and Andrea Carati whose comments allowed to improve the paper. This work has been developed with the partial support of the grant EC contract ERBCHRXCT940460 for the project "Stability and universality in classical mechanics", and of the grant CE n. CHRX–CT93–0330/DG, "Order and chaos in conservative dynamical systems".

2. Scheme of the proof

We present here the scheme of the proof of theorem 1.1; in order to simplify the discussion we begin by considering the case $m = 0$.

The first step of the proof consists in putting NLS in resonant Birkhoff normal form (following essentially [30,23,13,4]) up to an exponentially small remainder. Then we observe that the forth order normal form is integrable (averaged system). So, the idea is to study NLS as a perturbation of the nonlinear integrable system given by the forth order normal form and to prove a Nekhoroshev type theorem for it. It also turns out that, when expressed in action angle variable and restricted to a surface of constant L^2 norm, the Hamiltonian of such an integrable system is convex, and this is well known to simplify significantly the proof of Nekhoroshev's theorem[8,24,32].

The rest of the proof (following [1]) is based on a variant of the technique of approximation by periodic orbits introduced by Lochak[24]. Namely, we begin by fixing a completely resonant finite dimensional invariant torus of the averaged system, and we construct a resonant normal form close to such a torus. Then we exploit the normal form together with convexity, with the conservation of the L^2 norm and with the approximate conservation of the harmonic energy (which follows from the first Birkhoff normalization) to show long time stability close to the considered resonant torus (see theorem 6.2 below). Finally we conclude the proof of theorem 1.1 by showing that any initial datum close to a finite dimensional torus is close to a *resonant torus* enough to obtain long time stability.

To actually transform the above scheme into a proof there are some technical difficulties. The first is that one of its main ingredients is the conservation of the L^2 norm; so we have to show that this property is not destroyed by the normal form transformations. This is obtained by proving a normal form lemma for systems with symmetry (see theorem 4.4 below). By the way, the same technique that allows to keep into account symmetry allows also to deal with the case $m \neq 0$.

Some more difficulties are related to the fact that we work here close to an equilibrium point, but the question is very technical, and its discussion is deferred to sects. 5 and 6.

3. Discussion

We begin by discussing the relation of our result with previous results on small amplitude solutions of NLS.

In the case of Dirichlet boundary conditions KAM theory allows to prove that for any finite integer N, there exist Cantor families of N–dimensional invariant tori which accumulate at the origin of the phase space[21,35,23] (see also [22, 33]); moreover, such tori are covered by quasi-periodic motions and have zero Lyapunov exponents. However, N dimensional tori are ensured to exist only in a ball of radius R_N (having fixed an H^s norm) with R_N tending to zero very quickly as $N \to \infty$; so such quasi-periodic motions (although very important) correspond to exceptional initial data. Exploiting the vanishing of Lyapunov exponents one can deduce also informations on the dynamics close to invariant tori; precisely, corresponding to each KAM torus there exists a threshold ϵ_* such that for any $\epsilon < \epsilon_*$ solutions starting $\epsilon/2$ close to the considered torus remain ϵ close to it up to times $1/\epsilon$. However KAM theory gives no information on the size of ϵ_* and on its dependence on the torus, so one has no informations on the size of the portion of the phase space where the dynamics is controlled by KAM tori. We also point out that in the case of periodic boundary conditions the situation is much less clear. Indeed in this case the only KAM–like available results[15,13] give no informations on the Lyapunov exponents of the tori (see however the preliminary result [36]).

In [4] small amplitude solutions of NLS were studied by a suitable averaging method, and long time stability was proved (both for the case of Dirichlet and periodic boundary conditions) for "approximatively non-resonant" initial data. More precisely, corresponding to any N–dimensional Diophantine vector we constructed a continuous family of N dimensional tori (labelled by the energy E) with the property that initial data $\mathcal{O}(E)$ close (in H^1 norm) to the torus of energy E give rise to solutions remaining $\mathcal{O}(E)$ close to it up to times exponentially long with a negative power of E. Moreover, close to the majority of these (continuous) families of tori there is a (Cantor) family of KAM tori. So this result clarify stability properties close to non-resonant tori.

The present paper removes the above non-resonance condition, and indeed stability over exponentially long times is proved for initial data close to *any* finite dimensional small amplitude invariant torus of the linearized system. This is obtained by developing the so called geometric part of Nekhoroshev theorem (as described in the previous section). We think that this should allow to study also some situations more general than a small neighborhood of an equilibrium point. Finally it is shown here that, introducing a suitable space of very smooth functions it is possible to remove also the limitation of being close to finite dimensional tori, obtaining a result valid for *any smooth small amplitude solution*. By the way this could have some relations with some recent works [12, 34] where very smooth almost periodic solutions to some nonlinear PDE's were constructed (and may be with the results of [16, 31]). We point out that, the result of the present paper is obtained by exploiting the conservation of the L^2 norm, which played no role in KAM theory and in [4].

Finally we recall the papers [12] and [3], on long time stability in some PDE's different from (0.1).

In [12] Bourgain proved a Nekhoroshev type result of the kind of my theorem 1.1 (see also corollary 6.4 below) for the equation

$$iu_t + u_{xx} + Vu = \epsilon \frac{\partial H}{\partial u^*}(u, u^*) \ , \quad u(0,t) = u(\pi, t) = 0 \ ,$$

under the assumption that the Dirichlet eigenvalues of $-\partial_{xx} + V$ satisfy a suitable (very general) non-resonance relation. We think that Bourgain's result is an extension to PDE's of the long time stability results close to non-resonant equilibrium points which, in finite dimensional systems, are usually obtained by Birkhoff normal form (see e.g. [18,17]); for a result of this kind in infinite lattices of coupled particles see [10]. On the contrary the present paper (as well as [1,2,4]) is based on the idea that "resonance is stabilizing" (see [6,7,11,24]), and indeed we deal here with a neighborhood of a resonant equilibrium point (which is not covered by Bourgain's result), and we obtain control of the dynamics for more general initial data and times much longer than those considered in [12]. However, we point out that in the present case the control of the dynamics is obtained exploiting also some additional properties of the nonlinearity, in particular we need integrability of some truncated Birkhoff normal form and its convexity. Moreover, the present method does not allow to control higher Sobolev norms.

Also in [3] the neighborhood of a completely resonant equilibrium point was studied. In that paper, under very general assumptions on the nonlinearity, some approximate periodic orbits which are stable over exponentially long times were constructed. Such orbits were also explicitly computed for the nonlinear wave equation. The method of [3] is much more general than the method of the present paper, but it allows to deal only with particular initial data, and moreover (due to the conservation of the L^2 norm) it does not apply directly to NLS.

4. Normal form of a system with approximate symmetry

Consider a real Banach space $\mathcal{P}$ endowed by a weak symplectic form Ω. Then the application $\mathcal{P} \ni X \mapsto \Omega(.,X) \in \mathcal{P}^*$ is an injective linear map ($\mathcal{P}^*$ being the dual of $\mathcal{P}$). We recall that, if it is also surjective then Ω is said to be strongly non-degenerate. Here (as usual in Hamiltonian PDE's) we will deal with weak symplectic forms. Consider now the inverse $J : \mathcal{P}^* \supset D(J) \to \mathcal{P}$ of the above map. We recall that J is called the Poisson tensor, and that in the case of weak symplectic form which are not strongly nondegenerate, its domain $D(J)$ is strictly contained in $\mathcal{P}^*$.

Remark 4.1. *If $\mathcal{P}$ is reflexive, then it is easy to prove*[25] *that $D(J)$ is dense in $\mathcal{P}^*$.*

Given a differentiable function h its Hamiltonian vector field or symplectic gradient is defined by Jdh (when it exists); given two differentiable functions h and h_1 their Poisson bracket, if it exists, is defined by

$$\{h, h_1\}(z) := dh_1(z)Jdh(z) \ .$$

On a set $\mathcal{G} \subset \mathcal{P}$ consider a Hamiltonian system of the form

$$H(z) := h_\omega(z) + \hat{h}(z) + f(z) + f_1(z) + r(z) + g(N_1(z), ..., N_l(z)) \ , \tag{4.1}$$

where h_ω is the unperturbed Hamiltonian which will be assumed to generate a periodic flow, $\hat{h}$ is a part of the Hamiltonian already in normal form and small, f the main part of the perturbation, f_1 is a higher order part of the perturbation, and r a very small remainder (actually in the case of interest it will be exponentially small), whose role will be clear in a while; $N_1, ..., N_l$ are approximate constants of motion, precisely we assume that they commute with h_ω, $\hat{h}$, f and f_1, but not with r. The splitting of the perturbation in three terms, namely f, f_1 and r, is justified by the following remark. It is quite typical in applications that the perturbation naturally appears as composed by a main part, the average of which is easily computed, and a higher order part, which is perhaps even unknown. We denote by f the main part. In our case we have a further splitting of higher order terms into a part that commutes with the functions N_i, and a part which does not commute. We denote these two parts by f_1 and r, respectively. In the NLS case, r is also much smaller than f_1.

We present now in a qualitative way our normal form result. To start with consider the (oversimplified) case $f_1 = r = g = 0$, where (4.1) reduces to

$$H_o := h_\omega + \hat{h} + f \ . \tag{4.2}$$

In this case our result consists of the construction of a canonical transformation $\mathcal{T}$ putting (4.2) in resonant normal form, namely such that

$$H_o \circ \mathcal{T} = h_\omega + \hat{h} + \langle f \rangle + Z + \mathcal{R} \ , \tag{4.3}$$

where $\langle f \rangle$ is the average of f along the unperturbed flow, i.e.

$$\langle f \rangle(z) := \frac{1}{T}\int_0^T f(\Phi^t(z))dt \ , \tag{4.4}$$

where Φ^t is the flow generated by Jdh_ω. The function Z satisfies $\{h_\omega, Z\} = 0$ (i.e. it is in resonant normal form) and is of order higher than $\langle f \rangle$, while $\mathcal{R}$ is a remainder exponentially small with the size of $f + \hat{h}$.

Proceeding step by step, let now $f_1 \neq 0$ be a perturbation of higher order. One easily realizes that its presence only modifies the explicit form of Z, but changes neither its order of magnitude nor the order of magnitude of $\mathcal{R}$, so that the normal form (4.3) remains valid with essentially the same estimates.

Consider now the case where also $g \neq 0$. In view of the assumption that each part of the Hamiltonian commutes with each of the functions N_i's, one is tempted to use the standard Marsden–Weinstein reduction theory (under suitable assumptions) to get rid of the function $g(N_1, \ldots, N_l)$. However this it not convenient for the following reasons. (i) Marsden–Weinstein's reduction theory fails when the flow of some of the N_i's is not differentiable as in the case we will tackle. (ii) At some points of the phase space the reduction may fail even if the flow of all N_i's is regular. (iii) The phase space of the reduced system is in general a manifold with a non trivial topology, so that the development of perturbation theory involves some extra subtleties that are avoided by our approach. (iv) We will consider also the case where the system has some approximate constants of motion. Here one should be able to pull back the normalizing transformation from the reduced phase space to the original one; this could be nontrivial. So (following essentially [14]), we will proceed in a different way: looking carefully at the construction of the normalizing transformation $\mathcal{T}$, one can see that $\mathcal{T}$ itself leaves the functions N_i's invariant, i.e., one has $N_i \circ \mathcal{T} = N_i$. It follows that one can simply forget the function g which, being invariant under $\mathcal{T}$, does not affect the construction of the normal form.

Finally, consider the case where also r is different from 0. It is clear that the normalizing transformation changes it into $r \circ \mathcal{T}$, so that in particular it does not change its size.

The proof of the normal form lemma will be based on the method of Lie transform, i.e. we will construct the normalizing transformation as the composition of a sequence of transformations $\mathcal{T}_1, ..., \mathcal{T}_L$ each one being the time one flow of a suitable Hamiltonian vector field. The number L of steps will be fixed by minimizing the estimate of the remainder, and this will produce the exponential estimate. We illustrate at a formal level the first step.

Given a function $\chi_1 : \mathcal{P} \supset \mathcal{U} \to \mathbb{R}$, ($\mathcal{U} \supset \mathcal{G}$ open) consider the differential equation

$$\dot{z} = Jd\chi_1(z) ,$$

and denote by $\mathcal{T}_1$ the corresponding time one flow. Assume $Jd\chi_1$ small (of the same order as Jdf), then we have

$$\mathcal{T}_1(z) = z + Jd\chi_1(z) + \mathcal{O}(\|Jd\chi_1\|^2) ;$$

assume also $\{\chi_1, N_i\} \equiv 0$, and calculate $H \circ \mathcal{T}_1$; we have

$$\begin{aligned} H(\mathcal{T}_1(z)) = {} & h_\omega(z) + \hat{h}(z) + g(N_1(z), ..., N_l(z)) \\ & + \{\chi_1, h_\omega\}(z) + f(z) + \text{higher order terms} . \end{aligned} \tag{4.5}$$

So we look for a χ_1 such that the first order part of perturbation, i.e. $\{\chi_1, h_\omega\} + f$ is reduced to the average $\langle f \rangle$ of f with respect to the unperturbed flow. Hence, χ_1 has to satisfy

$$\{\chi_1, h_\omega\}(z) + f(z) = \langle f \rangle(z) .$$

It is easily seen that the wanted χ_1 is given by

$$\chi_1(z) := \frac{1}{T}\int_0^T t\left[f(\Phi^t(z)) - \langle f\rangle(\Phi^t(z))\right] dt \tag{4.6}$$

(see lemma 8.4 below), and that, with this definition, one has $\{\chi_1, N_i\} \equiv 0$, so that $N_i \circ \mathcal{T}_1 = N_i$. This completes at a formal level the first step the construction. Then one has to add quantitative estimates, and in particular to show that the "higher order terms" of (4.5) are really small. To this end we have to specify in a better way our assumptions and then to actually perform the estimates. In this section we will give only the precise statement of the normal form lemma, while the technical estimates are deferred to sect. 7.

We introduce now the domain where perturbation theory will be developed. So, consider the complex extension of the set $\mathcal{G}$, by a parameter R, namely define

$$\mathcal{G}_R := \bigcup_{z\in\mathcal{G}} B_R(z)\ ,$$

where $B_R(z) \subset \mathcal{P}^{\mathbf{C}}$ is the (complex) ball of radius R and center z, and $\mathcal{P}^{\mathbf{C}} := \mathcal{P} \oplus i\mathcal{P}$ is the complexification of $\mathcal{P}$. Then we introduce the following family of norms. Let $0 < \rho \le R$ be a parameter, and $h : \mathcal{G}_\rho \to \mathcal{B}$ with $\mathcal{B}$ a Banach space, be an analytic map; we put

$$\left|h\right|_\rho := \sup_{z\in\mathcal{G}_\rho} \left\|h(z)\right\|_{\mathcal{B}}\ . \tag{4.7}$$

If $h : \mathcal{G}_\rho \to \mathbf{C}$ has analytic symplectic gradient Jdh, we put

$$N_\rho\left(h\right) := \sup_{z\in\mathcal{G}_\rho} \left\|Jdh(z)\right\|_{\mathcal{P}}\ . \tag{4.8}$$

We come to the assumptions on the Hamiltonian

0) (Analyticity) There exists $R > 0$ such that h_ω, $\hat{h}$, f, f_1, r and N_i $(i = 1, ..., l)$ can be extended to analytic functions on $\mathcal{G}_R$, and there exist constants E, E_1, E_0, and λ such that

$$|f|_R \le E\ ,\quad |f_1|_R \le E_1\ ,\quad |\hat{h}|_R \le E_0\ ,\quad \left|r\right|_R \le \lambda\ ;$$

moreover $g(N_1(z), ..., N_l(z))$ is analytic on $\mathcal{G}_R$.

Concerning h_ω we assume

1) There exists a set $\mathcal{D}$ dense in $\mathcal{G}_R$ such that the symplectic gradient Jdh_ω exists in $\mathcal{D}$, namely, $dh(z) \in D(J)$ for $z \in \mathcal{D}$.
2) Jdh_ω generates a continuous flow Φ^t which leaves $\mathcal{D}$ and $\mathcal{G}_\rho$ $(\forall\, \rho \le R)$ invariant. Moreover, $\Phi^t(z)$ is time differentiable for $z \in \mathcal{D}$.
3) for any fixed t the map $z \mapsto \Phi^t(z)$ is analytic, and its differential satisfies the estimate

$$\sup_{z\in\mathcal{G}_R} \left\|d\Phi^t(z)\right\|_{\mathcal{P},\mathcal{P}} \le 1\ .$$

4) (Periodicity) There exists T such that the flow Φ^t is periodic with period T, namely

$$\Phi^t = \Phi^{T+t} \quad \forall t \in \mathbb{R} \ .$$

Remark 4.2. *By remark 4.1 one has that if $\mathcal{P}$ is reflexive and dh_ω is non constant then 1) is automatically satisfied.*

Concerning the other parts of the Hamiltonian we assume

5) (Smoothness of the vector fields) $\hat{h}$, f, f_1 and r have symplectic gradient which are analytic from $\mathcal{G}_R$ to $\mathcal{P}$. Moreover there exist constants F, F_0, F_1, Λ such that

$$N_R\left(\hat{h}\right) \leq F_0 \ , \quad N_R\left(f\right) \leq F \ , \quad N_R\left(f_1\right) \leq F_1 \ , \quad N_R\left(r\right) \leq \Lambda \ . \tag{4.9}$$

6) $\hat{h}$ is in normal form, namely

$$\left\{h_\omega, \hat{h}\right\} \equiv 0 \ . \tag{4.10}$$

7) (Symmetry) On $\mathcal{G}_R$ one has

$$\{f, N_i\} \equiv \{f_1, N_i\} \equiv \left\{\hat{h}, N_i\right\} \equiv 0 \ , \ \forall i = 1, ..., l \ , \tag{4.11}$$

and

$$N_i(\Phi^t(z)) = N_i(z) \ , \ \forall t \in \mathbb{R} \ . \tag{4.12}$$

Remark 4.3. *By (5) the Poisson brackets in (4.10) and (4.11) are well defined on $\mathcal{G}_R$; on the contrary the Poisson brackets of N_i with h_ω could be defined only on $\mathcal{D}$. From (4.12) it follows that*

$$\{h_\omega, N_i\}(z) = 0 \ , \quad \forall z \in \mathcal{D} \ .$$

Under the above assumptions we have the following theorem.

Theorem 4.4. *Define*

$$\epsilon := 2^7 e \frac{T}{R}(F + F_0) \ , \quad E^* = \max\left\{\frac{12}{5}\frac{F(E+E_0)}{F+F_0}, \frac{5}{4}E + E_0\right\} \ ;$$

assume $\epsilon < 1/2$, $F_1 \leq \epsilon F$ and $E_1 \leq \epsilon E$ then there exists an analytic canonical transformation $\mathcal{T} : \mathcal{G}_{R/2} \to \mathcal{G}_R$ with the following properties

1.. $\mathcal{T}$ is close to identity and leaves invariant N_i, namely it satisfies

$$\| \mathbb{1} - \mathcal{T}\|_{R/2} \leq 2TF \ , \qquad N_i \circ \mathcal{T} = N_i \ , \quad \forall i = 1, ..., l \tag{4.13}$$

2. *$\mathcal{T}$ puts the Hamiltonian in resonant normal form up to an exponentially small remainder, namely one has*

$$H(\mathcal{T}(z)) = h_\omega(z) + \hat{h}(z) + \langle f\rangle(z) + Z(z) + \mathcal{R}(z) + g(N_1(z), ..., N_l(z)) + \bar{r}(z) \ , \quad (4.14)$$

where

2.a $\langle f\rangle$ is the average (4.4) of f with respect to the unperturbed flow.

2.b Z is in normal form, namely $\{Z, h_\omega\} \equiv 0$, and of order higher than $\langle f\rangle$, namely it is estimated by

$$N_{R/2}\left(Z\right) \le 4\epsilon F \ , \quad |Z|_{R/2} \le 2E^*\epsilon \ , \quad (4.15)$$

2.c $\mathcal{R}$ is an exponentially small remainder estimated by

$$N_{R/2}\left(\mathcal{R}\right) \le 2e\epsilon F \exp\left(-\frac{1}{\epsilon}\right) \ , \quad |\mathcal{R}|_{R/2} \le 2e\epsilon E^* \exp\left(-\frac{1}{\epsilon}\right) \ ,$$

2.d $\bar{r} := r \circ \mathcal{T}$ is estimated by

$$N_{R/2}\left(\bar{r}\right) \le 4\Lambda \ , \quad |\bar{r}|_{R/2} \le \lambda \ .$$

2.e the transformed Hamiltonian is approximatively symmetric, namely one has

$$\{\langle f\rangle, N_i\} = \{Z, N_i\} = \{\mathcal{R}, N_i\} = 0 \quad , \forall i = 1, ..., l.$$

The proof of this theorem (whose statement is close to theorem 4.1 of [1]) is obtained by adapting to the infinite dimensional case with symmetry the proof of the main theorem of [9]; it will be given in detail in sect. 7.

5. Normal form close to completely resonant tori of nonlinear Schrödinger equation

We study here the nonlinear Schrödinger equation (0.1) in the phase space $\mathcal{P} := H_0^1([0,\pi], \mathbf{C})$ (which will be considered as a *real* Hilbert space) endowed by the symplectic form

$$\Omega(u, v) := \mathrm{Re} \int_0^\pi u i v^* dx \ , \quad (5.1)$$

and by the scalar product

$$\langle u, v\rangle := \frac{1}{2} \mathrm{Re} \int_0^\pi u_x^* v_x dx \ .$$

Canonical coordinates (p_k, q_k) are given by the real and the imaginary part of the k–th Fourier component of u, namely

$$p_k := \mathrm{Re}\, u_k \ , \quad q_k := \mathrm{Im}\, u_k \ .$$

The Hamiltonian function of (0.1) is given by

$$H(u) := \frac{1}{2}\int_0^\pi |u_x|^2 \, dx + \frac{1}{2} m \int_0^\pi |u|^2 dx + \int_0^\pi \Psi(|u|^2) dx \ , \tag{5.2}$$

where Ψ is such that $\Psi' = \varphi$, $\Psi(0) = 0$ and $\Psi'(0) = 0$.

It is useful to expand Ψ in Taylor series, namely to write

$$\Psi\,(y) = \sum_{k\geq 2} \Psi_k y^k \ ,$$

and moreover to rescale u in such a way that $\Psi_2 = \pi/4$. We will use the notations

$$h_0(u) := \frac{1}{2}\int_0^\pi |u_x|^2 \, dx = \sum_{k\geq 1} k^2 \frac{p_k^2 + q_k^2}{2} \ , \quad f_{(0)}(u) := \frac{\pi}{4}\int_0^\pi |u|^4 \, dx \ ,$$
$$f_{(1)}(u) := \sum_{i\geq 3} \Psi_i \int_0^\pi |u|^{2i} dx \ , \quad N(u) := \frac{1}{2}\int_0^\pi |u|^2 dx \ ,$$

so that (5.2) can be rewritten as

$$H(u) = h_0(u) + f_{(0)}(u) + f_{(1)}(u) + mN(u) \ , \tag{5.3}$$

which is of the form (4.1) with $h_\omega := h_0$, $f := f_{(0)}$, $f_1 := f_{(1)}$, $g := mN$, $\hat{h} \equiv 0$, $r \equiv 0$. To apply theorem 4.4, we define the set $\mathcal{G}$ to be the origin of the phase space, and extend it to a complex domain $\mathcal{G}^0_{R_0}$ using a parameter R_0, which will be eventually related to the norm of the initial datum, namely we put

$$\mathcal{G}^0_{R_0} := \left\{ u \in H^1_0 \ : \ \|u\|_{H^1} < R_0 \right\} \ .$$

It is clear from Sobolev embedding theorem that, provided R_0 is small enough, the different parts of the Hamiltonian are analytic on $\mathcal{G}^0_{R_0}$, and that the vector fields of $f_{(0)}$ and of $f_{(1)}$ are analytic on the same domain; moreover there exists a constant C_{12} such that

$$\left|f_{(0)}\right|_{R_0} \leq C_{12} R_0^4 \ , \quad \left|f_{(1)}\right|_{R_0} \leq C_{12} R_0^6 \ ,$$
$$N_{R_0}\left(f_{(0)}\right) \leq C_{12} R_0^3 \ , \quad N_{R_0}\left(f_{(1)}\right) \leq C_{12} R_0^5 \ .$$

It is easy to check that also the other assumptions of theorem 4.4 are fulfilled, so that it can be applied giving the forthcoming theorem 5.1. In its statement and below *we will use the notation $a \preceq b$ to mean "there exists a constant C dependent only on φ, in particular* *independent of R_0, on the period T and on the domain, such that $a \leq Cb$", and similarly we will write $a \asymp b$ to mean "there exists a constant C with the above property such that $a = Cb$". The same properties are fulfilled by all constants whose existence is states in the propositions and theorems below.*

Theorem 5.1. *Consider the Hamiltonian system (5.2); there exists a positive $R_\sharp$ and an analytic canonical transformation $\mathcal{T}_{(0)} : \mathcal{G}^0_{R_\sharp/2} \to \mathcal{G}^0_{R_\sharp}$, with the following properties. For any positive $R_0 \leq R_\sharp$ one has*

1. *$\mathcal{T}_{(0)}$ leaves invariant N, namely $N \circ \mathcal{T}_{(0)} = N$, and satisfies*

$$N_{R_0/2}\left(1\!\!1 - \mathcal{T}_{(0)} \right) \preceq R_0^3 \ ; \tag{5.4}$$

the same estimate holds for $\mathcal{T}^{-1}_{(0)}$.

2. *$\mathcal{T}_{(0)}$ puts the Hamiltonian (5.3) in resonant normal form, namely one has*

$$H_{(1)} := H \circ \mathcal{T}_{(0)} = h_0 + \langle f_{(0)} \rangle + Z_{(0)} + \mathcal{R}_0 + mN \tag{5.5}$$

where

2a. *$\langle f_{(0)} \rangle$ is the average of $f_{(0)}$ with respect to the unperturbed flow, which is explicitly given by*

$$\langle f_{(0)} \rangle (u) := \frac{1}{2}\Big(N(u) \Big)^2 - \frac{1}{8} \sum_{k \geq 1} |u_k|^4 \ ; \tag{5.6}$$

2b. *$Z_{(0)}$ is in normal form, i.e. it satisfies $\left\{ h_0, Z_{(0)} \right\} = 0$ and commutes with N, namely $\left\{ N, Z_{(0)} \right\} = 0$; it is analytic on $\mathcal{G}^0_{R_0/2}$ together with its symplectic gradient, and it satisfies the estimates*

$$N_{R_0/2}\left(Z_{(0)} \right) \preceq R_0^5 \ , \quad \left| Z_{(0)} \right|_{R_0/2} \preceq R_0^6 \ . \tag{5.7}$$

2c. *$\mathcal{R}_0$ is a remainder analytic on $\mathcal{G}^0_{R_0/2}$ together with its symplectic gradient; it satisfies $\{ \mathcal{R}_0, N \} = 0$ and the estimates*

$$N_{R_0/2}\left(\mathcal{R}_0 \right) \preceq R_0^5 \exp\left[- \left(\frac{R_\sharp}{R_0} \right)^2 \right] \ , \quad |\mathcal{R}_0|_{R_0/2} \preceq R_0^6 \exp\left[- \left(\frac{R_\sharp}{R_0} \right)^2 \right] \ . \tag{5.8}$$

Proof. The statement is an immediate consequence of theorem 4.4. The only point to be proved is formula (5.6) which gives the average of the main part of the perturbation. This formula can be obtained by explicit calculation exactly as in [23], proof of lemma 5 (see also [4]). So the proof is omitted. $\square$

Remark 5.2. *By (5.4) and the mean value theorem one has*

$$\left| \sum_{k \geq 1} k^2 Z_k - \sum_{k \geq 1} (k^2 Z_k \circ \mathcal{T}_{(0)}) \right|_{3R_0/8} \preceq R_0^4 \ . \tag{5.9}$$

We prepare now the Hamiltonian for a further application of theorem 4.4, in which h_0 will be considered as an approximate constant of motion. We aim to obtain a resonant normal form close to a periodic torus of the integrable system with Hamiltonian

$$\frac{1}{8}\sum_{k\geq 1}|u_k|^4 = \frac{1}{2}\sum_{k\geq 1} {Z_k}^2 \ .$$

We fix a finite set of indices $\mathcal{I}\subset \mathbf{N}\setminus\{0\}$, with cardinality N and introduce action angle variables for the oscillators labelled by the indices in $\mathcal{I}$. Namely we define new canonical variables $(Z_k,\phi_k)_{k\in\mathcal{I}}$ by

$$p_k = \sqrt{2Z_k}\cos\phi_k \ , \quad q_k = \sqrt{2Z_k}\sin\phi_k \ , \quad k\in\mathcal{I} \ , \tag{5.10}$$

and we retain the variables p,q for all the other oscillators. This is an analytic canonical transformation outside of the set $\{Z_k = 0 \text{ for some } k\in\mathcal{I}\}$.

To construct the resonant torus around which we will normalize the system we fix a real positive ν, and a (finite) collection of strictly positive integer numbers $\{\ell_k\}_{k\in\mathcal{I}}$. Correspondingly we define a completely resonant frequency vector $(\nu_1,\nu_2,...)$ by

$$\nu_k := \begin{cases} \nu\ell_k & \text{if} \quad k\in\mathcal{I} \\ 0 & \text{if} \quad k\notin\mathcal{I} \end{cases} .$$

We will assume also

$$\sum_{k\geq 1} k^2\nu_k \leq \frac{R_0^2}{e^2 8} \ . \tag{5.11}$$

Then we translate the origin of the actions defining

$$I_k := Z_k - \nu_k \ . \tag{5.12}$$

From now on we will use the canonical coordinates (I,ϕ,p,q); we will denote by

$$z := (I,\phi,p,q)$$

the collection of all the coordinates and by

$$\mathcal{P} := (\,\mathrm{I\!R}_+)^N \times \mathrm{I\!R}^N \times \ell^2_1 \times \ell^2_1 \ni z$$

the phase space; here we denoted by ℓ^2_1 the space of the sequences $\{p_k\}_{k\notin\mathcal{I}}$ such that kp_k is square summable. We will work around the torus $I=p=q=0$. To rewrite the Hamiltonian in the new variables it is useful to introduce the auxiliary functions J_k defined by

$$J_k(z) := \begin{cases} Z_k = \frac{p_k^2+q_k^2}{2} & \text{if} \quad k\notin\mathcal{I} \\ I_k & \text{if} \quad k\in\mathcal{I} \end{cases} ,$$

which are the translated actions. With these notations the Hamiltonian (5.5) takes the form

$$H_{(1)} := h_1 - \hat{h} + Z_{(0)} + \mathcal{R}_0 + \left[mN + h_0 + NC_\nu + \frac{1}{2}N^2 \right] \ . \tag{5.13}$$

where

$$h_1(z) := -\sum_{k\in\mathcal{I}} \nu_k I_k \ , \quad \hat{h} := \frac{1}{2}\sum_{k\geq 1} J_k^2 \ , \quad h_0 := \sum_{k\geq 1} k^2 J_k \ ,$$
$$N := \sum_{k\geq 1} J_k \ , \quad C_\nu := \sum_{k\geq 1} \nu_k \ ,$$

and $Z_{(0)}$, $\mathcal{R}_0$ are the same functions as in (5.5) but calculated in the new variables.

We construct now the domain in which theorem 4.4 will be applied again. To this end fix a positive constant $C_R > 1$ and a positive quantity R_1 satisfying

$$C_R R_0^{3/2} \leq R_1 \leq \frac{R_0}{10}. \tag{5.14}$$

Then we introduce a new R_1–dependent norm on the phase space, by

$$\|z\|_{\mathcal{P}} = \|(I,\phi,p,q)\|_{\mathcal{P}} := \sum_{i\in\mathcal{I}} \frac{i^2|I_i|}{R_1} + \sup_{i\in\mathcal{I}} R_1|\phi_i| + \sqrt{\sum_{i\notin\mathcal{I}} i^2 \frac{|p_i|^2+|q_i|^2}{2}} \ , \tag{5.15}$$

which is equivalent to the H^1 norm. *From now on we will use only this new norm, except when explicitly specified.* Then consider the torus

$$\mathcal{G}^\nu := \{ z\in\mathcal{P} \ : \ I_i = p_l = q_l = 0 \ \forall i\in\mathcal{I}, \ l\notin\mathcal{I} \} \ ,$$

and extend it to the complex using the balls of radius $2R_1$ in the new norm obtaining the domain $\mathcal{G}^\nu_{2R_1}$.

Remark 5.3. *By lemma 8.7 below (which exploits (5.11) and (5.14)) one has $\mathcal{G}^\nu_{2R_1} \subset \mathcal{G}^0_{\frac{R}{2}}$; moreover, provided*

$$\min_{i\in\mathcal{I}} \{|\nu_i|\} > (2R_1)^2 \ , \tag{5.16}$$

one has that $\mathcal{G}^\nu_{2R_1}$ does not intersect the set $\{Z_i = 0$ for some $i\in\mathcal{I}\}$, and therefore $H_{(1)}$ as given by (5.13) is analytic on $\mathcal{G}^\nu_{2R_1}$. Notice also that (5.16) is implied by

$$\nu > 4R_1^2 \ . \tag{5.17}$$

Corresponding to the new norm and the new domain we will use the notations

$$N^\nu_\rho(h) := \sup_{z\in\mathcal{G}^\nu_\rho} \|Jdh(z)\|_{\mathcal{P}} \ , \quad |h|^\nu_\rho := \sup_{z\in\mathcal{G}^\nu_\rho} |h(z)| \ .$$

We point out that according to the definition of the symbols $\preceq$ and $\asymp$, the constants realizing the corresponding relations are independent of the domain; in particular they are independent of $\mathcal{I}$, of $\mathcal{C}(\mathcal{I})$, of $N := \#\mathcal{I}$ and of the constant C_R (the same holds for the constants whose existence is ensured by the forthcoming theorems).

We have now

Theorem 5.4. *There exists a positive constant $\mu_\sharp$ such that, if*

$$\mu := \frac{1}{\nu}\left(R_1^2 + \frac{R_0^6}{R_1^2}\right) < \mu_\sharp \ , \tag{5.18}$$

then there exists an analytic canonical transformation $\mathcal{T}_{(1)} : \mathcal{G}^\nu_{R_1/2} \to \mathcal{G}^\nu_{R_1}$ with the following properties

1) it is close to the identity, namely it satisfies

$$\left| \mathbb{1} - \mathcal{T}_{(1)} \right|_{R_1/2} \preceq \frac{R_0^6}{\nu R_1} \, (< R_1 \mu) \ , \tag{5.19}$$

and leaves invariant N and h_0, namely one has $h_0 \circ \mathcal{T}_{(1)} = h_0$ and $N \circ \mathcal{T}_{(1)} = N$

2) it puts (5.13) in normal form, namely one has

$$H_{(1)} \circ \mathcal{T}_{(1)} = h_1 - \hat{h} + Z_{(1)} + \mathcal{R}_1 + \mathcal{R}_0 + \left[mN + h_0 + NC_\nu + \frac{1}{2}N^2 \right] \ , \tag{5.20}$$

where $Z_{(1)}$ is a normal form term commuting with h_0, and h_1, namely

$$\left\{h_0, Z_{(1)}\right\} = \left\{h_1, Z_{(1)}\right\} = 0 \ ;$$

$\mathcal{R}_1$ is a remainder commuting with h_0, and each term of the Hamiltonian commutes with N; moreover the following estimates hold

$$\begin{aligned}
N^\nu_{R_1/2}(Z_{(1)}) &\preceq \frac{R_0^6}{R_1} \ , & \left| Z_{(1)} \right|^\nu_{R_1/2} &\preceq R_0^6 \ , \\
N^\nu_{R_1/2}(\mathcal{R}_1) &\preceq \frac{R_0^6}{R_1} \mu \exp\left[-\left(\frac{\mu_\sharp}{\mu} \right) \right] \ , & \left| \mathcal{R}_1 \right|^\nu_{R_1/2} &\preceq R_0^6 \mu \exp\left[-\left(\frac{\mu_\sharp}{\mu} \right) \right] \ , \\
N^\nu_{R_1/2}(\mathcal{R}_0) &\preceq \frac{R_0^6}{R_1} \exp\left[-\left(\frac{R_\sharp}{R_0} \right)^2 \right] \ , & \left| \mathcal{R}_0 \right|^\nu_{R_1/2} &\preceq R_0^6 \exp\left[-\left(\frac{R_\sharp}{R_0} \right)^2 \right] \ .
\end{aligned} \tag{5.21}$$

Proof. We want to apply theorem 4.4 with $h_\omega := h_1$, $\hat{h}$ as defined above, $f := Z_{(0)}$, $f_1 \equiv 0$, $r := \mathcal{R}_0$, the approximate constants of motion given by $N_1 := N$, $N_2 := h_0$, $g(N_1, N_2) := mN + h_0 + NC_\nu + \frac{1}{2}N^2$, and the domain defined by $\mathcal{G}_R := \mathcal{G}^\nu_{R_1}$. We begin by verifying the analyticity assumptions. To this end remark that, provided $\mu_\sharp$ is smaller than

$1/4$, by (5.18) one has that eq. (5.17) is automatically satisfies and therefore the Hamiltonian is analytic in $\mathcal{G}^{\nu}_{2R_1}$. So assumption (0) of the theorem holds with

$$E \asymp R_0^6 \ , \quad \lambda \asymp R_0^6 \exp\left[-\left(\frac{R_\sharp}{R}\right)^2\right] \ ,$$

and $E_0 = R_1^4$; indeed, by lemma 8.7 one has $\mathcal{G}^{\nu}_{R_1} \subset \mathcal{G}^{0}_{R_0}$, and

$$\left|\hat{h}\right|^{\nu}_{R_1} \leq \frac{1}{2}\sum_{i\geq 1}\left|J_i^2\right| < \frac{1}{2}\left[\sum_{i\in\mathcal{I}}\left|i^2 I_i\right|^2 + \sum_{i\notin\mathcal{I}}\left(i^2\frac{|p_i|^2+|q_i|^2}{2}\right)^2\right] \leq \frac{1}{2}R_1^4 + \frac{1}{2}R_1^4 \ .$$

Assumptions (1–4) are satisfied, and one has $T = \frac{2\pi}{\nu}$. The analyticity assumptions in (5) on the vector fields are satisfied by the equivalence of the norm (5.15) and the H^1 norm. We need now to estimate *in the new norm* the vector fields of the perturbation. This is obtained using lemma 8.8, eqs. (5.7), (5.8) and (5.14), from which we have

$$F \asymp \frac{R_0^6}{R_1} \ , \quad \Lambda \asymp \frac{R_0^6}{R_1}\exp\left[-\left(\frac{R_\sharp}{R_0}\right)^2\right] \ .$$

To obtain F_0 we must estimate $Jd\hat{h}$. We have (in $\mathcal{G}^{\nu}_{R_1}$)

$$\left|\left(Jd\hat{h}\right)_{\phi_i}\right| = \left|\frac{\partial\hat{h}}{\partial I_i}\right| = |I_i| \leq R_1^2$$

$$\left\|(Jd\hat{h})_{pq}\right\|^2 \leq \sum_{k\notin\mathcal{I}}|J_k|^2\left(|p_k|^2+|q_k|^2\right) \leq 2R_1^4\sum_{k\notin\mathcal{I}}\frac{|p_k|^2+|q_k|^2}{2} < 2R_1^6 \ ,$$

from which we can choose $F_0 \asymp R_1^3$. Also conditions (6–7) are satisfied. Substituting the above values in the definition of the quantities ϵ and E^* involved in theorem 4.4, we get

$$\epsilon \asymp \frac{1}{\nu}\left(R_1^2 + \frac{R_0^6}{R_1^2}\right) \ , \quad E^* \asymp R_0^6 \ ,$$

so that (5.18) with a suitable $\mu_\sharp$ is the smallness condition of theorem 4.4, and so the application of such a theorem gives the result. □

Remark 5.5. *The term $\mathcal{R}_0$ of the remainder is much smaller than $\mathcal{R}_1$; indeed, by (5.11), one has that ν is of order R_0^2, and therefore, taking into account (5.14) one has that μ is of order smaller than R_0 while the argument in the exponential estimate of $\mathcal{R}_0$ is proportional to R_0^{-2}.*

Remark 5.6. *By (5.19) the mean value theorem and the Cauchy estimate one has*

$$|\hat{h} - \hat{h}\circ\mathcal{T}_{(1)}|_{3R_1/8} \preceq \frac{R_1^2 R_0^6}{\nu}\left(< R_1^4\frac{R_1^2}{\nu}\right) \ . \tag{5.22}$$

6. Stability in nonlinear Schrödinger equation.

From theorem 5.4 one can conclude that in the domain $\mathcal{G}^{\nu}_{R_1/2}$ the time derivatives of the quantities h_0, h_1 and $\hat{h}-Z_{(1)}$ are exponentially small. Moreover, since $|Z_{(1)}|$ is small, $\hat{h}$ is an approximate constant of motion up to exponentially long times. The functions $\hat{h}$ and h_0 will be used as Lyapunov functions to bound the diffusion of the actions. To this end the following lemma is useful

Lemma 6.1. *Assume that $z\in\mathcal{P}$ (i.e. it is real), and that for some $\rho,\rho_1>0$ one has*

$$h_0(z)=\sum_{k\geq 1}k^2J_k(z)\leq\rho^2\ ,\quad \hat{h}(z)=\frac{1}{2}\sum_{k\geq 1}J_k^2(z)\leq\frac{1}{2}\rho_1^4\ , \tag{6.1}$$

then

$$\sum_{k\geq 1}k^2|J_k|\leq 2\mathcal{C}(\mathcal{I})\rho_1^2+\rho^2\ , \tag{6.2}$$

where $\mathcal{C}(\mathcal{I}):=\sum_{i\in\mathcal{I}}i^2$, moreover one has $z\in\mathcal{G}^{\nu}_d$ with

$$d=\mathcal{C}(\mathcal{I})\frac{\rho_1^2}{R_1}+\sqrt{\rho^2+\rho_1^2\mathcal{C}(\mathcal{I})}\ . \tag{6.3}$$

Proof. From the second of (6.1) one has $|I_k|\leq\rho_1^2$ for any $k\in\mathcal{I}$. Substituting in the first of (6.1), one gets

$$\sum_{i\notin\mathcal{I}}i^2J_i\leq\rho^2+\rho_1^2\mathcal{C}(\mathcal{I})\ ,$$

from which (6.2) immediately follows. To prove the second assertion just use the definition of the norm and of $\mathcal{G}^{\nu}_d$. □

It is now easy to prove stability close to the periodic tori of the averaged system; in the statement of the corresponding theorem we will denote

$$\beta:=\frac{1}{8\mathcal{C}(\mathcal{I})}\ . \tag{6.4}$$

Theorem 6.2. *There exists a positive constant Δ, such that, if the initial datum z_0 is real and belongs to $\mathcal{G}^{\nu}_{\beta R_1}$, if μ is small enough (i.e. (5.18) holds) and if the constant C_R in (5.14) satisfies*

$$C_R\geq\Delta\sqrt{\mathcal{C}(\mathcal{I})}\ , \tag{6.5}$$

then for

$$|t|\preceq\frac{1}{\nu\mu^2}\exp\left(\frac{\mu_\sharp}{\mu}\right)\ , \tag{6.6}$$

one has that the solution $z(t)$ of (5.13) satisfies $z(t)\in\mathcal{G}^{\nu}_{R_1/4}$ and

$$\sum_{i\geq1} i^2\left|J_i(z(t))\right|\leq\frac{R_1^2}{72\mathcal{C}(\mathcal{I})}\ . \tag{6.7}$$

Remark 6.3. *If one denotes by z' the new variables introduced by the transformation $\mathcal{T}_{(1)}$, namely $z=\mathcal{T}_{(1)}(z')$, then theorem 6.2 is formulated in terms of the old variables z. In particular the distance of the initial datum from the torus $\mathcal{G}^{\nu}$ has to be computed in such (simpler) variables.*

Proof. The proof is obtained by applying theorem 5.4 and lemma 6.1. First we assume $z_0\in\mathcal{G}^{\nu}_{R_1\beta}\cap\mathcal{P}$ with a still undetermined β. Denote $z_0':=\mathcal{T}_{(1)}^{-1}(z_0)$. Denote by C_{13} the constant realizing (5.19), then, provided $C_{13}\mu+\beta<1/4$, one has $z_0'\in\mathcal{G}^{\nu}_{R_1/4}$.

We study now the dynamics of $\hat{h}$. Exploiting the definition of $\hat{h}$ one has

$$|\hat{h}(z_0)|\leq\frac{1}{2}R_1^4\beta^2(1+\beta^2)\ ,$$

from which, by (5.22),

$$|\hat{h}(z_0')|\leq\frac{1}{2}R_1^4\beta^2(1+\beta^2)+C_{14}\frac{R_1^2R_0^6}{\nu}\ ,$$

where C_{14} is the constant realizing (5.22). To obtain informations on $\hat{h}(t)$ we use the fact that the evolution of $\hat{h}_1:=-\hat{h}+Z_{(1)}$ is exponentially slow (see (6.9) below). We first need informations on $\hat{h}_1(z_0')$, this is given by

$$|\hat{h}_1(z_0')|\leq\frac{1}{2}R_1^4\beta^2(1+\beta^2)+C_{14}\frac{R_1^2R_0^6}{\nu}+C_{15}R_0^6\leq\frac{1}{2}R_1^4\beta^2(1+\beta^2)+2C_{15}R_0^6\ , \tag{6.8}$$

where C_{15} is a constant which exists by the estimate (5.21) of $Z_{(1)}$, and we assumed $C_{15}/C_{14}>\mu(>R_1^2/\nu)$. Now one has

$$\begin{aligned}\left|\frac{d}{dt}\hat{h}_1\right|^{\nu}_{R_1/4}&=\left|\left\{\hat{h}_1,\mathcal{R}_1+\mathcal{R}_0\right\}\right|^{\nu}_{R_1/4}=\left|d\hat{h}_1(Jd\mathcal{R}_1+Jd\mathcal{R}_0)\right|^{\nu}_{R_1/4}\\&\leq\frac{4}{R_1}\left|\hat{h}_1\right|^{\nu}_{R_1/2}\left(N^{\nu}_{R_1/4}(\mathcal{R}_1)+N^{\nu}_{R_1/4}(\mathcal{R}_0)\right)\\&\preceq\left(R_1^4+R_0^6\right)\frac{R_0^6}{R_1^2}\mu\exp\left[-\left(\frac{\mu_\sharp}{\mu}\right)\right]=R_0^6\nu\mu^2\exp\left[-\left(\frac{\mu_\sharp}{\mu}\right)\right]\\\left|\dot{h}_0\right|^{\nu}_{R_1/4}&\preceq R_0^5\exp\left[-\left(\frac{R_\sharp}{R_0}\right)^2\right]\ ,\end{aligned} \tag{6.9}$$

where we assumed that R_1 is so small that

$$\mu \exp\left[-\left(\frac{\mu_\sharp}{\mu}\right)\right] \succeq \exp\left[-\left(\frac{R_\sharp}{R_0}\right)^2\right] ,$$

which is possible by remark 5.5. Then, using (6.6) one has

$$|\hat{h}_1(z'(t))| \le \frac{1}{2} R_1^4 \beta^2 (1+\beta^2) + 3C_{15} R_0^6 , \tag{6.10}$$

provided $z'(t) \in \mathcal{G}^\nu_{R_1/4}$. From (6.10) we have

$$\left|\hat{h}(z'(t))\right| \le \left|\hat{h}_1(z'(t))\right| + \left|Z_{(1)}(z'(t))\right| \le \frac{1}{2} R_1^4 \beta^2 (1+\beta^2) + 4C_{15} R_0^6 , \tag{6.11}$$

and therefore

$$\left|\hat{h}(z(t))\right| \le \frac{1}{2} R_1^4 \beta^2 (1+\beta^2) + 5C_{15} R_0^6 \le \left(\frac{1}{2}\beta^2(1+\beta^2) + \frac{5C_{15}}{C_R^4}\right) R_1^4 , \tag{6.12}$$

where we took into account (5.14); from this, provided

$$C_R^4 > \frac{10 C_{15}}{\beta^2(1+\beta^2)} \tag{6.13}$$

one has

$$\left|\hat{h}(z(t))\right| \le R_1^4 \beta^2 (1+\beta^2) , \tag{6.14}$$

and also

$$\left|\hat{h}(z'(t))\right| \le R_1^4 \beta^2 (1+\beta^2) . \tag{6.15}$$

Eq. (6.13) will give rise to (6.5). Concerning h_0 we have

$$|h_0(z_0)| = |h_0(z_0')| \le \beta R_1^2 (1+\beta) ,$$

from which, by remark 5.5 and (6.9) one has

$$|h_0(z(t))| = |h_0(z'(t))| \le 2\beta R_1^2 (1+\beta) . \tag{6.16}$$

Using lemma 6.1, (6.15), and (6.16) we obtain that provided β satisfies

$$\sqrt{2\beta}\left[\mathcal{C}(\mathcal{I})\sqrt{\beta(1+\beta^2)} + \sqrt{1+\beta+\mathcal{C}(\mathcal{I})\sqrt{1+\beta^2}}\right] < \frac{1}{4} \tag{6.17}$$

$z'(t)$ belongs to $\mathcal{G}^\nu_{R_1/4}$. Since $\beta < 1$, one has that the l.h.s. of (6.17) is less than

$$\sqrt{2\beta}\left[\sqrt{2\beta}\,\mathcal{C}(\mathcal{I}) + \sqrt{2+2\mathcal{C}(\mathcal{I})}\right] < 2\beta\mathcal{C}(\mathcal{I}) + \sqrt{8\beta\mathcal{C}(\mathcal{I})} ,$$

so, (6.17) is satisfied by $\beta = [8\mathcal{C}(\mathcal{I})]^{-1}$. From lemma 6.1, (6.14), (6.16) one gets also $z(t) \in \mathcal{G}^{\nu}_{R_1/4}$, and the estimate (6.7). Finally we define Δ by $\Delta^4 := 10C_{15}2^6$, and so (6.13) becomes (6.5). $\square$

Proof of theorem 1.1

Define $R_0 := 16\varepsilon$, so that $u^0 \in \mathcal{G}^0_{R_0/16}$; define $v^0 := \mathcal{T}^{-1}_{(0)}(u^0)$, where $\mathcal{T}_{(0)}$ is the transformation defined by theorem 5.1, and, for $i \in \mathcal{J}$, define

$$\omega_i := |v^0_i|^2/2 = Z_i(v^0) \ .$$

We approximate ω_i by a completely resonant vector. To this end we use Dirichlet theorem according to which, for any $Q \geq 1$ there exists an integer vector ℓ_i, $i \in \mathcal{J}$ and a positive integer number $q \leq Q$ such that

$$\left|\omega_i - \frac{\omega \ell_i}{q}\right| \leq \frac{\omega}{qQ^{1/(n-1)}} \ , \quad \forall i \in \mathcal{J} \ , \tag{6.18}$$

where

$$\omega := \max_i \omega_i \ .$$

Let $\mathcal{I} \subset \mathcal{J}$ be the set of the i's such that $\ell_i \neq 0$. We wont to apply theorem 6.2 with $\nu_i = \omega\ell_i/q$, $i \in \mathcal{I}$ and $\nu := \omega/q$. To this end we must define R_1 in such a way that the initial datum is in $\mathcal{G}^{\nu}_{\beta R_1}$ (with $\mathcal{G}^{\nu}$ defined as the torus on which the actions coincide with ν_i $i \in \mathcal{I}$ and vanish otherwise), and verify the hypotheses of the theorem, namely that μ is small. We must also verify that (5.11), (5.14) and (6.5) hold.

Decompose v^0 as follows

$$v^1(x) := \sum_{k\in\mathcal{J}} v^0_k \sqrt{\frac{2}{\pi}} \sin(kx) \ , \quad v^2(x) := v^0(x) - v^1(x) \ ;$$

We denote by

$$z^1 = (\{I_i, \phi_i\}_{i\in\mathcal{I}}, \{p^1_i, q^1_i\}_{i\in\mathcal{J}\setminus\mathcal{I}})$$

the coordinates of v^1 as defined by (5.10), (5.12), and similarly by $z^2 := \{p_i, q_i\}_{i\notin\mathcal{J}}$ the coordinates of v^2. We first define R_1 in such a way that the r.h.s. of (6.18) is of order R_1^2. Precisely, we put

$$R_1 := C_R \sqrt{\frac{\omega}{qQ^{1/(n-1)}}} \ , \tag{6.19}$$

where C_R and Q are still undetermined. We fix now C_R by imposing the condition $z^1 \in$

$\mathcal{G}^{\nu}_{\beta R_1/2}$. To this end we estimate the distance of z^1 from $\mathcal{G}^{\nu}$ by

$$\begin{aligned}\left\|z^1\right\| &= \sum_{i\in\mathcal{I}} i^2\frac{|I_i|}{R_1}+\sqrt{\sum_{i\in\mathcal{J}\setminus\mathcal{I}} i^2\frac{(p_i^1)^2+(q_i^1)^2}{2}}\\ &= \sum_{i\in\mathcal{I}} i^2\frac{|\omega_i-\nu_i|}{R_1}+\sqrt{\sum_{i\in\mathcal{J}\setminus\mathcal{I}} i^2\left|\frac{(p_i^1)^2+(q_i^1)^2}{2}-\nu_i\right|}\\ &\le \mathcal{C}(\mathcal{I})\frac{\omega}{R_1}\frac{1}{qQ^{1/(n-1)}}+\sqrt{\sum_{i\in\mathcal{J}\setminus\mathcal{I}} i^2\frac{\omega}{qQ^{1/(n-1)}}}\\ &\le \mathcal{C}(\mathcal{I})\frac{R_1}{C_R^2}+\frac{R_1}{C_R}\sqrt{\mathcal{C}(\mathcal{J}\setminus\mathcal{I})}\ ,\end{aligned}\tag{6.20}$$

and this is less than $\beta R_1/2$ provided

$$C_R>\frac{\sqrt{\mathcal{C}(\mathcal{J}\setminus\mathcal{I})}+\sqrt{\mathcal{C}(\mathcal{J}\setminus\mathcal{I})+2\beta\mathcal{C}(\mathcal{I})}}{\beta}\ ,$$

which is implied by

$$C_R>16\mathcal{C}(\mathcal{J})^{3/2}\ .$$

So we fix now

$$C_R:=\Delta_1\mathcal{C}(\mathcal{J})^{3/2}\ ,\tag{6.21}$$

where $\Delta_1:=\max\{\Delta,16\}$, so that also (6.5) is satisfied, and $z^1\in\mathcal{G}^{\nu}_{\beta R_1/2}$.

We now fix Q in order to minimize the value of μ. So, we estimate μ:

$$\mu=\frac{q}{\omega}\left(\frac{C_R^2\omega}{qQ^{1/(n-1)}}+\frac{R_0^6qQ^{1/(n-1)}}{C_R^2\omega}\right)\le\frac{C_R^2}{Q^{1/(n-1)}}+\frac{R_0^6Q^{\frac{2n-1}{n-1}}}{\omega^2C_R^2}\ ;$$

in order to obtain a μ as small as possible we fix Q by

$$Q^{1/(n-1)}=\left(\frac{\omega^2}{R_0^6}\right)^{1/2n}=\left(\frac{\omega}{R_0^3}\right)^{1/n}\ .\tag{6.22}$$

We remark that we did not insert C_R in the definition of Q (as it would have been more natural) since only with the definition (6.22) we will be able to satisfy (5.14). Eq. (6.22) gives $\mu\le 2C_R^2(\frac{R_0^3}{\omega})^{1/n}$. Before giving a more explicit expression for such a quantity we show that (by (1.3)) $z^2\in\mathcal{G}^{\nu}_{R_1\beta/2}$, so that the initial datum z^1+z^2 is in $\mathcal{G}^{\nu}_{\beta R_1}$. We have

$$\begin{aligned}\left\|v^2\right\|^2=\left\|v^2\right\|^2_{H^1}=\sum_{i\notin\mathcal{J}} i^2Z_i(v^0)&\le\sum_{i\notin\mathcal{J}} i^2Z_i(\mathcal{T}_{(0)}(v^0))+\left|\sum_{i\notin\mathcal{J}}\left(i^2Z_i(\mathcal{T}_{(0)}(v^0))-i^2Z_i(v^0)\right)\right|\\ &\le C_1\mathcal{C}(\mathcal{J})R_0^3+C_{16}R_0^4\le 2C_1\mathcal{C}(\mathcal{J})R_0^3\ ,\end{aligned}\tag{6.23}$$

where C_{16} is a constant which exists by (5.9), and the last inequality is ensured by $C_1 \geq C_{16}R_0$ (which is a smallness condition on R_0 which will trivially hold). So, if $2\sqrt{2C_1\mathcal{C}(\mathcal{J})^2} < C_R\beta$, which is just a smallness condition on the free constant C_1, one has by (5.14) (which will be verified in a while)

$$\left\|v^2\right\| \leq \sqrt{2C_1}\mathcal{C}(\mathcal{J}) < \frac{\beta}{2}C_R R_0^{3/2} < \frac{\beta R_1}{2} \ .$$

In order to obtain an estimate of μ in terms of R_0 only, we need an estimate of ω in terms of it. We have

$$\begin{aligned}\left\|v^1\right\|_{H^1}^2 = \left\|v^0\right\|_{H^1}^2 - \left\|v^2\right\|_{H^1}^2 &\geq \left(\left\|u^0\right\| - \left\|u^0 - v^0\right\|\right)^2 - \left\|v^2\right\|^2 \\ &\geq \frac{R_0^2}{2^8} - C_{17}R_0^4 - 2C_1\mathcal{C}(\mathcal{J})^2R_0^3 \geq \frac{R_0^2}{2^9} \ ,\end{aligned}$$

where we used (5.4) and (6.23), and we assumed

$$2C_1\mathcal{C}(\mathcal{J})R_0 < \frac{1}{2^{10}} \ , \quad C_{17}R_0 < \frac{1}{2^{10}} \ , \tag{6.24}$$

(to be checked below), from which

$$\frac{R_0^2}{2^7} \geq \omega \geq \frac{R_0^2}{\Omega} \ , \tag{6.25}$$

where

$$\Omega := 2^9 n \max_{i \in \mathcal{I}} i^2 \ ,$$

and

$$\mu \leq 2C_R^2(\Omega R_0)^{1/n} \preceq \mathcal{C}(\mathcal{J})^5 R_0^{1/n} \ , \tag{6.26}$$

and also, inserting in (6.19)

$$C_R R_0^{3/2} \leq R_1 \leq C_R \frac{R_0^{1+1/2n}}{2^{7(1-1/n)/2}} \ ,$$

and so, provided

$$R_0^{1/2n} < \frac{2^{7(1-1/n)/2}}{10C_R} \preceq \frac{1}{\mathcal{C}(\mathcal{J})^{3/2}} \ , \tag{6.27}$$

(5.14) is satisfied. Finally, if

$$R_0^{1/n} \leq \frac{C_{18}}{\mathcal{C}(\mathcal{J})^5} \tag{6.28}$$

with a suitable C_{18}, (6.26) and all the smallness conditions are satisfied and also (5.18) holds.

We verify now that (5.11) holds. One has

$$\sum_{k\in\mathcal{J}} k^2\nu_k \le \sum_{k\in\mathcal{J}} k^2\left|\nu_k-\omega_k\right| + \sum_{k\in\mathcal{J}} k^2\omega_k \le \mathcal{C}(\mathcal{J})\frac{\omega}{qQ^{1/(n-1)}} + \left\|v^1\right\|_{H^1}^2$$
$$\le \frac{\mathcal{C}(J)}{2^7}R_0^{2+1/n} + \frac{R_0^2}{2^7} < \frac{R_0^2}{8e^2}\ ,$$

by (6.28).

So, we can apply theorem 6.2 which ensures that, for

$$|t| \le C_{19}\exp\left(\frac{R_*}{\mathcal{C}(\mathcal{J})^5 R_0^{1/n}}\right)$$

one has

$$\sum_{k\ge1} k^2\left|Z_k(v(t))-\nu_k\right| \preceq R_1^2 \preceq \mathcal{C}(\mathcal{J})^3 R_0^{2+1/n}\ .$$

To obtain our statement we have to go back to $Z_k(u(t))-Z_k(u^0)$. We have

$$\sum_{k\in\mathcal{J}} k^2\left|Z_k(u(t))-Z_k(u^0)\right| + \sum_{k\notin\mathcal{J}} k^2 Z_k(u(t)) \le \sum_{k\ge1} k^2\left|Z_k(u(t))-Z_k(v(t))\right|$$
$$+\sum_{k\ge1} k^2\left|Z_k(v(t))-\nu_k\right| + \sum_{k\in\mathcal{J}} k^2\left|Z_k(v^0)-Z_k(u^0)\right| + \sum_{k\notin\mathcal{J}} k^2\left|Z_k(v^0)-\nu_k\right|$$
$$\preceq R_0^4 + \mathcal{C}(\mathcal{J})^3 R_0^{2+1/n} \preceq \mathcal{C}(\mathcal{J})^3 R_0^{2+1/n}\ .$$

□

We remark that, from the above proof also theorem 1.2 follows.

Now we comment briefly on some technical points related to the proof of theorem 1.1.

We recall that we have first performed a normalization up to an exponentially small remainder, transforming NLS into a perturbation of the integrable system $h_0+\langle f_{(0)}\rangle$, and that (in the new variables) h_0 is an approximate constant of motion for the complete system; this last fact allowed to consider NLS as a perturbation of $\langle f_{(0)}\rangle$ only. In our scheme this was crucial; indeed, the frequencies of motion of the system $h_0+\langle f_{(0)}\rangle$ are of order one, while the differences between them are of order ϵ, so the dynamics of $h_0+\langle f_{(0)}\rangle$ involves two time scales. Systems of this kind are quite difficult to be studied by Lochak method, which is based on the simultaneous approximation of all frequencies. In our approach such a problem completely disappears (a different solution has been recently given by Niederman [28]).

Some more technical difficulties are related to the singularity of action angle variables at the origin. We solved them by introducing action angle variables only for those degrees of freedom which at time zero (and therefore for all the times we are interested in) are well separated from the origin, and by using the averaging method developed in [1], which does not make use of particular variables in order to construct normal forms.

We also tried to obtain a very good control on the motion of the actions, in order to get an estimate qualitatively stronger than that given by energy conservation. For this reason we introduced the parameter R_1 and left it free until the last step of the proof. This also allowed to obtain the optimal exponent of ϵ in the expression of the stability time.

Proof of corollary 1.4.

Fix an initial datum in $\mathcal{A}_{\alpha,\beta}$, and define $\mathcal{J} := \{1, ..., n\}$ with a still undetermined n. Then one has

$$\sum_{i\geq n+1} i^2 \frac{|u_i^0|^2}{2} \leq \frac{n^2}{e^{\beta n^\alpha}} K^2 \ ,$$

and this is smaller than $C_1 \mathcal{C}(\mathcal{J})^2 \varepsilon^3$ (i.e. the second of (1.3) holds) if

$$\frac{n^2 K^2}{e^{\beta n^\alpha}} \leq C_1 \mathcal{C}(\mathcal{J})^2 \varepsilon^3 \ ,$$

which is implied by

$$e^{-\beta n^\alpha} \leq \frac{C_1}{K^2} \varepsilon^3 \ . \tag{6.29}$$

Fix now

$$n = n_\varepsilon := 2 \times \text{integer part of} \left[\left(\frac{6}{\beta} |\ln \varepsilon| \right)^{1/\alpha} \right] \ ,$$

and assume that $\varepsilon^3 < C_1/K^2$, then (6.29) is satisfied, and therefore the second of (1.3) holds. Finally one has

$$\mathcal{C}(\mathcal{J})^5 \varepsilon^{1/n} \preceq n^{15} e^{-|\ln \varepsilon|/n} \preceq |\ln \varepsilon|^{15/\alpha} \exp \left[- \left(\frac{\beta}{6} \right)^{1/\alpha} \frac{1}{2} |\ln \varepsilon|^{\frac{\alpha-1}{\alpha}} \right] < \eta(\varepsilon) \ ,$$

provided ε is small enough. Using this, corollary 1.4 simply follows. $\square$

We point out that along the same lines it is possible to obtain (in the spirit of [12]) a stability result for initial data having only finite Sobolev regularity. This is given by the following

Corollary 6.4. *Fix $N \geq 1$, and $K > 0$, denote $\mathcal{C}_N := \sum_{i=1}^N i^2$ and assume*

$$\varepsilon := \left\| u^0 \right\|_{H^1} \leq \min \left\{ \left(\frac{\varepsilon_*}{\mathcal{C}_N^5} \right)^N , \frac{C_1^{1/3}}{K^{2/3}} \right\} \ .$$

Fix $s > 3|\ln \varepsilon| / \ln(N+1)$, if

$$\left\| u^0 \right\|_{H^s} := \left(\sum_{k\geq 1} k^{2s} \frac{|u_k|^2}{2} \right)^{1/2} \leq K$$

then one has

$$\sum_{1\leq k\leq N} k^2 |Z_k(t) - Z_k(0)| + \sum_{k>N} k^2 \, |Z_k(t)| \leq C_2 \mathcal{C}_N^3 \varepsilon^{2+1/N} \ , \tag{6.30}$$

for all times t satisfying

$$|t| \leq C_3 \exp \left(\frac{\varepsilon_*}{\mathcal{C}_N^5 \varepsilon^{1/N}} \right) \ . \tag{6.31}$$

7. Proof of theorem 4.4

We prove here the abstract normal form theorem 4.4. The main step of the proof is an iterative lemma which gives the transformation increasing by one the order of normalization.

Lemma 7.1. *(Iterative Lemma) Consider the Hamiltonian*

$$H^{(k)}(z) = h_\omega(z) + \hat{h}(z) + Z^{(k)}(z) + f^{(k)}(z) + r^{(k)}(z) + g(N_1(z), ..., N_l(z)) \ , \tag{7.1}$$

fix $\delta < R/(k+1)$, assume that h_ω, $\hat{h}$, N, g satisfy the hypotheses (0-7) of sect. 4, that $Z^{(k)}$, $f^{(k)}$, $r^{(k)}$ are analytic on $\mathcal{G}_{R-k\delta}$ together with their symplectic gradients, that $Z^{(k)}$ and $f^{(k)}$, commute with N_i for all i, and that

$$\begin{aligned}
&N_{R-k\delta}\left(Z^{(k)}\right) \le \begin{cases} 0 & \text{if} \quad k=0 \\ F\sum_{i=0}^{k-1} {\epsilon_s}^i & \text{if} \quad k \ge 1 \end{cases} , \quad |Z^{(k)}|_{R-k\delta} \le \begin{cases} 0 & \text{if} \quad k=0 \\ E & \text{if} \quad k=1 \\ E^\sharp \sum_{i=0}^{k-1} {\epsilon_s}^i & \text{if} \quad k \ge 2 \end{cases} \\
&N_{R-k\delta}\left(f^{(k)}\right) \le {\epsilon_s}^k F \ , \qquad\qquad |f^{(k)}|_{R-k\delta} \le E^\sharp {\epsilon_s}^k \\
&N_{R-k\delta}\left(r^{(k)}\right) \le \begin{cases} \Lambda & \text{if} \quad k=0 \\ \Lambda\prod_{i=0}^{k-1}\left(1+\frac{3FT}{\delta}{\epsilon_s}^i\right) & \text{if} \quad k \ge 1 \end{cases} , \quad |r^{(k)}|_{R-k\delta} \le \lambda \ ,
\end{aligned} \tag{7.2}$$

with

$$\epsilon_s := \frac{T}{\delta}(18F + 5F_0) \ , \quad E^\sharp := \max\{E, \frac{2F(3E+E_0)}{9F+5F_0}\} \ , \tag{7.3}$$

then, if $\epsilon_s < 1/2$ there exists a canonical transformation $\mathcal{T}_k$ satisfying

$$N_i \circ \mathcal{T}_k = N_i \ , \quad |\mathcal{T}_k - \mathbb{1}|_{R-\delta(k+1)} \le TF{\epsilon_s}^k \ , \tag{7.4}$$

such that $H^{(k)} \circ \mathcal{T}_k$ has the form (7.1) and satisfies (7.2) with $k+1$ in place of k, moreover $f^{(k+1)}$ and $Z^{(k+1)} := Z^{(k)} + \langle f^{(k)} \rangle$, commute with N_i for all i, and

$$r^{(k+1)} = r^{(k)} \circ \mathcal{T}_k \ .$$

Proof. As anticipated in sect. 4 we look for $\mathcal{T}_k$ as the time one map of the Hamiltonian vector field of a function χ_k which is analytic together with its symplectic gradient $Jd\chi_k$ (method of Lie transform). So, consider the differential equation

$$\dot{z} = Jd\chi_k(z) \ ; \tag{7.5}$$

by standard theory one has that, provided $N_{R-k\delta}(\chi_k)$ is small enough and $z_0 \in \mathcal{G}_{R-\delta(k+1)}$, the solution of (7.5) exists for all t with $|t| \le 1$. So, we can define the flow $\mathcal{T}_k^t : \mathcal{G}_{R-(k+1)\delta} \to \mathcal{G}_{R-\delta k}$, and in particular the corresponding time one map $\mathcal{T}_k := \mathcal{T}_k^t|_{t=1}$, which is an analytic canonical transformation, close to the identity.

Remark 7.2. *Assuming* $\{\chi, N_i\} = 0$ *one has*

$$N_i \circ \mathcal{T}_k = N_i \ ;$$

moreover for any h *with regular symplectic gradient one has*

$$\{h \circ \mathcal{T}_k, N_i\} = \{h \circ \mathcal{T}_k, N_i \circ \mathcal{T}_k\} = \{h, N_i\} \ ,$$

so that in particular

$$\{h, N_i\} = 0 \implies \{h \circ \mathcal{T}_k, N_i\} = 0$$

i.e. the transformed of a symmetric function is symmetric.

Assume now $\{\chi, N_i\} = 0$. We have

$$\begin{aligned} H^{(k)} \circ \mathcal{T}_k &= h_\omega + \hat{h} + Z^{(k)} + r^{(k)} \circ \mathcal{T}_k + g \\ &+ \left[\{\chi_k, h_\omega\} + f^{(k)}\right] + (h_\omega \circ \mathcal{T}_k - h_\omega - \{\chi_k, h_\omega\}) \\ + \left(\hat{h} \circ \mathcal{T}_k - \hat{h}\right) &+ \left(Z^{(k)} \circ \mathcal{T}_k - Z^{(k)}\right) + \left(f^{(k)} \circ \mathcal{T}_k - f^{(k)}\right) \ . \end{aligned}$$

We use lemma 8.4 to determine χ_k in such a way that the square bracket is reduced to $\langle f^{(k)} \rangle$; from lemma 8.4 we also have

$$N_{R-k\delta}\left(\chi_k\right) \leq {\epsilon_s}^k F T \ , \tag{7.6}$$

and $\{\chi_k, N_i\} = 0$; moreover, using lemma 8.1 the estimate (7.4) and the estimate of $N_{R-\delta(k+1)}\left(Z^{(k+1)}\right)$ follow. The estimate of $r^{(k+1)}$ follows from lemma 8.2. Define $f^{(k+1)}$ as the sum of the terms in ronde brackets. So, $f^{(k+1)}$ is easily estimated using lemmas 8.3 and 8.5 below, giving

$$N_{R-(k+1)\delta}\left(f^{(k+1)}\right) \leq \left(\frac{10}{\delta}{\epsilon_s}^k F + \frac{5}{\delta}F_0 + \frac{5}{\delta}F\sum_{i=0}^{k-1}{\epsilon_s}^i + \frac{5}{\delta}F{\epsilon_s}^k\right) N_{R-k\delta}\left(\chi_k\right) \ . \tag{7.7}$$

If $k = 0$ then the third term is not present and (7.7) is less than

$$(15F + 5F_0)\frac{T}{\delta}F{\epsilon_s}^k < F{\epsilon_s}^{k+1} \ ,$$

if $k \geq 1$, exploiting ${\epsilon_s}^k < 1/2$ one has that (7.7) is less than

$$(18F + 5F_0)\frac{T}{\delta}F{\epsilon_s}^k = F{\epsilon_s}^{k+1} \ .$$

We come to the estimates of the moduli of functions. For $k = 0$, using lemmas 8.3, 8.5 one has

$$\left|Z^{(1)}\right|_{R-\delta} \leq E \ ,$$

and

$$\left|f^{(1)}\right|_{R-\delta} \le \frac{FT}{\delta}(6E+2E_0) \le E^\sharp \epsilon_s \ ,$$

where we used the definitions of ϵ_s and $E^\sharp$. For $k \ge 1$ one has

$$\left|f^{(k+1)}\right|_{R-k\delta} \le \frac{TF{\epsilon_s}^k}{\delta}(9E^\sharp+2E_0) \le E^\sharp {\epsilon_s}^{k+1} \ ,$$

where we used again the definitions of ϵ_s and $E^\sharp$. □

Proof of theorem 4.4. The Hamiltonian (4.1) satisfies the hypotheses of lemma 7..1 with $k=0$, with $r+f_1$ in place of $r^{(0)}$, f in place of $f^{(0)}$; $E_1+\lambda$, and $F_1+\Lambda$ in place of λ and Λ respectively. So we apply lemma 7.1 one time with $\delta = R/4$; this is possible provided

$$\frac{4T}{R}\left(18F+5F_0\right) \le \frac{1}{2} \ ;$$

but the l.h.s. is less than ϵ. So we have that there exists a canonical transformation $\tilde{\mathcal{T}} : \mathcal{G}_{\frac{3R}{4}} \to \mathcal{G}_R$ with

$$\left\|z - \tilde{\mathcal{T}}(z)\right\|_{\frac{3R}{4}} \le TF \ ,$$

such that

$$H \circ \tilde{\mathcal{T}} = h_\omega + \hat{h}^{(1)} + \tilde{f} + \tilde{r} + g \ , \tag{7.8}$$

with $\hat{h}^{(1)} = \hat{h} + \langle f_{(0)} \rangle$, $\tilde{f} := f^{(1)} + f_1 \circ \tilde{\mathcal{T}}$, and

$$\begin{aligned}
N_{\frac{3R}{4}}\left(\hat{h}^{(1)}\right) &\le F_0 + F =: \tilde{F}_0 \ , \\
\left|\hat{h}^{(1)}\right|_{\frac{3R}{4}} &\le E_0 + E =: \tilde{E}_0 \ , \\
N_{\frac{3R}{4}}\left(\tilde{f}\right) &\le \frac{4T}{R}\left(18F+5F_0\right)F + F_1\left(1+\frac{12FT}{R}\right) \le 2\epsilon F =: \tilde{F} \ , \\
\left|\tilde{f}\right|_{\frac{3R}{4}} &\le \frac{4T}{R}\left(18F+5F_0\right)\max\left\{E, \frac{2F(3E+E_0)}{9F+5F_0}\right\} + E_1 \\
&\le \frac{\epsilon}{4}\max\left\{E, \frac{2F(3E+E_0)}{9F+5F_0}\right\} + E_1 =: \tilde{E} \ , \\
N_{\frac{3R}{4}}\left(\tilde{r}\right) &\le \Lambda\left(1+\frac{12FT}{R}\right) =: \tilde{\Lambda} \ , \\
\left|\tilde{r}\right|_{\frac{3R}{4}} &\le \lambda \ .
\end{aligned}$$

By remark 7.2 it is also clear that $\left\{\tilde{f}, N_i\right\} = 0$. So, (7.8) satisfy again the hypotheses of lemma 7.1 with $k=0$ with $\hat{h}^{(1)}$, $\tilde{f}$, $\tilde{r}$ in place of $\hat{h}$, $f^{(0)}$, $r^{(0)}$, and $\frac{3R}{4}$, $\tilde{F}_0$, $\tilde{F}$, $\tilde{E}$, $\tilde{E}_0$ $\tilde{\Lambda}$ in place of R, F_0, F, E, E_0 Λ. We fix now $\delta = R/(4k)$ and apply k times the lemma, so we obtain an Hamiltonian of the form (4.14), with the estimates

$$N_{\frac{R}{2}}(Z) \leq 2\tilde{F} \ , \quad |Z|_{\frac{R}{2}} \leq 2\tilde{E}^{\sharp} \ ,$$
$$N_{\frac{R}{2}}(\mathcal{R}) \leq \tilde{F}{\epsilon_s}^k \ , \quad |\mathcal{R}|_{\frac{R}{2}} \leq \tilde{E}^{\sharp}{\epsilon_s}^k \ ,$$
$$N_{\frac{R}{2}}(\bar{r}) < \tilde{\Lambda}\prod_{l=0}^{\infty}\left(1+{\epsilon_s}^{l+1}\right) < \tilde{\Lambda}e^{2\epsilon_s}$$

with

$$\tilde{E}^{\sharp} := \max\left\{\tilde{E}, \frac{2\tilde{F}(3\tilde{E}+\tilde{E}_0)}{9\tilde{F}+5\tilde{F}_0}\right\} \ , \quad \epsilon_s := \left[\frac{4T}{R}(18\tilde{F}+5\tilde{F}_0)e\right]\frac{k}{e} \ . \tag{7.9}$$

Denote by ϵ the square bracket, then we determine k in order to minimize the estimate of $\mathcal{R}$, this gives

$$k = \text{integer part}\left(\frac{1}{\epsilon}\right) \Longrightarrow {\epsilon_s}^k \leq e\exp\left(-\frac{1}{\epsilon}\right) \ ,$$

namely the exponential estimate of the remainder. To obtain the statement of theorem 4.4 one has just to insert the definition of the quantities with tildes in terms of the original quantities. In lemma 8.6 the calculation is done for $\tilde{E}^{\sharp}$, the other calculations are simpler and are omitted. □

8. Technical lemmas

Lemma 8.1. *Let χ be analytic in $\mathcal{G}_\rho$ together with its symplectic gradient, fix $\delta < \rho$. Assume $N_\rho(\chi) < \delta$ and consider the time t flow T^t of the corresponding Hamiltonian vector field. Then, for $|t| \leq 1$, one has*

$$|T^t - \mathbb{1}|_{\rho-\delta} \leq N_\rho(\chi) \ . \tag{8.1}$$

Proof. It is just an application of the equality

$$z(t) - z(0) = \int_0^t \frac{dz}{dt}(s)ds = \int_0^t Jd\chi(z(s))ds \ .$$

□

Lemma 8.2. *Consider χ as above and let $h : \mathcal{G}_\rho \to \mathbf{C}$ be an analytic function with vector field analytic in $\mathcal{G}_\rho$, fix $0 < \delta < \rho$ assume $N_\rho(\chi) \leq \delta/3$, then, for $|t| \leq 1$, one has*

$$N_{R-\delta}\left(h \circ T^t\right) \leq \left(1 + \frac{3}{\delta}N_\rho(\chi)\right)N_\rho(h)$$

Proof. First remark that, since T^t is a canonical transformation one has

$$Jd(h \circ T^t)(z) = dT^{-t}(T^t(z))Jdh(T^t(z)) \ , \tag{8.2}$$

from which

$$Jd(h \circ T^t)(z) = \left(dT^{-t}(T^t(z)) - \mathbb{1}\right) Jdh(T^t(z)) + Jdh(T^t(z)) \ .$$

To estimate the first term fix $\delta_1 := \delta/3$; we have

$$\begin{aligned} \sup_{z \in \mathcal{G}_{\rho-3\delta_1}} \left\| dT^{-t}(T^t(z)) - \mathbb{1} \right\| &\leq \sup_{z \in \mathcal{G}_{\rho-2\delta_1}} \left\| dT^{-t}(z) - \mathbb{1} \right\| \\ &\leq \frac{1}{\delta_1} \sup_{z \in \mathcal{G}_{\rho-\delta_1}} \left\| T^{-t}(z) - z \right\| \leq \frac{1}{\delta_1} N_\rho(\chi) \ . \end{aligned} \tag{8.3}$$

Going back to δ and adding the trivial estimate of the second term one has the thesis. □

Lemma 8.3. *Consider χ as above, and let $h : \mathcal{G}_\rho \to \mathcal{B}$ be an analytic function, fix $0 < \delta < \rho$, and assume $N_\rho(\chi) < \delta/3$, then, for $|t| \leq 1$ one has*

$$\left| h \circ T^t - h \right|_{\rho-\delta} \leq \frac{2}{\delta} N_\rho(\chi) \, |h|_\rho \ . \tag{8.4}$$

Moreover, if h takes values in $\mathbf{C}$ and its symplectic gradient is analytic on $\mathcal{G}_\rho$ one has

$$N_{\rho-\delta}\left(h \circ T^t - h\right) \leq \frac{5}{\delta} N_\rho(h) \, N_\rho(\chi) \tag{8.5}$$

Proof. To obtain (8.4) fix $\delta_1 := \delta/2$, one has

$$\sup_{z \in \mathcal{G}_{\rho-2\delta_1}} \left\| h(T^t(z)) - h(z) \right\| \leq \sup_{z \in \mathcal{G}_{\rho-\delta_1}} \|dh(z)\| \sup_{z \in \mathcal{G}_{\rho-2\delta_1}} \left\| T^t(z) - z \right\| \leq \frac{1}{\delta_1} |h|_\rho N_{\rho-\delta_1}(\chi) \ ;$$

expressing δ_1 in terms of δ one has (8.4).

We come to (8.5). One has

$$Jd(h \circ T^t)(z) - Jdh(z) = \left(dT^{-t}(T^t(z)) - \mathbb{1}\right) Jdh(T^t(z)) + \left[Jdh(T^t(z)) - Jdh(z)\right] \ .$$

The norm of the square bracket is estimated using (8.4). The other term was already estimated in (8.3), so we have the thesis. □

Lemma 8.4. *Let $h \in \mathcal{G}_\rho$ be analytic together with its symplectic gradient and such that $\{h, N_i\} = 0$ for $i = 1, ..., l$; consider the homological equation*

$$\{h_\omega, \chi\} + \langle h \rangle = h \ ; \tag{8.6}$$

where

$$\langle h\rangle(z) := \frac{1}{T}\int_0^T h(\Phi^t(z))dt$$

It has a solution

$$\chi(z) := \frac{1}{T}\int_0^T t\left[h(\Phi^t(z)) - \langle h\rangle(\Phi^t(z))\right]dt\ . \tag{8.7}$$

Moreover one has that $\langle h\rangle$, χ and their symplectic gradients are analytic on $\mathcal{G}_\rho$ and satisfy the estimates

$$N_\rho\left(\langle h\rangle\right) \le N_\rho\left(h\right)\ \ ,|\langle h\rangle|_\rho \le |h|_\rho\ ,$$
$$N_\rho\left(\chi\right) \le TN_\rho\left(h\right)\ \ ,|\chi|_\rho \le T|h|_\rho\ ,$$

and $\{\chi, N_i\} = \{\langle h\rangle, N_i\} = 0$, $\forall i = 1, ..., l$.

Proof. Apart from the statement concerning the commutation with the N_i's this is exactly lemma 5.3 of [1]. We give its proof for the sake of completeness. First remark that by our assumption Φ^t is a canonical transformation, namely one has

$$\Omega\left(d\Phi^t(z)X, d\Phi^t(z)Y\right) = \Omega(X,Y)\ ,\quad \forall X,Y\in\mathcal{P}\ .$$

Denote $G := h - \langle h\rangle$, then it is easy to see that, for z in the domain $\mathcal{D}$ of differentiability (with respect to time) of Φ, the function χ defined by (8.7) solves (8.6) (just compute the time derivative of $\chi(\Phi^t(z))$). Then, exploiting the fact that Φ^t is a canonical transformation it easy to see that

$$Jd\chi(z) = \frac{1}{T}\int_0^T t[d\Phi^{-t}(\Phi^t(z))](JdG)(\Phi^t(z))dt\ , \tag{8.8}$$

which is continuous in z; so $\{h_\omega, \chi\}$ is continuous on $\mathcal{G}_\rho$. Since the two sides of (8.6) coincide on the dense set $\mathcal{D}$ they coincide everywhere, and therefore χ is the solution of (8.6). To obtain the estimates of the vector field of χ just use (8.8).

To prove that $\langle h\rangle$ commutes with N_i just use its definition; it gives

$$\begin{aligned}\left\{N_i, \langle h\rangle\right\}(z) &= \frac{1}{T}\int_0^T \left\{N_i, h\circ\Phi^t\right\}(z)dt\\ &= \frac{1}{T}\int_0^T \left\{N_i\circ\Phi^t, h\circ\Phi^t)\right\}(z)dt = \frac{1}{T}\int_0^T \left\{N_i, h\right\}(\Phi^t(z))dt = 0\ ,\end{aligned}$$

where we used (4.12). Similarly one obtains the result for χ. □

Although h_ω has unbounded symplectic gradient the symplectic gradient of $h_\omega\circ T - h_\omega$ is bounded, more precisely we have

Lemma 8.5. *Let χ be the solution of the homological equation (8.6). Denote by T^t the flow of the Hamiltonian vector field associated to χ, and by $T := T^1$ the corresponding time one map, moreover denote*

$$\ell(z) := h_\omega(T(z)) - h_\omega(z) - \left\{\chi, h_\omega\right\}(z)\ ,$$

then, for any $\delta < \rho$, the symplectic gradient of ℓ is analytic on $\mathcal{G}_{\rho-\delta}$, and satisfies

$$|\ell|_{\rho-\delta} \le \frac{4}{\delta} |h|_\rho N_\rho(\chi) \ , \quad N_{\rho-\delta}(\ell) \le \frac{10}{\delta} N_\rho(\chi) N_\rho(h) \, . \tag{8.9}$$

Proof. One has

$$h_\omega(T(z)) - h_\omega(z) = \int_0^1 \frac{d}{dt} h_\omega(T^t(z)) dt = \int_0^1 (h(T^t(z)) - \langle h \rangle (T^t(z))) dt \ ,$$

where we used the homological equation to calculate $\{h_\omega, \chi\}$. Denote again $G := h - \langle h \rangle$, then one has

$$\ell(z) = \int_0^1 (G(T^t(z)) - G(z)) dt \ ;$$

Using (8.4) and $|h - \langle h \rangle| \le 2|h|$ one gets the first of equations (8.9). The second is obtained similarly, but using (8.5). $\square$

Lemma 8.6. *The quantity $\tilde{E}^\sharp$, cf. (7.9) is estimated by*

$$\tilde{E}^\sharp \le \max\left\{ \frac{12}{5} \frac{F(E+E_0)}{F+F_0}, \frac{5}{4} E + E_0 \right\} . \tag{8.10}$$

Proof. One has

$$\tilde{E}^\sharp = \max\left\{ \tilde{E}, \frac{2\tilde{F}(3\tilde{E}+\tilde{E}_0)}{9\tilde{F}+5\tilde{F}_0} \right\} \ .$$

To estimate the second term in the maximum we estimate $\tilde{E}$ as follows

$$\tilde{E} = \frac{\epsilon}{4} \max\left\{ E, \frac{2F}{9F+5F_0}(3E+E_0) \right\} + \epsilon E$$
$$< \frac{\epsilon}{4} \max\left\{ E, \frac{2}{9}(3E+E_0) \right\} + \epsilon E < \frac{\epsilon}{4}(E+E_0) + \epsilon E = \frac{5}{4} E + E_0 \ .$$

So we have

$$\frac{2\tilde{F}(3\tilde{E}+\tilde{E}_0)}{9\tilde{F}+5\tilde{F}_0} = \frac{4\epsilon F}{18\epsilon F + 5F + 5F_0}(3\tilde{E}+\tilde{E}_0)$$
$$< \frac{4\epsilon F}{5(F+F_0)} \left[3\epsilon \left(\frac{5}{4} E + E_0 \right) + E + E_0 \right] \le \frac{4}{5} \epsilon \frac{F}{F+F_0} \left[\frac{15}{8} E + \frac{3}{2} E_0 + E + E_0 \right]$$
$$< \frac{4}{5} \epsilon \frac{F}{F+F_0} (3E + 3E_0) \ ,$$

from which (8.10) follows. $\square$

Lemma 8.7. *Assume $z \in \mathcal{G}^{\nu}_{2R_1}$, denote by u the function whose coordinates are z, then one has*

$$u \in \mathcal{G}^{0}_{R_0/2} \ .$$

Proof. We have

$$u(x) = \sum_{k\in\mathcal{I}} \sqrt{2(\nu_k + I_k)}e^{i\phi_k}\sin(kx) + \sum_{k\notin\mathcal{I}}(p_k + iq_k)\sin(kx) \ ,$$

and therefore

$$\|u\|^2_{H^1} \le \sum_{k\in\mathcal{I}} k^2 \left|\nu_k + I_k\right| \left|e^{i\phi_k}\right|^2 + \sum_{k\notin\mathcal{I}} k^2 \frac{|p_k|^2 + |q_k|^2}{2}$$

$$\le e^2\sum_{k\in\mathcal{I}} k^2\nu_k + e^2\sum_{k\in\mathcal{I}} k^2|I_k| + 4R_1^2 \le \frac{R_0^2}{8} + (e^2+4)R_1^2 \ ,$$

where we used (5.11); using (5.14) we get the thesis. □

Lemma 8.8. *Let $h : \mathcal{G}^{\nu}_{2R_1} \to \mathbf{C}$ be analytic together with its symplectic gradient, assume*

$$|h|_{2R_1} \le C_a \ , \tag{8.11}$$

$$\sup_{z\in\mathcal{G}^{\nu}_{2R_1}} \|Jdh(z)\|_{H^1} \le C_b \ , \tag{8.12}$$

then one has

$$N^{\nu}_{R_1}(h) \le C_b + 2\frac{C_a}{R_1} \ . \tag{8.13}$$

Proof. It is clear that the norm of the (pq) components of Jdh is less than C_b. To estimate the (I,ϕ) components we use (8.11). One has

$$\|(Jdh)_{I,\phi}\| = \sum_{i\in\mathcal{I}} \frac{i^2}{R_1}\left|\frac{\partial h}{\partial \phi_i}\right| + \sup_{i\in\mathcal{I}} R_1 \left|\frac{\partial h}{\partial I_i}\right| \ .$$

We begin by the first term. First remark that, by the definition of the norm h is analytic in the angle ϕ_i in the strip $|\mathrm{Im}\,\phi_i| < 2$, so we have

$$\sup_{z\in\mathcal{G}^{\nu}_{R_1}} \left|\frac{\partial h}{\partial\phi_i}(z)\right| < \frac{1}{1}\sup_{z\in\mathcal{G}^{\nu}_{2R_1}} |h(z)| < C_a \ .$$

With respect to I_i the analyticity strip of h is $2R_1^2$, so we have

$$\sup_{z\in\mathcal{G}^{\nu}_{R_1}} \left|\frac{\partial h}{\partial I_i}\right| \le \frac{1}{R_1^2}\sup_{z\in\mathcal{G}^{\nu}_{2R_1}} |h(z)| < \frac{C_a}{R_1^2} \ ,$$

from which the thesis follows. □

Appendix. On finite order integer functions and the spaces $\mathcal{A}_{\alpha,\beta}$

In this appendix we recall some important facts of the theory of integer functions of finite order (for more details see e.g. [20]), and give the connections with the spaces $\mathcal{A}_{\alpha,\beta}$ introduced in sect. 1.

Define Weierstrass *prime factors* as

$$E_0(z) := (1-z)\ ; E_p(z) := (1-z)\exp\left(\frac{z}{1}+\frac{z^2}{2}+...+\frac{z^p}{p}\right)\ ,\quad p\geq 1$$

then, given an integer function $u : \mathbf{C}\to\mathbf{C}$, Weierstrass theorem ensures that it can be factored according to

$$u(z) = z^k e^{g(z)}\prod_{n\geq 1} E_{p_n}\left(\frac{z}{z_n}\right)\ ,$$

where g is an integer function, p_n a suitable sequence of integer numbers, and z_n the sequence of the zeros of u.

In the case of integer functions of finite order it is possible to say more. First we recall the definition of order and type of an integer function. Given an integer function u define $M(r,u) := \sup_{\theta\in[0,2\pi]}\left|u(re^{i\theta})\right|$; then the order ρ of u is defined by

$$\rho := \limsup_{r\to\infty}\frac{\ln\ln M(r,u)}{\ln r}\ ,$$

while the type γ of u is defined by

$$\gamma := \limsup_{r\to\infty}\frac{\ln M(r,u)}{r^\rho}\ .$$

Remark that if ρ and γ are both positive and finite then one has $M(r,u)\leq C\exp\gamma r^\rho$.

An important example of integer function of finite order is constructed as follows: consider a sequence $S := \{z_n\}_{n\geq 1}$ of complex numbers, such that $z_n\neq 0$, and $|z_n|\leq|z_{n+1}|$; define its convergence exponent σ as the infimum of the real τ's for which the series

$$\sum_{n\geq 1} r_n^{-\tau}\ , \tag{A.1}$$

converges (here $r_n := |z_n|$). Assume $\sigma<\infty$; let $p := [\sigma]$ if σ is not integer; if it is an integer, let $p := \sigma-1$ if the series (A.1) with $\tau=\sigma$ converges, $p := \sigma$ if it diverges.

Then one has

Proposition A.1. *The canonical product*

$$\prod_{n=1}^{\infty} E_p\left(\frac{z}{z_n}\right)$$

converges and is an integer function of order σ. Moreover the set of its zeros coincides with S.

One of the main results in the theory of finite order integer functions is Hadamard's factorization theorem

Theorem A.2. *Let u be an integer function of order ρ, then one has*

$$u(z) = z^m e^{g(z)} \prod_{k=1}^{\infty} E_p\left(\frac{z}{z_k}\right) \ ,$$

where $p \leq \rho$, and g is a polynomial of order less or equal to ρ.

The relation between finite order integer functions and the spaces $\mathcal{A}_{\alpha,\beta}$ was essentially established by Pólya and Wiener [29] who proved the following

Proposition A.3. *Consider the function*

$$u(z) := \sum_{k\in\mathbf{Z}} c_k e^{ikz} \ ;$$

it is an analytic integer function of order ρ if and only if

$$\lim_{k\to\infty} \frac{\ln|\ln|c_k||}{\ln|k|} = \frac{\rho}{\rho-1} \ . \tag{A.2}$$

To give a precise statement one has to pay attention also to the type of the considered function or equivalently to the fact that (A.2) does not imply $u \in \mathcal{A}_{\alpha,\beta}$ with $\alpha = \rho/(\rho-1)$ and a finite β. This leads to the following

Corollary A.4. *Let $u \in \bigcup_{\beta>0} \mathcal{A}_{\alpha,\beta}$ then it extends to a skew symmetric 2π periodic integer function of order less or equal to $\rho := \alpha/(\alpha-1)$. Viceversa, let u be a skew symmetric 2π periodic integer function of order strictly less than ρ then $u \in \bigcap_{\beta>0} \mathcal{A}_{\alpha,\beta}$ with $\alpha := \rho/(\rho-1)$.*

It could also be proved that if $u \in \bigcup_{\beta>0} \mathcal{A}_{\alpha,\beta}$ has order exactly equal to $\alpha/(\alpha-1)$ then it has finite type.

Finally remark that any function of $\mathcal{A}_{\alpha,\beta}$ can be represented by Hadamard's theorem.

References

[1] D. Bambusi, A. Giorgilli: *Exponential Stability of States Close to Resonance in Infinite Dimensional Hamiltonian Systems.* Jour. Stat. Phys. **71**, p. 569-606 (1993).

[2] D. Bambusi: *Exponential Stability of Breathers in Hamiltonian Chains of Weakly Coupled Oscillators.* Nonlinearity **9** 433-457 (1996).

[3] D. Bambusi, N.N. Nekhoroshev: *A property of exponential stability in nonlinear wave equation near the fundamental linear mode.* Preprint 16/1996 Dipartimento di Matematica, Università di Milano.

[4] D. Bambusi: *Long time stability of some small amplitude solutions in nonlinear Schrödinger equations* Commun. Math. Phys. **189** 205-226 (1997)

[5] D. Bättig, A.M. Bloch, J.C. Guillot, T. Kappeler: *On the Symplectic structure of the phase space for periodic KdV, Toda, and defocusing NLS* Duke Math. Journal **79**, 549–604 (1995).

[6] G. Benettin, L. Galgani, A. Giorgilli: *Numerical Investigations on a Chain of Weakly Coupled Rotators in the Light of Classical Perturbation Theory.* Nuovo Cimento B, **89**, 103-119 (1985).

[7] G. Benettin, L. Galgani, A. Giorgilli: *Classical Perturbation Theory for Systems of Weakly Coupled Rotators .* Nuovo Cimento B, **89**, 89-102 (1985).

[8] G. Benettin, G. Gallavotti: *Stability of motion near resonances in quasi–integrable Hamiltonian systems.* J. Stat. Phys. **44**, 293–338 (1986).

[9] G. Benettin, L. Galgani, A. Giorgilli: *Realization of Holonomic Constraints and Freezing of High Frequency Degrees of Freedom in the Light of Classical Perturbation Theory, Part I.* Commun. Math. Phys. **113**, 87-103 (1987).

[10] 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).

[11] G. Benettin, L. Galgani, A. Giorgilli: *Realization of Holonomic Constraints and Freezing of High Frequency Degrees of Freedom in the Light of Classical Perturbation Theory, Part II.* Commun. Math. Phys. **121**, 557-601 (1989).

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

[13] J. Bourgain: Nonlinear Schrödinger equations. Preprint 1995

[14] R.C. Churchill, M. Kummer, D.L. Rod: *On averaging, reduction, and symmetry* J. Diff. Eq. **49**, 359–414 (1983).

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

[16] J. Fröhlich, T. Spencer and C.E. Wayne: *Localization in Disordered, Nonlinear Dynamical Systems.* J. Stat. Phys. **42**, 247-274 (1986).

[17] A. Giorgilli: *Rigorous results on the power expansions for the integrals of a hamiltonian system near an elliptic equilibrium* Ann. Inst. Henri Poincaré, **48** 423–439 (1988).

[18] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani, C. Simó: *Effective Stability for a Hamiltonian System near an Elliptic Equilibrium Point, with an Application to the Restricted Three body Problem.* J. Diff. Eq. **77**, 167–198 (1989).

[19] B.Grebert, T.Kappeler: *Action-angle variables for NLS and applications*, in preparation.

[20] E. Hille: Analytic Function Theory. Ginn and Company (Boston 1962)

[21] S.B. Kuksin: *Hamiltonian Perturbations of Infinite-dimensional Linear Systems with Imaginary Spectrum.* Funktsional. Anal. i Prilozhen. **21**, (3), 22-37 (1987) = Funct. Anal. Appl. **21** (1987).

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

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

[24] P. Lochak: *Canonical Perturbation Theory via Simultaneous Approximation.* Uspekhi Math. Nauk, **47** (6)59–140, 1992. English translation: Russ. Math. Surv. **47** 57–133 (1992).

[25] J.E. Marsden: Lectures on Geometric Methods in Mathematical Physics. SIAM (Philadelphia, 1981).

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

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

[28] L. Niederman: *Nonlinear stability around an elliptic equilibrium point in an Hamiltonian system.* Preprint 1997.

[29] G.Pólya, N. Wiener: *On the oscillation of the derivatives of a periodic function.* Trans. Amer. Math. Soc. **52** 249 (1942).

[30] J. Pöschel: *On Lower Dimensional Tori in Hamiltonian Systems.* Math. Z. **202** 559-608 (1989)

[31] J. Pöschel: *Small Divisors with Spatial Structure in Infinite Dimensional Hamiltonian Systems.* Commun. Math. Phys. **127**, 351-393 (1990).

[32] J. Pöschel: *Nekhoroshev Estimates for Quasi–convex Hamiltonian Systems.* Math. Z. **213** 187–216 (1993).

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

[34] J. Pöschel: *On the construction of almost periodic solutions for a nonlinear Schrödinger equation* Preprint 95-13 Univerität Stuttgart (1995).

[35] C.E. Wayne: *Periodic and Quasi-periodic Solutions of Nonlinear Wave Equation via KAM Theory.* Commun. Math. Phys. **127**, 479-528 (1990).

[36] J. You: *Perturbation of lower dimensional tori for hamiltonian systems.* Preprint 1997.