INVARIANT KAM TORI AND GLOBAL STABILITY FOR HAMILTONIAN SYSTEMS

ANTONIO GIORGILLI
Dipartimento di Matematica dell' Università di Milano
and Gruppo Nazionale di Fisica Matematica del CNR,
Via Saldini 50, 20133 Milano, Italy.

ALESSANDRO MORBIDELLI
CNRS, Observatoire de la Côte d'Azur
BP 229, 06304 – NICE Cedex 4, France

Abstract. We point out a deep connection between KAM theorem and Nekhoroshev's theorem. Precisely, we reformulated the construction by Arnold of the set of invariant tori using Nekhoroshev's theorem as a basic tool. We prove in this way the existence of a hierarchic structure of nested domains characterized by a diffusion speed exponentially decreasing at each step. The set of KAM tori appears as the domain characterized by vanishing diffusion speed.

1. Introduction and statement of the result

In a previous paper[1] we studied the stability in the neighbourhood of an invariant KAM torus for a Hamiltonian system in the light of Nekhoroshev's theory. The aim of the present work is to produce a global version of that result which brings to light a strong connection existing between the Nekhoroshev[2][3] and the KAM[4][5][6] theorems.

The main point is that Arnold's construction of the set of invariant tori can be completely reformulated using Nekhoroshev's theorem as a basic iterative tool. Indeed, a careful reading of the usual proof of Nekhoroshev's theorem allows one to extract the following information: on the one hand, one has stability over times of the order of $O(\exp(1/\varepsilon))$ all over the phase space; on the other hand, in a subset of phase space characterized by absence of resonances with order smaller than $O(1/\varepsilon)$ the Hamiltonian can be given the form of an integrable system with a perturbation of size $O(1/\exp(1/\varepsilon))$. In such a domain the theorem of Nekhoroshev can be applied again, and the same procedure can be iterated infinitely many times, giving rise to a process similar to that of Arnold's proof of KAM theorem. Actually, the only technical difference is that the quadratic step of Arnold is replaced by a Nekhoroshev step reducing the perturbation to be exponentially small.

The interest is that one finds a complete connection between the theorem by Nekhoroshev and KAM theorem, which can be expressed, in rough words, by saying that the applicability of Nekhoroshev theorem implies the existence of a structure of nested domains, centered on invariant tori, characterized by stability times increasing with exponential steps. This result has been announced in a previous short communication[7]. We refer to that paper for a more complete discussion, and for a brief sketch of the proof. The aim of the present paper is to provide a complete and self–contained proof.

We consider a canonical system of differential equations with Hamiltonian

$$H(p,q) = h(p) + f(p,q) \ , \tag{1}$$

where $p \in \mathcal{G} \subset \mathbf{R}^n$ are action variables, and $q \in \mathbf{T}^n$ are angles. The Hamiltonian will be assumed to be real analytic in $\mathcal{G} \times \mathbf{T}^n$, and to admit a holomorphic extension to a complex domain $\mathcal{G}_\varrho \times \mathbf{T}^n_\xi$, where ϱ and ξ are positive constants, and

$$\mathcal{G}_\varrho = \bigcup_{p\in\mathcal{G}} B_\varrho(p) \ , \quad \mathbf{T}^n_\xi = \{q \in \mathbf{C}^n \ : \ |\,\mathrm{Im}\, q_j| < \xi \, , \, 1 \le j \le n\} \ . \tag{2}$$

Here, $B_\varrho(p)$ denotes the open complex ball of radius ϱ and center p, namely

$$B_\varrho(p) = \{p' \in \mathbf{C}^n \ : \ |p' - p| < \varrho\} \ ,$$

where $|\cdot|$ denotes the norm $|p| = \max_j |p_j|$. The symbol $\|\cdot\|$ will denote the Euclidean norm. We shall also denote

$$\omega(p) = \frac{\partial h}{\partial p} \ , \quad A(p) = \frac{\partial^2 h}{\partial p \partial p} \tag{3}$$

the frequencies and the Hessian matrix of the unperturbed Hamiltonian $h(p)$, respectively. With a little abuse, we shall denote by $\mathrm{Vol}(\mathcal{G}_\varrho)$ the volume of the real part of the complex domain $\mathcal{G}_\varrho$.

Moreover, we shall say that a torus $\mathcal{T}$ is (δ, T)–stable if, for every orbit satisfying $(p(0), q(0)) \in \mathcal{T}$, one has $|p(t) - p(0)| < \delta$ for all times t such that $|t| < T$.

Theorem: *Consider the Hamiltonian (1), real holomorphic in the phase space $\mathcal{G}_\varrho \times \mathbf{T}^n_\xi$, where $\mathcal{G} \subset \mathbf{R}^n$, and ϱ and ξ are positive constants. Assume that there are positive constants ε, m, M, V and Ω such that*

$$\sup_{(p,q)\in\mathcal{G}_\varrho\times\mathbf{T}^n_\xi} |f(p,q)| \le \varepsilon \ ; \tag{a}$$

$$\sup_{p\in\mathcal{G}_\varrho} \|A(p)v\| \le M\|v\| \quad \text{for all } v \in \mathbf{R}^n \ ; \tag{b}$$

$$\inf_{p\in\mathcal{G}_\varrho} |A(p)v\cdot v| \ge m\|v\|^2 \quad \text{for all } v \in \mathbf{R}^n \ ; \tag{c}$$

$$(d) \qquad \sup_{p\in\mathcal{G}_\varrho} \|\omega(p)\| \le \Omega \ ;$$

$$(e) \qquad \mathrm{Vol}(\omega(\mathcal{G}_{\varrho/2})) \ge V\Omega^{n-1} \ .$$

Then there exists a positive constant ε_* *such that for every* $|\varepsilon| < \varepsilon_*$ *the following holds true: there are positive constants* T_*, $C_1 \ldots C_4$ *and* $\overline{\mu}$, *a positive sequence* $\{\mathcal{D}^{(r)}\}_{r\ge0}$ *of nested domains, with* $\mathcal{D}^{(0)} = \mathcal{G}_{\varrho/2} \times \mathbf{T}^n$ *and three sequences* $\{\varepsilon_r\}_{r\ge0}$, $\{\varrho_r\}_{r\ge0}$ *and* $\{x_r\}_{r\ge0}$, *of positive numbers, monotonically decreasing to zero as*

$$\begin{aligned} x_r &= \frac{\overline{\mu} C_1^r \varepsilon_r}{\varrho_r^2} \ , \\ \varepsilon_0 &= \varepsilon \qquad , \quad \varepsilon_r < C_2^r \varepsilon_{r-1} \exp\left(-x_{r-1}^{-1/(4n)}\right) \ , \\ \varrho_0 &= \varrho \qquad , \quad \varrho_r < C_3 x_{r-1}^{1/4} \varrho_{r-1} \ , \end{aligned}$$

such that:

(i) $\mathcal{D}^{(r+1)} \subset \mathcal{D}^{(r)}$ *;*

(ii) for every r, $\mathcal{D}^{(r)}$ *is a set of* n*–dimensional tori diffeomorphic to* $\mathcal{G}^{(r)}_{\varrho_r/2} \times \mathbf{T}^n$, *where* $\mathcal{G}^{(r)}_{\varrho_r/2}$ *has the form (2);*

(iii) $\mathrm{Vol}(\mathcal{D}^{(r+1)}) > (1 - C_4\varrho_r)\,\mathrm{Vol}(\mathcal{D}^{(r)})$ *;*

(iv) $\mathcal{D}^{(\infty)} = \bigcap_{r\ge0} \mathcal{D}^{(r)}$ *is a non empty set of invariant tori for the flow* φ^t, *and moreover one has* $\mathrm{Vol}(\mathcal{D}^{(\infty)}) > \exp\left(-\frac{5}{9}C_4\varrho_0\right)\mathrm{Vol}\,\mathcal{D}^{(0)}$*;*

(v) for every $p^{(r)} \in \mathcal{G}^{(r)}$ *the torus* $p^{(r)} \times \mathbf{T}^n \subset \mathcal{D}^{(r)}$ *is* $(\varrho_r/2, T_* \exp(x_r^{-1/(4n)})/2^{2(n+1)r})$*–stable;*

(vi) for every $p^{(r)} \in \mathcal{G}^{(r)}$ *there exists an invariant torus* $\mathcal{T} \subset B_{\varrho_r}(p^{(r)}) \times \mathbf{T}^n$

We add some comments concerning the threshold ε_* for the applicability of the theorem. The question is whether the value for applicability of Nekhoroshev's theorem is already enough for performing the process of iteration or if some further more restrictive condition is necessary. Actually from the proof it turns out that the iteration of the theorem requires a condition of the form

$$x_0 := \frac{\overline{\mu}\varepsilon}{\varrho^2} < O(n^{-n}) \ ,$$

while only the weaker condition $x_0 < 1$ is strictly required for Nekhoroshev's theorem (see condition (62) below and condition (4), respectively). However, it is an easy matter to see that the stronger condition above is essentially the same that one obtains by demanding the theorem of Nekhoroshev to be not just applicable, but also meaningful. Indeed, a straightforward estimate on the original Hamiltonian gives immediately a stability time of order ϱ/ε; thus, Nekhoroshev's theorem is significant if it gives a longer stability time, i.e., if

$$\exp\left[\left(\frac{\varrho^2}{\overline{\mu}\varepsilon}\right)^{\frac{1}{2n}}\right] > \frac{\varrho}{\varepsilon} \ .$$

This, taking into account that $\overline{\mu}$ is of the form α^n, with some constant $\alpha > 1$, actually gives a condition similar to the one above, at least for what concerns the dependence on n.

Let us add a brief comment about the contents of the paper. Although there are already several proofs of Nekhoroshev's theorem in the existing literature none of them is sufficient for our purposes because we need to extract all the necessary information for the iteration. For this reason we decided to include a self consistent proof of the theorem, taking into account the improvements made by several authors since the first version by Nekhoroshev. These improvements concern mainly the so called geometric part of the theorem. Precisely, the convexity of the Hamiltonian together with the conservation of energy are used in order to bound the motion inside the resonance regions. This idea was first exploited by Benettin and Gallavotti[8], and was used by Lochak[10][11], who succeeded in finding optimal values for the exponent in the exponential estimate. Unfortunately, the elegant formulation by Lochak cannot be used here for our purposes, because it gives no information on the non–resonant region. Thus, we follow the latest exposition by Pöschel[12] and Delshams and Gutierrez[13].

The scheme of the paper is as follows. In section 2 we state the Nekhoroshev theorem in a form adapted to iteration, postponing the details of the proof to the technical section 4. Section 3 contains the proof of the main theorem. Some very technical details are contained in two appendices.

2. Iterative version of Nekhoroshev's theorem

Proposition 1: *Let $H(p,q) = h(p) + f(p,q)$ be real holomorphic in the complex domain $\mathcal{G}_{2\delta} \times \mathbf{T}^n_\sigma$, where $\mathcal{G} \subset \mathbf{R}^n$ and and σ are positive constants. Denote $\omega(p) = \partial h/\partial p$, $A(p) = \partial^2 h/\partial p \partial p$. Assume that there are positive constants ε, m, M, V and Ω such that*

$$(a) \qquad \sup_{(p,q)\in\mathcal{G}_{2\delta}\times\mathbf{T}^n_\sigma} |f(p,q)| \le \varepsilon \ ;$$

$$(b) \qquad \sup_{p\in\mathcal{G}_{2\delta}} \|A(p)v\| \le M\|v\| \quad \text{for all } v \in \mathbf{R}^n \ ;$$

$$(c) \qquad \inf_{p\in\mathcal{G}_{2\delta}} |A(p)v\cdot v| \ge m\|v\|^2 \quad \text{for all } v \in \mathbf{R}^n \ ;$$

$$(d) \qquad \sup_{p\in\mathcal{G}_{2\delta}} \|\omega(p)\| \le \Omega \ ;$$

(*e*) $$\mathrm{Vol}(\omega(\mathcal{G}_\delta)) \geq V\Omega^{n-1}\ .$$

Then there exists positive constants $\overline{\mu}$ and T such that the following holds true: if

$$\frac{\overline{\mu}\varepsilon}{\delta^2} < 1 \tag{4}$$

then

(i) for every orbit $p(t)$ with initial point $p(0) \in \mathcal{G}_\delta$ one has

$$\|p(t) - p(0)\| < \delta \qquad \text{for all } |t| < T\exp\left[\left(\frac{\delta^2}{\overline{\mu}\varepsilon}\right)^{1/(4n)}\right]\ ;$$

(ii) there exist $\mathcal{G}' \subset \mathcal{G}_\delta$, a positive $\delta' < \delta/4$ and a real holomorphic canonical transformation $(p,q) = \mathcal{C}(p',q')$ mapping $\mathcal{G}'_{2\delta'} \times \mathbf{T}^n_{\sigma/4}$ to $\mathcal{G}_\delta \times \mathbf{T}^n_{5\sigma/16}$ such that the transformed Hamiltonian is holomorphic in the complex domain $\mathcal{G}'_{2\delta'} \times \mathbf{T}^n_{\sigma/4}$ and takes the form

$$H'(p',q') = h'(p') + f'(p',q')\ .$$

(iii) the transformed Hamiltonian $H'(p',q')$ satisfies (a)–(d) with new constants ε', m', M', and Ω' given by

$$\varepsilon' = 4\varepsilon\left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right)\exp\left[-\left(\frac{\delta^2}{\overline{\mu}\varepsilon}\right)^{1/(4n)}\right] \tag{5}$$

$$M' = M\left(1 + \frac{1}{2}\sqrt{\frac{\overline{\mu}\varepsilon}{\delta^2}}\right) \tag{6}$$

$$m' = m\left(1 - \frac{1}{2}\sqrt{\frac{\overline{\mu}\varepsilon}{\delta^2}}\right) \tag{7}$$

$$\Omega' = \Omega + \frac{m}{12\sqrt{n}B^nK^n}\left(\frac{\overline{\mu}\varepsilon}{\delta^2}\right)^{3/4}\delta\ ; \tag{8}$$

(iv) the volume of $\mathcal{G}'$ satisfies

$$\mathrm{Vol}(\mathcal{G}') \geq \left(1 - \frac{2^{n+2}M^{n+1}\delta}{Vm^n(B-\sqrt{2})B^{n-1}}\right)\mathrm{Vol}\,\mathcal{G}_\delta \tag{9}$$

and the property (e) remains valid with a new constant

$$V' = V\left[1 - \left(\frac{\Omega'}{\Omega}\right)^{n-1}\frac{2^{n+3}M\delta}{(B-\sqrt{2})B^{n-1}V}\right]\left(1 + \frac{m}{12\sqrt{n}B^nK^n\Omega}\left(\frac{\overline{\mu}\varepsilon}{\delta^2}\right)^{3/4}\delta\right)^{1-n}\ . \tag{10}$$

Estimated values of $\overline{\mu}$, T and δ' are

$$
\begin{aligned}
&\overline{\mu} = \max(\lambda_*, \mu_*, \nu_*) \ , \quad T = \frac{\sigma}{4\Omega}\left(\frac{1-e^{-\sigma/4}}{1+e^{-\sigma/4}}\right)^n \ ,\\
&\delta' = \frac{1}{4\sqrt{n}B^nK^n}\left(\frac{\overline{\mu}\varepsilon}{\delta^2}\right)^{1/4}\delta \ ,
\end{aligned}
\tag{11}
$$

where

$$
\begin{aligned}
&K = \left\lceil\frac{4(1+3\ln 2)}{\sigma}\right\rceil \ , \quad \lambda_* = \frac{3m(B-\sqrt{2})B^{2n}K^{2n}}{2M^2} \ ,\\
&\mu_* = \frac{2^{10}\sqrt{n}K^{2n}B^{2n}}{MK\sigma}\left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right)^n \ , \quad \nu_* = \frac{2^6n^2B^{2n}K^{2n}}{m}\left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right)^n \ ,\\
&B = 1+\sqrt{2}+\frac{4M}{m}\max\left\{6\sqrt{n}, \left(\frac{2^4M\delta}{V}\right)^{1/n}\right\} \ .
\end{aligned}
\tag{12}
$$

3. Proof of the main theorem

Set $\mathcal{G}^{(0)} = \mathcal{G}$, $\varepsilon_0 = \varepsilon$, $\varrho_0 = \varrho$, $\xi_0 = \xi$, $m_0 = m$, $M_0 = M$, $V_0 = V$ and $\Omega_0 = \Omega$. Apply proposition 1 with ε_0, $\varrho_0/2$, ξ_0, m_0, M_0, V_0 and Ω_0 in place of ε, δ, σ, m, M, V and Ω, respectively. With these values, define the constants K_0, B_0, T_0 and $\overline{\mu}_0$ according to (11) and (12), namely

$$
\begin{aligned}
&K_0 = \left\lceil\frac{4(1+3\ln 2)}{\xi_0}\right\rceil \ , \quad T_0 = \frac{\xi_0}{4\Omega_0}\left(\frac{1-e^{-\xi_0/4}}{1+e^{-\xi_0/4}}\right)^n \ ,\\
&B_0 = 1+\sqrt{2}+\frac{4M_0}{m_0}\max\left\{6\sqrt{n}, \left(\frac{2^3M_0\varrho_0}{V_0}\right)^{1/n}\right\} \ ,\\
&\overline{\mu}_0 = (B_0K_0)^{2n}\max\left\{\frac{3m_0(B_0-\sqrt{2})}{2M_0^2}\ ,\right.\\
&\qquad\qquad \left.\frac{2^{10}\sqrt{n}}{M_0K_0\xi_0}\left(\frac{1+e^{-\xi_0/4}}{1-e^{-\xi_0/4}}\right)^n \ , \ \frac{2^6n^2}{m_0}\left(\frac{1+e^{-\xi_0/4}}{1-e^{-\xi_0/4}}\right)^n\right\} \ .
\end{aligned}
\tag{13}
$$

Then proposition 1 can be applied provided

$$
x_0 := \frac{4\overline{\mu}_0\varepsilon_0}{\varrho_0^2} < 1 \ ,
$$

which is satisfied if ε is small enough.

Now we proceed by induction. Assuming that proposition 1 has been applied $r-1$ times, for $0 \le s < r$ we have constants ε_s, ϱ_s, ξ_s, m_s, M_s, Ω_s and V_s, domains $\mathcal{D}^{(s)} = \mathcal{G}^{(s)}_{\varrho_s/2} \times \mathbf{T}^n_{\xi_s}$, and Hamiltonians $H^{(s)}$ of the form (1) real analytic in $\mathcal{G}^{(s)}_{\varrho_s} \times \mathbf{T}^n_{\xi_s}$ and satisfying the hypotheses (a)–(e) of proposition 1. In order to perform the next step we recursively define

$$
\begin{aligned}
&\varepsilon_r = 4\varepsilon_{r-1}\left(\frac{1+e^{-\xi_{r-1}/4}}{1-e^{-\xi_{r-1}/4}}\right)^n \exp\left(-x_{r-1}^{-1/(4n)}\right) \ ,\\
&\varrho_r = \frac{1}{4\sqrt{n}B^n_{r-1}K^n_{r-1}} x^{1/4}_{r-1}\varrho_{r-1} \ , \quad \xi_r = \frac{\xi_{r-1}}{4} \ ,\\
&m_r = m_{r-1}\left(1-\frac{x^{1/2}_{r-1}}{2}\right) \ , \quad M_r = M_{r-1}\left(1+\frac{x^{1/2}_{r-1}}{2}\right) \ ,\\
&\Omega_r = \Omega_{r-1} + \frac{m_{r-1}}{24\sqrt{n}B^n_{r-1}K^n_{r-1}} x^{3/4}_{r-1}\varrho_{r-1} \ ,\\
&V_r = V_{r-1}\left(1-\frac{2^{n+2}\Omega_r^{n-1}M_{r-1}\varrho_{r-1}}{\Omega^{n-1}_{r-1}(B_{r-1}-\sqrt{2})B^{n-1}_{r-1}V_{r-1}}\right)\left(1+\frac{m_{r-1}x^{3/4}_{r-1}\varrho_{r-1}}{24\sqrt{n}B^n_{r-1}K^n_{r-1}\Omega_{r-1}}\right)^{1-n} \ ,
\end{aligned}
\tag{14}
$$

with

$$
x_r = \frac{4\overline{\mu}_r\varepsilon_r}{\varrho_r^2}
\tag{15}
$$

and with constants $\overline{\mu}_r$, B_r and K_r defined according to (11) and (12) as

$$
\begin{aligned}
&K_r = \left\lceil \frac{4(1+3\ln 2)}{\xi_r} \right\rceil \ ,\\
&B_r = 1+\sqrt{2}+\frac{4M_r}{m_r}\max\left\{6\sqrt{n}, \left(\frac{2^3M_r\varrho_r}{V_r}\right)^{1/n}\right\} \ ,\\
&\overline{\mu}_r = (B_rK_r)^{2n}\max\Bigg\{\frac{3m_r(B_r-\sqrt{2})}{2M_r^2} \ ,\\
&\qquad \frac{2^{10}\sqrt{n}}{M_rK_r\xi_r}\left(\frac{1+e^{-\xi_r/4}}{1-e^{-\xi_r/4}}\right)^n , \ \frac{2^6n^2}{m_r}\left(\frac{1+e^{-\xi_r/4}}{1-e^{-\xi_r/4}}\right)^n\Bigg\} \ .
\end{aligned}
\tag{16}
$$

If x_0 is small enough then the sequences x_r, ε_r and ϱ_r decrease to zero, the sequences M_r and Ω_r are monotonically increasing, but bounded, and the sequences m_r and V_r are monotonically decreasing, but bounded far from zero. More precisely, we prove that for every $r \ge 0$ one has

$$
m_r > m_\infty \ , \quad M_r < M_\infty \ , \quad \Omega_r < \Omega_\infty \ , \quad V_r > V_\infty \ ,
\tag{17}
$$

with some positive constants m_∞, M_∞, Ω_∞ and V_∞. This is quite evident form the form of the sequences. Since we want also to produce explicit estimates of the constants, we report the detailed computation in appendix A. An explicit condition on x_0 is given there by (62). We also prove that we can set

$$m_\infty = \tfrac{9}{10} m_0 \ , \quad M_\infty = \tfrac{10}{9} M_0 \ , \quad \Omega_\infty = (1+10^{-4})\Omega_0 \ , \quad V_\infty = \tfrac{5}{6} V_0 \ . \tag{18}$$

Thus, proposition 1 can be applied for every r.

Now we go back to the statement of the theorem. The sequences x_r, ε_r, ϱ_r are given by (15) and (14), so it is just matter of finding the constants C_1, C_2 and C_3, that clearly exist. Explicit values are computed in appendix A as

$$C_1 = 2^{7n+1}\lambda^{2n+1} \ , \quad C_2 = \frac{2^{3n+2}}{(1-e^{-\xi_0/4})^n} \ , \quad C_3 = \frac{1}{4\sqrt{n}B_0^n K_0^n} \ . \tag{19}$$

The statements (i) and (ii) are consequence of the statement (ii) of proposition 1, by identifying $\mathcal{D}^{(r)}$ with $\mathcal{G}_\delta \times \mathbf{T}^n$ and $\mathcal{D}^{(r+1)}$ with $\mathcal{G}'_{\delta'} \times \mathbf{T}^n$. The statement (iii) follows from the estimate (9), taking into account that $\mathrm{Vol}(\mathcal{D}^{(r)}) = (2\pi)^n \,\mathrm{Vol}(\mathcal{G}_\delta)$ and $\mathrm{Vol}(\mathcal{D}^{(r+1)}) > (2\pi)^n \,\mathrm{Vol}(\mathcal{G}')$, with the constant C_4 estimated by

$$C_4 = \frac{2^{n+2} M_\infty^{n+1}}{V_\infty m_\infty^n (B_0 - \sqrt{2}) B_0^{n-1}} \ .$$

Postponing the proof of (iv), we come now to (v). This follows from the statement (i) of proposition 1, just replacing δ with $\varrho_r/2$ and $\delta^2/(\overline{\mu}\varepsilon)$ with x_r. From (11) one has

$$T > \frac{\xi_r}{4\Omega_\infty}\left(\frac{1-e^{-\xi_0/4}}{1+e^{-\xi_0/4}}\right)^n = \frac{T_*}{[2^{2(n+1)}]^r}$$

where

$$T_* = \frac{\xi_0}{4\Omega_\infty}\left(\frac{1-e^{-\xi_0/4}}{1+e^{-\xi_0/4}}\right) \ ;$$

In order to prove (iv) we first show that the volume of $\mathcal{D}^{(\infty)}$ is positive. To this end, we iterate (9) as

$$\mathrm{Vol}(\mathcal{D}^{(r+1)}) \geq \prod_{s=0}^{r}\left(1 - \frac{2^{n+1} M_s^{n+1}\varrho_s}{V_s m_s^n (B_s - \sqrt{2}) B_s^{n-1}}\right) \mathrm{Vol}(\mathcal{D}^{(0)}) \ ,$$

and recall that the factor in parenthesis is always positive. However, we should prove that the infinite product is not zero. To this end we should prove that the sequence $\{\varrho_s\}$ decreases to zero fast enough. From the explicit estimates in appendix A we get indeed

$$\sum_{s=0}^{r} \varrho_s < \tfrac{10}{9}\varrho_0 \ . \tag{20}$$

Using this inequality and the definition above of the constant C_4 we get the estimate on the volume of $\mathcal{D}^{(\infty)}$ in (iv).

We now prove that $\mathcal{D}^{(\infty)}$ is a set of invariant KAM tori. Let $\tilde{\mathcal{D}}^{(r)} = \mathcal{G}^{(r)}_{\varrho_r} \times \mathbf{T}^n$, so that $\mathcal{D}^{(r)} \subset \tilde{\mathcal{D}}^{(r)}$. Since $\varrho_r \to 0$, we clearly have $\mathcal{D}^{(\infty)} = \bigcap_{r\ge 0} \tilde{\mathcal{D}}^{(r)}$. Let now $x \in \mathcal{D}^{(\infty)}$, so that $x \in \mathcal{D}^{(r)}$ for every r. Denote by $(p^{(r)}, q^{(r)})$ the coordinates of x in $\mathcal{G}^{(r)}_{\varrho_r/2} \times \mathbf{T}^n$. Then, by (i) of proposition 1, one has $\varphi^t x \in \tilde{\mathcal{D}}^{(r)}$ for $t < T_r$, say, with T_r monotonically increasing to infinity. Since r is arbitrary and $\tilde{\mathcal{D}}^{(s)} \subset \tilde{\mathcal{D}}^{(r)}$ for $s > r$, we also have $\varphi^t x \in \tilde{\mathcal{D}}^{(r)}$ for $|t| < T_s$ for all s, i.e., $\varphi^t x \in \tilde{\mathcal{D}}^{(r)}$ for all t and for all r. Thus, by definition of $\mathcal{D}^{(\infty)}$, we get $\varphi^t x \in \mathcal{D}^{(\infty)}$ for all t. This shows that the set $\mathcal{D}^{(\infty)}$ is invariant for the Hamiltonian flow. By the form of the Hamiltonian in $\tilde{\mathcal{D}}^{(r)}$ we also have $\mathrm{dist}(\varphi^t x, (p^{(r)}, q^{(r)} + \omega^{(r)}(p^{(r)})t)) < O(\varepsilon_r^{1/2})$ for $|t| < O(\varepsilon_r^{-1/2})$, where $\omega^{(r)}(p^{(r)})$ is the unperturbed frequency at step r. On the other hand, by construction we have $\omega^{(r)} \to \omega^{(\infty)}$ with $\omega^{(\infty)}$ non resonant and $|\omega^{(r)} - \omega^{(\infty)}| < O(\varepsilon_r)$. Thus we also have $\mathrm{dist}(\varphi^t x, (p^{(r)}, q^{(r)} + \omega^{(\infty)}(p^{(r)})t)) < O(\varepsilon_r^{1/2})$ for $|t| < O(\varepsilon_r^{-1/2})$. Since r is arbitrary, we conclude that the orbit $\varphi^t x$ is dense on a torus, which, by invariance belongs to $\mathcal{D}^{(\infty)}$. The same trivially holds for the closure of $\varphi^t x$, namely for the torus.

Concerning (vi), it is enough to remark that the theorem can be applied to every open ball $B_{\varrho_r}(p^{(r)}) \times \mathbf{T}^n$. This concludes the proof.

4. Proof of proposition 1

We follow the traditional scheme of proof of Nekhoroshev's theorem, composed by an analytic part and of a geometric part. Although several proofs are already available in the literature there is a problem of determining explicit values for all the constants which makes it unpleasant to reconstruct the proof making reference to several papers. For this reason we give here an essentially self contained exposition.

4.1 Technical settings.

In this section we introduce domains and norms to be used in the rest of the proofs. We start with the definition of resonance module. This is defined as a subgroup $\mathcal{M} \in \mathbf{Z}^n$ satisfying $\mathrm{span}(\mathcal{M}) \cap \mathbf{Z}^n = \mathcal{M}$; here both $\mathcal{M}$ and $\mathbf{Z}^n$ are considered as subsets of $\mathbf{R}^n$, and $\mathrm{span}(\mathcal{M})$ is the linear subspace in $\mathbf{R}^n$ generated by $\mathcal{M}$. The integer $\dim(\mathrm{span}(\mathcal{M}))$ will be called the *dimension* or *multiplicity* of the resonance module.

A set $\mathcal{W} \subset \mathbf{R}^n$ will be called a *non–resonance domain of type* $(\mathcal{M}, \alpha, \delta, N)$ in case

$$|k \cdot \omega(p)| > \alpha \quad \text{for all } p \in \mathcal{W}_\delta \ , \quad k \in \mathbf{Z}^n \setminus \mathcal{M} \text{ and } |k| \le N \ .$$

Here, $|k| = |k_1| + \ldots |k_n|$, α and δ are real positive parameters, N a positive integer and

$\mathcal{M}$ a resonance module. For any point $p \in \mathcal{W}$ the *plane of fast drift* $P_{\mathcal{M}}(p)$ is defined as

$$P_{\mathcal{M}}(p) = \{p' \in \mathcal{W} \, : \, p' - p \in \mathrm{span}(\mathcal{M})\} \tag{21}$$

In the rest of this section we recall now some known facts about generalization of Cauchy's estimates, canonical transformations defined through the Lie transform algorithm and normal forms. Our aim is only to establish a clear technical setting, without entering unnecessary details. An introductory exposition can be found for instance in [15].

Let $f(p,q)$ be a real analytic function on a domain $\mathcal{G} \times \mathbf{T}$, with an holomorphic extension to a complex domain $\mathcal{G}_\delta \times \mathbf{T}^n_{\sigma'}$ for some positive δ, σ', and consider the Fourier expansion

$$f(p,q) = \sum_k f_k(p) \exp(ik \cdot q) \ .$$

We say that f is in normal form with respect to a resonance module $\mathcal{M}$ in case the Fourier expansion of f has the form

$$f(p,q) = \sum_{k \in \mathcal{M}} f_k(p) \exp(ik \cdot q) \ ,$$

i.e., it contains only harmonics belonging to $\mathcal{M}$. In particular, if $\mathcal{M} = \{0\}$, then one has $f = f(p)$. The weighted Fourier norm of f is defined as

$$\|f\|_{\delta,\sigma} = \sum_k |f_k|_\delta \exp(|k|\sigma) \ ,$$

$|f_k|_\delta$ denoting the usual supremum norm over the complex domain $\mathcal{G}_\delta$. The definition is consistent for every non negative $\sigma < \sigma'$ in view of the known property of exponential decay of coefficients in the Fourier expansion of an analytic function.

The following properties generalize the Cauchy estimates for derivatives of analytic functions to the case of the weighted Fourier norm. If f and g have bounded norms $\|f\|_{\delta,\sigma}$ and $\|g\|_{(1-d')(\delta,\sigma)}$ with some non negative $d' < 1$, then:

(i) for $0 < d < 1$ and $1 \le j \le n$, one has

$$\left\| \frac{\partial f}{\partial p_j} \right\|_{(1-d)(\delta,\sigma)} \le \frac{1}{d\delta} \|f\|_{(\delta,\sigma)} \ , \quad \left\| \frac{\partial f}{\partial q_j} \right\|_{(1-d)(\delta,\sigma)} \le \frac{1}{ed\sigma} \|f\|_{(\delta,\sigma)} \ , \tag{22}$$

(ii) for $0 < d < 1 - d'$ one has

$$\|\{f,g\}\|_{(1-d'-d)(\delta,\sigma)} \le \frac{2}{ed(d+d')\delta\sigma} \|f\|_{\delta,\sigma} \|g\|_{(1-d')(\delta,\sigma)} \ . \tag{23}$$

Similarly, let $g_1, \ldots, g_s$ and f be analytic functions with bounded norms $\|g_j\|_{\delta,\sigma}$ and $\|f\|_{\delta,\sigma}$. Then for every positive $d < 1$ one has

$$(24)\qquad \|L_{g_s} \circ \ldots \circ L_{g_1} f\|_{(1-d)(\delta,\sigma)} \le \frac{s!}{e^2} \left(\frac{2e}{d^2\delta\sigma}\right)^s \|g_1\|_{(\delta,\sigma)} \cdot \ldots \cdot \|g_s\|_{(\delta,\sigma)} \|f\|_{(\delta,\sigma)} \ .$$

Here, we used the common notation $L_\chi \cdot = \{\chi, \cdot\}$, where $\{\cdot,\cdot\}$ denotes the Poisson bracket.

For a fixed non–negative integer N we shall denote by $\mathcal{F}_N$ the class of functions $f(p,q)$ which are trigonometric polynomials in q of degree N, i.e., admit a truncated Fourier expansion $f(p,q) = \sum_{|k|\le N} f_k(p) \exp(ik\cdot q)$.

With these settings, and having fixed a positive integer K, we expand the Hamiltonian as

$$(25)\qquad H(p,q) = H_0(p) + H_1(p,q) + \ldots \ ,$$

where $H_0(p) = h(p)$ and $H_s(p,q) \in \mathcal{F}_{sK}$ for $s \ge 0$. A convenient way is the following: referring to the Fourier expansion of $f(p,q)$ in the Hamiltonian, define $H_s(p,q) = \sum_{(s-1)K\le|k|<sK} f_k(p)\exp(ik\cdot q)$. With this definition, and using the property of exponential decay of coefficients, it is an easy matter to find the estimates on the norms

$$(26)\qquad \|H_s\|_{\delta,\sigma/2} \le e^{-(s-1)K\sigma/4} |f|_{\delta,\sigma} \left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right)^n$$

(see [14], lemma 8).

Having given a generating sequence $\chi = \{\chi_r\}_{r\ge1}$ of analytic functions, we define a canonical transformation via the linear operator

$$(27)\qquad T_\chi = \sum_{s\ge0} E_s \ , \quad \text{where } E_0 = \mathrm{Id} \ , \quad E_s = \sum_{j=1}^{s} \frac{j}{s} L_{\chi_j} E_{s-j} \ .$$

The operator T_χ is invertible, and an explicit form for T_χ^{-1} is:

$$(28)\qquad T_\chi^{-1} = \sum_{s\ge0} D_s \ , \quad \text{where } D_0 = \mathrm{Id} \ , \quad D_s = -\sum_{j=1}^{s} \frac{j}{s} D_{s-j} L_{\chi_j} \ .$$

Writing the canonical transformation as $p' = T_\chi p$, $q' = T_\chi q$, one has that a function $f'(p',q')$ is transformed to $f(p,q) = \bigl(T_\chi f'\bigr)(p,q)$. Similarly, for the inverse transformation $p = T_\chi^{-1} p'$, $q = T_\chi^{-1} q'$ one has $f'(p',q') = \bigl(T_\chi^{-1} f\bigr)(p',q')$.

A condition for convergence of the canonical transformation is the following: assume

$$(29)\qquad \|\chi_s\|_{\delta,\sigma} \le \frac{\beta^{s-1}}{s} G$$

with some positive β and G; assume also the convergence condition

$$\frac{2eG}{d^2\delta\sigma}+\beta\le\frac{1}{2}\ ; \tag{30}$$

then, for every positive $d<1/2$ both the operators T_χ and T_χ^{-1} define an analytic canonical transformation mapping the domain $\mathcal{G}_{(1-d)\delta}\times\mathbf{T}^n_{(1-d)\sigma}$ to $\mathcal{G}_\delta\times\mathbf{T}^n_\sigma$, with the properties

$$|p-p'|<d\delta\ ,\quad |q-q'|<d\sigma\ . \tag{31}$$

This claim can be proved using the estimates

$$\begin{aligned}\|E_s f\|_{(1-d)(\delta,\sigma)}&\le\left(\frac{2eG}{d^2\delta\sigma}+\beta\right)^s\|f\|_{\delta,\sigma}\\ \|D_s f\|_{(1-d)(\delta,\sigma)}&\le\left(\frac{2eG}{d^2\delta\sigma}+\beta\right)^s\|f\|_{\delta,\sigma}\ ,\end{aligned} \tag{32}$$

which holds for every positive $d<1$.

4.2 The Analytic part: local normal forms.

The scheme of the analytic part is the following: one considers a non–resonance domain $\mathcal{W}$ of type $(\mathcal{M},\alpha,\delta,N)$ with $N=rK$, r and K being positive integers. On the non resonance domain, one performs a canonical transformation which gives the Hamiltonian (25) the form

$$H^{(r)}=H_0+Z_1+\ldots+Z_r+\mathcal{R}^{(r)}\ ,\quad \mathcal{R}^{(r)}=\sum_{s>r}H_s^{(r)} \tag{33}$$

where $Z_j\in\mathcal{F}_{jK}$ is in normal form with respect to $\mathcal{M}$, $H_s^{(r)}\in\mathcal{F}_{sK}$, and $\mathcal{R}^{(r)}$ is a non–normalized remainder.

Denoting by p',q' the new canonical variables, it is an easy matter to check that any function $I_\lambda=\lambda\cdot p'$ with $\mathrm{span}(\mathcal{M})\perp\lambda\in\mathbf{R}^n$ is a first integral for the normalized Hamiltonian $H_0+Z_1+\ldots Z_r$. Thus, every orbit $(p'(t),q'(t))$ with $p'(0)\in\mathcal{W}$ lies on the plane of fast drift $P_{\mathcal{M}}(p'(0))$ until it leaves $\mathcal{W}$, if ever.

In the original variables p,q the picture of the flow is described by the following proposition, which constitutes the analytic part of Nekhoroshev's theorem.

Proposition 2: *Let $\mathcal{W}$ be a non–resonance domain of type $(\mathcal{M},\alpha,2\delta,N)$ with $N=rK$, r and K being positive integers and with some positive δ. Assume that the Hamiltonian $H=h(p)+f(p,q)$ is analytic in the domain $\mathcal{W}_{2\delta}\times\mathbf{T}^n_\sigma$ for some positive σ, and that the hypotheses (a) and (d) of the theorem are satisfied. Assume moreover that for every $p(0)\in\mathcal{W}$ there is $\omega^*\perp\mathcal{M}$ such that $\|\omega(p(0))-\omega^*\|<\delta_{\mathcal{M}}$ for some positive $\delta_{\mathcal{M}}$, and* *that*

$$\delta > \frac{2}{m}\left(\delta_{\mathcal{M}} + \sqrt{\delta_{\mathcal{M}}^2 + 6m\varepsilon}\right) \ . \tag{34}$$

Then the following holds true: there exists a constant T such that if

$$\mu := \left(\frac{2^{10}r\mathcal{E}}{e\alpha\delta\sigma} + 4e^{-K\sigma/4}\right) \le \frac{1}{2} \ , \quad \mathcal{E} = \varepsilon\left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right)^n \ , \tag{35}$$

then one has

$$|p(t) - p(0)| \le \delta \quad \text{for all } |t| < T\mu^{-r} \ .$$

An estimated value of T is:

$$T = \frac{\sigma}{4\Omega}\left(\frac{1-e^{-\sigma/4}}{1+e^{-\sigma/4}}\right)^n \ . \tag{36}$$

In order to be able to iterate the theorem of Nekhoroshev, we need some additional informations concerning non–resonance domains characterized by the module $\mathcal{M} = \{0\}$, i.e., completely non–resonance regions in action space. This is given by the following

Proposition 3: *Let $\mathcal{W}$ be a non–resonance domain of type $(\{0\}, \alpha, 2\delta, N)$ with $N = rK$. Assume that the Hamiltonian $h(p) + f(p,q)$ satisfies the hypotheses (a), (b) and (c) of the theorem. Assume moreover the hypothesis (35) of proposition 2 and the further condition*

$$\mathcal{E} < \frac{m\delta^2}{16n} \ . \tag{37}$$

Then there exists a canonical transformation $(p,q) = \mathcal{C}(p',q')$ mapping $\mathcal{W}_\delta \times \mathbf{T}^n_{\sigma/4} \to \mathcal{W}_{3\delta/2} \times \mathbf{T}^n_{3\sigma/8}$ and satisfying

$$|p - p'| < \frac{\delta}{4} \ , \quad |q - q'| < \frac{\sigma}{16} \tag{38}$$

which gives the Hamiltonian the form $H'(p',q') = h'(p') + f'(p',q')$; the Hamiltonian h' satisfies the hypotheses (a), (b), (c) and (d) with new constants ε', M', m' and Ω' estimated as

$$\varepsilon' = 4\mu^r\mathcal{E} \tag{39}$$

$$M' = M + \frac{2^3 n\mathcal{E}}{\delta^2} \tag{40}$$

$$m' = m - \frac{2^3 n\mathcal{E}}{\delta^2} \tag{41}$$

$$\Omega' = \Omega + \frac{8\sqrt{n}}{3\delta}\mathcal{E} \ . \tag{42}$$

The proof is deferred to appendix B.

4.3 The geometric part: geography of resonances

The aim of the geometric part can be briefly described as follows. One is given an open set $\mathcal{V} \in \mathbf{R}^n$ to be extended to a complex domain $\mathcal{V}_\delta$, an integer N specifying the highest resonance to be taken into account, and a real analytic map $\omega : \mathcal{V}_\delta \to \mathbf{C}^n$ satisfying

$$(43) \qquad \|A(p)v\| \le M\|v\| \ , \quad |A(p)v\cdot v| \ge m\|v\|^2 \quad \text{for all } v \in \mathbf{R}^n \ ,$$

where $A(p) = \partial\omega/\partial p$. Moreover, we assume the further condition

$$(44) \qquad \mathrm{Vol}(\omega(\mathcal{V})) > V\Omega^{n-1} \ , \quad \Omega = \sup_{p\in\mathcal{V}_\delta} \|\omega(p)\| \ ,$$

for some positive V. The complex extension is needed only in order to introduce non–resonance domains. Indeed the problem of the geometric part is precisely to construct a family of non–resonance domains of type $(\mathcal{M}, \alpha, \delta_{\mathcal{M}}, N)$ parameterized by a family of suitable non–resonance modules $\mathcal{M}$ and constants α and $\delta_{\mathcal{M}}$ depending on $\mathcal{M}$. This family of domains must be a covering of $\mathcal{V}$. The construction is made in the space $\mathbf{R}^n$ of frequencies and then mapped back to the action space $\mathcal{V}$ via the frequency mapping $\omega(p)$.

We start with some definitions.

1) N–moduli and resonance parameters. If $\mathcal{M} \subset \mathbf{Z}^n$ is a resonance module, we denote $\mathcal{M}_N = \{k \in \mathcal{M} \ : \ |k| \le N\}$, and call a module $\mathcal{M}$ with $\dim(\mathcal{M}) = s$ a N–module in case $\mathcal{M}_N$ contains s independent vectors $k \in \mathbf{Z}^n$. To the N–moduli we associate positive parameters $\beta_0 < \beta_1 < \ldots < \beta_n$, the index referring to the dimension of $\mathcal{M}$. They will be determined below. Moreover, we shall denote by $|\mathcal{M}|$ the volume of the parallelepiped generated by any basis of $\mathcal{M}$.

2) Resonant plane and resonant zone. To a N–module $\mathcal{M}$ we associate the *resonant plane* defined as

$$R_{\mathcal{M}} = \{\omega \in \mathbf{R}^n : k\cdot\omega = 0 \ \forall k \in \mathcal{M}\} \ .$$

The *resonant zone* is defined as

$$\mathcal{Z}_{\mathcal{M}} = \bigcup_{\omega\in R_{\mathcal{M}}} \{\omega' \in \mathbf{R}^n \ : \ \|\omega' - \omega\| < \delta_{\mathcal{M}}\} \ ,$$

where $\delta_{\mathcal{M}} = \beta_s/|\mathcal{M}|$.

3) Resonant region of order s. To the family of all N–moduli of fixed $\dim(\mathcal{M}) = s$ we associate the *resonant region*

$$\mathcal{Z}^*_s = \bigcup_{\dim\mathcal{M}=s} \mathcal{Z}_{\mathcal{M}} \ .$$

By definition, $\mathcal{Z}^*_0 = \mathbf{R}^n$; we also set $\mathcal{Z}^*_{n+1} = \emptyset$.

4) Resonant block. To a N–module $\mathcal{M}$ of $\dim\mathcal{M} = s$ we associate the *resonant block*

$$\mathcal{B}_{\mathcal{M}} = (\mathcal{Z}_{\mathcal{M}} \setminus \mathcal{Z}^*_{s+1}) \ .$$

We now give the following

Lemma 1: *Let the parameters $\beta_1, \ldots, \beta_n$ satisfy*

$$\beta_s > BN\beta_{s-1} \ , \quad s = 1, \ldots, n$$

with a constant $B > \sqrt{2}$. Then the following holds true:

i) the resonant blocks are a covering of $\mathbf{R}^n$, i.e.

$$\mathbf{R}^n = \bigcup_{\mathcal{M}} \mathcal{B}_{\mathcal{M}}$$

the union being over the set of all N–moduli.

ii) For every N–module $\mathcal{M}$ and for every $\omega \in \mathcal{B}_{\mathcal{M}}$ one has

$$|k \cdot \omega| > \alpha_{\mathcal{M}} \text{ for } k \notin \mathcal{M} \ , \ |k| \leq N \ ,$$

with $\alpha_{\mathcal{M}} = (B - \sqrt{2})N\delta_{\mathcal{M}}$.

The property i) is a trivial consequence of the definition. For the proof of ii) see [12], geometric lemma, p. 201.

The main result of the geometric part is given by the following

Proposition 4: *Consider an open domain $\mathcal{V} \subset \mathbf{R}^n$ with a frequency map $\omega(p)$ satisfying (43) and (44) on $\mathcal{V}_\delta$ for some positive δ. Let also N be a positive integer, and define*

$$\beta_0 = \frac{2M\delta}{(B - \sqrt{2})B^n N^n} \ , \quad \beta_s = BN\beta_{s-1} \ , \quad 0 < s \leq n \ , \tag{45}$$

with a constant $B > \sqrt{2}$. Let also

$$\mathcal{W}_{\mathcal{M}} = \{p \in \mathcal{V} \ : \ \omega(p) \in \mathcal{B}_{\mathcal{M}}\} \ .$$

Thus the following holds true:

i) the domains $\mathcal{W}_{\mathcal{M}}$ cover $\mathcal{V}$, i.e.

$$\mathcal{V} = \bigcup_{\mathcal{M}} \mathcal{W}_{\mathcal{M}}$$

the union being over the set of all N–moduli;

ii) for every $p \in \mathcal{W}_{\mathcal{M}}$ there exists $\omega^* \perp \mathcal{M}$ such that

$$\|\omega(p) - \omega^*\| < \delta_{\mathcal{M}} \ .$$

iii) for every N–module $\mathcal{M}$, $\mathcal{W}_{\mathcal{M}}$ is a non–resonant domain of type

$$(\mathcal{M}, \frac{\alpha_{\mathcal{M}}}{2}, \frac{\alpha_{\mathcal{M}}}{2\sqrt{n}MN}, N)$$

with $\alpha_{\mathcal{M}} = (B-\sqrt{2})N\delta_{\mathcal{M}}$; moreover one has $\alpha_{\mathcal{M}}/(2\sqrt{n}MN) < \delta$.

Proof. i) and ii) are a trivial consequence of the construction. In order to prove iii), let $p \in \mathcal{W}_{\mathcal{M},\alpha_{\mathcal{M}}/(2MN)}$; then there exists $p_0 \in \mathcal{W}_{\mathcal{M}}$, with $|p-p_0| < \alpha_{\mathcal{M}}/(2\sqrt{n}NM)$. So, for $k \notin \mathcal{M}$, $|k| \le N$, we get

$$\begin{aligned} |k\cdot\omega(p)| &\ge |k\cdot\omega(p_0)| + |k\cdot(\omega(p)-\omega(p_0))| \\ &\ge \alpha_{\mathcal{M}} - \sqrt{n}NM\|p-p_0\| > \frac{\alpha_{\mathcal{M}}}{2}\ . \end{aligned}$$

as claimed. The inequality $\alpha_{\mathcal{M}}/(2\sqrt{n}MN) < \delta$ is true in view of the definition of β_0 in (45). *Q.E.D.*

We finally state the following lemma on the volume of the non–resonance domain $\mathcal{W}_0$.

Lemma 2: *With the same hypotheses of proposition 4, but with the stronger condition on the constant B*

$$B > \sqrt{2} + \frac{4M}{m}\left(\frac{4M\delta}{V}\right)^{1/n}\ ,$$

the nonresonant part $\mathcal{W}_0$ of $\mathcal{V}$ satisfies

$$\mathrm{Vol}(\mathcal{W}_0) > \left(1 - \frac{2^{n+2}M^{n+1}\delta}{Vm^nB^{n-1}(B-\sqrt{2})}\right)\mathrm{Vol}(\mathcal{V})\ . \tag{46}$$

Remark that in view of the condition on B the constant in parentheses is always positive.

Proof. Consider the image $\omega(\mathcal{V})$ through the frequency map $\omega(p)$. We first estimate $\mathrm{Vol}\left(\mathcal{Z}_1^*\cap\omega(\mathcal{V})\right)$. To this end, let $\mathcal{M}$ be any resonant N–module with $\dim(\mathcal{M}) = 1$. By definition, we have

$$\mathcal{Z}_{\mathcal{M}}\cap\omega(\mathcal{V}) \subset \bigcup_{\omega\in R_{\mathcal{M}}\cap\omega(\mathcal{V})} B_{\delta_{\mathcal{M}}}(\omega)\ ,$$

where $R_{\mathcal{M}}$ is the $(n-1)$ dimensional plane associated to $\mathcal{M}$. In view of the second condition in (44) on ω we have $\mathrm{diam}(R_{\mathcal{M}}\cap\omega(\mathcal{V})) < \Omega$, and so also

$$\mathrm{Vol}(\mathcal{Z}_{\mathcal{M}}\cap\omega(\mathcal{V})) \le 2\Omega^{n-1}\delta_{\mathcal{M}} < 2\Omega^{n-1}\beta_1\ . \tag{47}$$

Using the known fact that the number of integer vectors k satisfying $|k| \le N$ does not exceed 2^nN^{n-1}, we immediately get, using (44)

$$\mathrm{Vol}(\mathcal{Z}_1^*\cap\omega(\mathcal{V})) < 2^{n+1}\Omega^{n-1}N^{n-1}\beta_1 < \frac{2^{n+1}N^{n-1}}{V}\beta_1\,\mathrm{Vol}(\omega(\mathcal{V}))\ .$$

Mapping back everything to $\mathcal{V}$, we estimate

$$\mathrm{Vol}(\omega^{-1}(\mathcal{Z}_1^* \cap \omega(\mathcal{V}))) < \frac{2^{n+1}N^{n-1}M^n}{Vm^n}\beta_1 \,\mathrm{Vol}(\mathcal{V}) \ .$$

Thus, remarking that $\mathcal{W}_0 = \mathcal{V} \setminus \omega^{-1}(\mathcal{Z}_1^* \cap \omega(\mathcal{V}))$ and substituting β_1 as given by (45) we get (46). *Q.E.D.*

4.4 Choice of the parameters and conclusion of the proof

We first obtain estimates depending on the parameters r, the order of normalization, and K, the initial Fourier cutoff. The exponential estimate of Nekhoroshev will come out from a proper choice of the latter parameters. Thus, let us fix $r \geq 1$, $K \geq 1$, $N = rK$ and an arbitrary constant B satisfying the condition of lemma 2. Considering the domain $\mathcal{G}_{2\delta}$ we apply the geometric proposition 4 to $\mathcal{V} = \mathcal{G}_\delta$. The hypotheses (43) and (44) are satisfied in view of (b)–(e) of proposition 1. Thus, we cover the domain $\mathcal{G}_\delta$ with a family of nonresonance domains $\mathcal{W}_\mathcal{M}$ of type $(\mathcal{M}, \alpha_\mathcal{M}/2, \alpha_\mathcal{M}/(2\sqrt{n}MN), N)$, with $\mathcal{M}$ running over the set of all N–moduli. We recall that the constants $\alpha_\mathcal{M}$ are defined together with the constants $\delta_\mathcal{M}$ and $\beta_0, \ldots, \beta_n$ as

$$\begin{aligned}
\alpha_\mathcal{M} &= (B - \sqrt{2})rK\delta_\mathcal{M} \ ,\\
\delta_\mathcal{M} &= \frac{\beta_s}{|\mathcal{M}|} \ , \quad s = \dim \mathcal{M} \ ,\\
\beta_0 &= \frac{2M\delta}{(B - \sqrt{2})B^n r^n K^n} \ ,\\
\beta_s &= BrK\beta_{s-1} \ .
\end{aligned} \tag{48}$$

To every domain we apply the analytic proposition 2, remarking that the parameters α and δ must be replaced by $\alpha_\mathcal{M}/2$ and $\alpha_\mathcal{M}/(4\sqrt{n}MrK)$, respectively, depending on the resonance module $\mathcal{M}$, while $\delta_\mathcal{M}$ is given by (48). Having this in mind, conditions (35) and (34) become

$$\mu_\mathcal{M} := \left(\frac{2^{13}\sqrt{n}r^2MK\mathcal{E}}{e\alpha_\mathcal{M}^2\sigma} + 4e^{-K\sigma/4}\right) \leq \frac{1}{2} \tag{49}$$

and

$$\frac{\alpha_\mathcal{M}}{4\sqrt{n}MrK} > \frac{2}{m}\left(\delta_\mathcal{M} + \sqrt{\delta_\mathcal{M}^2 + 6m\varepsilon}\right) \ , \tag{50}$$

with ε given by condition (a) of proposition 1. With these conditions proposition 2 gives the estimate

$$|p(t) - p(0)| < \frac{\alpha_\mathcal{M}}{4\sqrt{n}MrK} \quad \text{for } |t| < T\mu_\mathcal{M}^{-r} \ , \tag{51}$$

with T given by (36). With $\alpha_{\mathcal{M}}$ as in (48), we replace condition (50) by

$$\frac{(B-\sqrt{2})}{4\sqrt{n}M} > \frac{2}{m}\left[1+\sqrt{1+\frac{6m\varepsilon}{\delta_{\mathcal{M}}^2}}\right] ,$$

which is satisfied by imposing

$$\frac{6m\varepsilon}{\delta_{\mathcal{M}}^2} \le 1 , \quad B > \sqrt{2} + \frac{24\sqrt{n}M}{m} . \tag{52}$$

We look now for uniform estimates with respect to $\mathcal{M}$, and therefore uniform also in $\mathcal{G}_\delta$. Using the trivial inequality $\delta_{\mathcal{M}} > \beta_0$, the first of (52) is replaced by

$$\frac{\lambda_* \varepsilon r^{2n}}{\delta^2} \le 1 , \quad \lambda_* = \frac{6m(B-\sqrt{2})^2 B^{2n} K^{2n}}{4M^2} . \tag{53}$$

Concerning the condition (49), using

$$\alpha_{\mathcal{M}} > \alpha_0 := \frac{2M\delta}{B^n r^{n-1} K^{n-1}}$$

we immediately get

$$\mu_{\mathcal{M}} < \mu_0 := \frac{\mu_* r^{2n}\varepsilon}{2e\delta^2} + 4e^{-K\sigma/4} , \quad \mu_* = \frac{2^{12}\sqrt{n}K^{2n-1}B^{2n}}{M\sigma}\left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right) , \tag{54}$$

so that condition (49) can be replaced by $\mu_0 \le 1/2$. We come now to the choice of r and K. To this end we determine r and K by the conditions

$$4e^{-K\sigma/4} < \frac{1}{2e} , \quad \overline{\mu}\frac{r^{2n}\varepsilon}{\delta^2} < 1 , \quad \text{with} \quad \overline{\mu} = \max(\mu_*, \lambda_*) ,$$

with λ_* as in (53); notice that this implies $\mu_0 \le 1/e$, so that the condition above on μ_0 is satisfied. We set

$$K = \left\lceil \frac{4(1+3\ln 2)}{\sigma} \right\rceil , \quad r = \left\lfloor \left(\frac{\delta^2}{\overline{\mu}\varepsilon}\right)^{1/(4n)} \right\rfloor , \tag{55}$$

where $\lfloor\cdot\rfloor$ denotes the integer part, and $\lceil\cdot\rceil$ the integer part plus 1. The choice $1/(4n)$ instead of the more natural $1/(2n)$ will be necessary below, in proving statement (iii).

Since r must be a positive integer, we have the condition $\overline{\mu}\varepsilon < \delta^2$, which is assumed in the statement of proposition 1. With the choice above of r and K, and using $\mu_0 < 1/e$, the statement (51) is replaced by the uniform estimate (i) in proposition 1, where the constant T is given in (36), and is already independent of $\mathcal{M}$. Thus, the statement (i) of proposition 1 is proven.

The statement (ii) follows from proposition 3. First, we check the condition (37). To this end, substituting $\alpha_0/(4\sqrt{n}MrK)$ in place of δ we rewrite the condition as

$$\frac{\nu_*\varepsilon r^{2n}}{\delta^2} < 1 \ , \quad \nu_* = \frac{2^6 n^2 B^{2n} K^{2n}}{m} \left(\frac{1+e^{-\sigma/4}}{1-e^{-\sigma/4}}\right)^n \ .$$

This is satisfied in view of the definition (11) of $\overline{\mu}$. Next, we evaluate δ' as $\alpha_0/(4\sqrt{n}MrK)$, which gives the third of (11). Finally, the form of the transformed Hamiltonian is explicitly stated by proposition 3.

The statement (iii) is also consequence of proposition 3. Indeed: the new constants M', m' and Ω' are computed from (40), (41) and (42), substituting again $\alpha_0/(8\sqrt{n}MrK)$ in place of δ, and using $\overline{\mu} \geq \nu_*$ in view of the definition of $\overline{\mu}$; concerning ε', replace in (39) μ with $\mu_0 < 1/e$, and use r as given by (55).

We come finally to statement (iv). Using the estimate (46) of lemma 2 we immediately find (9). The estimate (10) requires a short discussion. We follow the scheme of proof of lemma 2, up to the estimate (47). This estimate must be changed taking into account a further enlargement of the resonant zones $\mathcal{Z}_{\mathcal{M}}$ due to the replacement of the frequency map $\omega(p)$ with $\omega'(p')$. This is estimated by $\Omega' - \Omega$, with Ω' given by (8). On the other hand, replacing Ω with Ω' in (47) we get

$$\mathrm{Vol}(\mathcal{Z}_{\mathcal{M}} \cap \omega'(\mathcal{V})) < 2{\Omega'}^{n-1}(\beta_1 + |\Omega' - \Omega|) \ .$$

We claim $|\Omega' - \Omega| < \beta_1$. For, by (45) we get

$$\beta_1 = \frac{2M\delta}{(B-\sqrt{2})B^{n-1}r^{n-1}K^{n-1}} > \frac{2M}{B^n K^{n-1}} \left(\frac{\overline{\mu}\varepsilon}{\delta^2}\right)^{1/4} \delta \ ,$$

where (55) has been used, and our claim follows by straightforward comparison with (8). Using this, we estimate

$$\begin{aligned} \mathrm{Vol}(\mathcal{G}'_{\delta'}) > \mathrm{Vol}(\omega'(\mathcal{G}')) &> \left[1 - \left(\frac{\Omega'}{\Omega}\right)^{n-1} \frac{2^{n+2}N^{n-1}\beta_1}{V}\right] \mathrm{Vol}(\omega(\mathcal{G}_\delta)) \\ &> \left[1 - \left(\frac{\Omega'}{\Omega}\right)^{n-1} \frac{2^{n+2}N^{n-1}\beta_1}{V}\right] \Omega^{n-1} V \ , \end{aligned}$$

so that using (8) we conclude

$$V' = V \left[1 - \left(\frac{\Omega'}{\Omega}\right)^{n-1} \frac{2^{n+2}N^{n-1}\beta_1}{V}\right] \left(1 + \frac{m}{12\sqrt{n}B^n K^n \Omega} \left(\frac{\overline{\mu}\varepsilon}{\delta^2}\right)^{3/4} \delta\right)^{1-n} \ .$$

Substituting β_1 as given by (45) we get (10). By the definition (12) of B and by the trivial inequality $\Omega' < 2\Omega$ (use $\delta < \Omega$) one immediately checks $V' > 0$. This concludes the proof.

A. Explicit estimates for the main theorem

In view of (16) and of definition (14) of ξ_r we easily get

$$K_0 < K_{r-1} < K_r \le 4K_{r-1} \le 4^r K_0 \ ,$$

On the other hand, we also have the elementary inequality

$$\left(\frac{1+e^{-\xi_r/4}}{1-e^{-\xi_r/4}}\right) < 4\left(\frac{1+e^{-\xi_{r-1}/4}}{1-e^{-\xi_{r-1}/4}}\right) < 4^r\left(\frac{1+e^{-\xi_0/4}}{1-e^{-\xi_0/4}}\right) . \tag{56}$$

Moreover, $K_r\xi_r \ge K_{r-1}\xi_{r-1}/2$ and $K_r\xi_r \ge K_0\xi_0/2$. All these inequalities are clearly true for all $r \ge 0$. Let us now assume, by induction, that (17) holds for $0 \le s < r$; this is true, of course, for r=1. Putting this hypothesis and the inequalities above in (16), for $0 \le s < r$ we get

$$\begin{aligned} B_s < \lambda B_0 < \lambda B_{s-1} \ , \quad B_s - \sqrt{2} < \lambda(B_0 - \sqrt{2}) < \lambda(B_{s-1} - \sqrt{2}) \ , \\ \overline{\mu}_s < 2^{6n+1}\lambda^{2n+1}\overline{\mu}_{s-1} \ , \quad \overline{\mu}_s < 2^{s(6n+1)}\lambda^{2n+1}\overline{\mu}_0 \end{aligned} \tag{57}$$

with

$$\lambda = \frac{M_\infty^{1+1/n} m_0 V_0^{1/n}}{M_0^{1+1/n} m_\infty V_\infty^{1/n}} \ .$$

We now proceed to estimating the sequence x_r. To this end, recalling (15) and using the estimates (57) we first get, for $0 \le s \le r$,

$$\begin{aligned} x_s &= 2^{6n+7}\lambda^{2n+1} n B_{s-1}^{2n} K_{s-1}^{2n} \left(\frac{1+e^{-\xi_{s-1}/4}}{1-e^{-\xi_{s-1}/4}}\right)^n \sqrt{\frac{4\overline{\mu}_{s-1}\varepsilon_{s-1}}{\varrho_{s-1}^2}} \exp\left(-x_{s-1}^{-1/(4n)}\right) \\ &< a b^s x_{s-1}^{1/2} \exp\left(-x_{s-1}^{-1/(4n)}\right) \ , \end{aligned}$$

where

$$a = n B_0^{2n} K_0^{2n} \lambda^{4n+1} 2^{6n+7} \left(\frac{1+e^{-\xi_0/4}}{1-e^{-\xi_0/4}}\right)^n \ , \quad b = 2^{6n} \ .$$

It is also useful to remark that in view of $n \ge 1$ and $K_0 \ge 1$ we have the numerical estimates

$$B_0 > 26 \ , \quad b \ge 2^6 \ . \tag{58}$$

We prove now that, provided x_0 is small enough (the condition on x_0 will be expressed in (62) below), one has

$$\sum_{s=0}^{r} x_s^{1/2} < \frac{\pi^2}{96} \ . \tag{59}$$

Indeed, consider the infinite sequence $\{y_r\}_{r\geq 0}$ recursively defined as

$$y_0 = x_0 \ , \quad y_r = ab^r \exp(-y_{r-1}^{-1/(4n)}) \ . \tag{60}$$

It is evident that if $x_0 < 1$ and if this sequence is decreasing then one has $x_s \leq y_s$ for $0 \leq s \leq r$. We first prove by induction that for every r one has

$$y_r < \frac{\eta}{[(r+1)\ln b + \ln a]^{4n}} \tag{61}$$

provided η is sufficiently small, in particular $\eta < 1$, and provided the inequality is satisfied for $r = 0$. To this end, using (60) and (61), by the induction hypothesis we get

$$\ln y_r = (r \ln b + \ln a) - y_{r-1}^{-1/(4n)} < (r \ln b + \ln a)(1 - \eta^{-1/(4n)}) \ ,$$

so that it is enough to demand that the r.h.s. is smaller than the r.h.s. of (61), i.e. that

$$(\eta^{-1/(4n)} - 1)(r \ln b + \ln a) > 4n \ln[(r+1)\ln b + \ln a] - \ln \eta \ .$$

This is true provided

$$\eta^{-1/(4n)} - 1 + \frac{\ln \eta}{r \ln b + \ln a} > \frac{4n \ln[(r+1)\ln b + \ln a]}{r \ln b + \ln a} \ .$$

Remarking that the l.h.s. is clearly increasing with r (recall $\eta < 1$), while the r.h.s. is clearly decreasing, it is enough to satisfy the inequality for $r = 0$. This gives the condition

$$\eta^{-1/(4n)} - 1 + \frac{\ln \eta}{\ln a} > \frac{4n \ln(\ln b + \ln a)}{\ln a} \ ,$$

which is clearly satisfied for η small enough. For instance, an explicit estimate is easily found by using

$$\frac{1}{2}\eta^{-1/(4n)} + \frac{\ln \eta}{\ln a} > 0 \quad \text{for} \quad \eta < \left(\frac{\ln a}{8n}\right)^{4n} \ ,$$

which gives the condition

$$\eta < \left[\frac{8n \ln(\ln b + \ln a)}{\ln a} + 1\right]^{-4n} \ .$$

For $r = 0$, recalling also that $y_0 = x_0$, (61) gives the condition on x_0

$$x_0 < \left[(\ln b + \ln a)\left(\frac{8n \ln(\ln b + \ln a)}{\ln a} + 1\right)\right]^{-4n} \ , \tag{62}$$

which is clearly satisfied provided ε is small enough. This ensures that (61) is satisfied for every r.

Concerning (59), just remark that in view of $n > 1$ one has

$$x_r < y_r < \frac{\eta}{[(r+1)\ln b]^4} \ ;$$

this readily gives

$$\sum_{s\geq 0} x_s^{1/2} < \frac{\eta^{1/2}}{\ln^2 b}\sum_{s\geq 0}\frac{1}{s^2} = \frac{\eta^{1/2}\pi^2}{6\ln^2 b} \ ,$$

so that (59) follows in view of (58), which gives $\ln b > 4$, and in view of $\eta < 1$. Furthermore the claim that the sequences ε_r and ϱ_r decrease to zero, is proven in view of the convergence of x_r to 0 and of the definition (14).

We come now to the proof of (17). For m_r and M_r we use the elementary estimates

$$\left(1-\frac{y}{2}\right)^{-1} \leq 1+y \quad \text{for} \quad 0\leq y\leq 1 \ . \tag{63}$$

and

$$\ln\prod_{s=0}^{r}(1+x_s^{1/2}) < \sum_{s=0}^{r} x_s^{1/2} < \frac{\pi^2}{96}$$

where (59) has been used. This immediately gives

$$\prod_{s=0}^{r}\left(1+\tfrac{x_s^{1/2}}{2}\right) < \exp\left(\tfrac{\pi^2}{96}\right) = 1.108\ldots \ , \quad \prod_{s=0}^{r}\left(1-\tfrac{x_s^{1/2}}{2}\right) > \exp\left(-\tfrac{\pi^2}{96}\right) = 0.902\ldots \ ,$$

so that, in view of (14) the first two inequalities (17) are satisfied with m_∞ and M_∞ given, e.g., by (18). Concerning the estimate for Ω_r, with a trivial use of the convexity condition we have $\Omega_{r-1} > m_{r-1}\varrho_{r-1}$. Using the expression of Ω_r in (14) we immediately get

$$\Omega_r < \Omega_{r-1}\left(1+\frac{x_{r-1}^{3/4}}{24\sqrt{n}B_{r-1}^n K_{r-1}^n}\right) \ .$$

Proceeding by induction, this easily gives

$$\Omega_0 < \ldots < \Omega_{r-1} < \Omega_r < \Omega_0\prod_{s=0}^{r-1}\left(1+\frac{x_s^{3/4}}{24\sqrt{n}B_0^n K_0^n}\right) < \Omega_0\exp\left(\frac{\pi^2}{59904}\right) \ , \tag{64}$$

so that (17) is satisfied with Ω_∞ given, e.g., by (18). The estimate for V_r requires a short discussion. Firstly, using $m_s\varrho_s < \Omega_s$, $B_s > B_0 > 26$, $K_0 > 1$, $n \geq 1$ and (59) we get

$$\prod_{s=0}^{r}\left(1+\frac{m_s\varrho_s x_s^{3/4}}{24\sqrt{n}B_s^n K_s^n\Omega_s}\right) < \prod_{s=0}^{r}\left(1+\frac{x_r^{3/4}}{24\sqrt{n}B_0^n K_0^n}\right) < \exp\left(\frac{\pi^2}{48\cdot 26^n}\right) \ . \tag{65}$$

Secondly, by the definition of ϱ_r in (14) and by (61) we also have (recall that $\ln b > 4$)

$$\varrho_s < \frac{1}{4\sqrt{n}B_0^n K_0^n} \cdot \frac{\varrho_{s-1}}{s\ln b} < \frac{\varrho_0}{(16\sqrt{n}B_0^n K_0^n)^s s!} \ .$$

This gives $\sum_{s=0}^{r} \varrho_s < \frac{10}{9}\varrho_0$, namely (20). Thirdly, using (63) we get

$$\begin{aligned}
&\prod_{s=0}^{r-1}\left[1-\left(\frac{\Omega_{s+1}}{\Omega_s}\right)^{n-1}\frac{2^{n+2}M_s\varrho_s}{(B_s-\sqrt{2})B_s^{n-1}V_s}\right]^{-1}\\
&\qquad<\prod_{s=0}^{r-1}\left[1+\left(\frac{\Omega_\infty}{\Omega_0}\right)^{n-1}\frac{2^{n+3}M_\infty\varrho_s}{(B_0-\sqrt{2})B_0^{n-1}V_\infty}\right]\\
&\qquad<\exp\left[\left(\frac{\Omega_\infty}{\Omega_0}\right)^{n-1}\frac{10\cdot 2^{n+3}M_\infty\varrho_0}{9(B_0-\sqrt{2})B_0^{n-1}V_\infty}\right] \ .
\end{aligned} \tag{66}$$

On the other hand, by (13) we have

$$B_0-\sqrt{2} > \frac{4M_0}{m_0}\left(\frac{8M_0\varrho_0}{V_0}\right)^{1/n} \ ;$$

recalling also (64), and using, e.g., $\Omega_\infty/\Omega_0 < 10/9$ and (18) we bound the expression (66) by

$$\exp\left(\frac{2\cdot 5^{n+1}V_0}{9^{n+1}V_\infty}\right)$$

Replacing the latter expression and (65) in the definition (14) for V_r we get the condition

$$V_r > V_0\exp\left(-\frac{2\cdot 5^{n+1}V_0}{9^{n+1}V_\infty}\right)\exp\left(-\frac{(n-1)\pi^2}{48\cdot 26^n}\right)$$

so that (17) is satisfied for some $V_\infty < V_0$. For instance, we demand the r.h.s. of the latter expression be greater than V_∞, and this is satisfied with V_∞ given, e.g., by (18). This proves that the sequences m_r, M_r, Ω_r and V_r defined by (14) satisfy (18) for every $r \geq 0$.

Finally, we come to (19). All constant are computed from the sequences x_r, ε_r and ϱ_r given by (15) and (14). In particular, C_1 is estimated using (57) and recalling that $\overline{\mu}_0 \equiv \overline{\mu}$, and (56) is used for C_2.

B. Proof of propositions 2 and 3

Starting with the Hamiltonian (25) we look for a canonical transformation generated via the truncated sequence $\chi^{(r)} = \{\chi_1, \ldots, \chi_r\}$ which gives the transformed Hamiltonian the form (33). To this end, we solve with respect to the unknowns $\chi_1, \ldots, \chi_r$ and

$Z_1, \ldots, Z_r$ the equation $T_\chi^{(r)} H^{(r)} = H$. Using the linearity of T_χ and the expansions (25) and (33) we readily get equations

$$Z_s - L_{H_0} \chi_s = \Psi_s \ , \quad 1 \le s \le r \ , \tag{67}$$

where

$$\Psi_1 = H_1 \ , \quad \Psi_s = H_s - \sum_{j=1}^{s-1} \frac{j}{s} (L_{\chi_j} H_{s-j} + E_{s-j} Z_j) \quad \text{for } 1 < s \le r \ . \tag{68}$$

These equations are deduced as follows. Write the equation $T_{\chi^{(r)}} H^{(r)} = H$ as

$$\sum_{s\ge 0} E_s H_0 + \sum_{s\ge 1} \sum_{j=1}^{s} E_{s-j} Z_j + T_{\chi^{(r)}} \mathcal{R}^{(r)} = \sum_{s\ge 0} H_s \ ,$$

where the operators $\{E_s\}_{s\ge 0}$ are recursively defined by (27). This equation is satisfied by solving the system

$$\begin{aligned} & E_s H_0 + \sum_{j=1}^{s} E_{s-j} Z_j = H_s \ , \quad 1 \le s \le r \\ & \sum_{s>r} E_s H_0 + \sum_{s>r} \sum_{j=1}^{r} Z_j + T_{\chi^{(r)}} \mathcal{R}^{(r)} = \sum_{s>r} H_s \ . \end{aligned} \tag{69}$$

The second equation just defines $\mathcal{R}^{(r)}$ but it is actually useless. The first equation for $s = 1$ immediately gives (68), while, for $s > 1$, it can be rewritten as

$$Z_s - L_{H_0} \chi_s = H_s - \sum_{j=1}^{s-1} \frac{j}{s} L_{\chi_j} E_{s-j} H_0 - \sum_{j=1}^{s-1} E_{s-j} Z_j =: \Psi_s \ .$$

It remains to prove that the right hand side coincides with Ψ_s as given by (68). To this end, replace here $E_{s-j} H_0$ as given by the first of (69) and split $E_{s-j} Z_j$, getting

$$\begin{aligned} \Psi_s = H_s - \sum_{j=1}^{s-1} \frac{j}{s} L_{\chi_j} H_{s-j} - \sum_{j=1}^{s-1} \frac{j}{s} E_{s-j} Z_j \\ + \sum_{j=1}^{s-1} \frac{j}{s} L_{\chi_j} \sum_{m=1}^{s-j} E_{s-m-j} Z_m - \sum_{j=1}^{s-1} \frac{s-j}{s} E_{s-j} Z_j \ . \end{aligned}$$

We show now that the last two terms cancel each other. Indeed, using the recursive definition of E_{s-j}, and exchanging the sums, one gets

$$\begin{aligned}-\sum_{j=1}^{s-1}\frac{s-j}{s}E_{s-j}Z_j &= -\sum_{j=1}^{s-1}\frac{s-j}{s}\sum_{m=1}^{s-j}\frac{m}{s-j}L_{\chi_m}E_{s-j-m}Z_j\\ &= -\sum_{m=1}^{s-1}\frac{m}{s}L_{\chi_m}\sum_{j=1}^{s-1}E_{s-j-m}Z_j\ .\end{aligned}$$

This cancels out the next to last term above, as claimed.

The solution of equation (67) is a well known matter. Denoting by $c_k(p)$ the known coefficients of the Fourier expansion of Ψ_s, we put $Z_s = \sum_{k\in\mathcal{M}} c_k(p)\exp(ik\cdot q)$ (the normal form part of Ψ_s), and determine χ_s by solving in the usual way the equation $-L_{H_0}\chi_s = \Psi_s - Z_s$; precisely, the coefficients of χ_s are determined as $ic_k(p)/(k\cdot\omega(p))$. By the way, this also shows that $\chi_s \in \mathcal{F}_s$ and $Z_s \in \mathcal{F}_s$, provided $\Psi_s \in \mathcal{F}_s$. The easy proof that $\Psi_s \in \mathcal{F}_s$ is made by induction.

Let now $\mathcal{W}$ be a non–resonance domain of type $(\mathcal{M},\alpha,2\delta,N)$ with $N = rK$. It is an easy matter to see that the Hamiltonian (25) can be given the normal form (33) with respect to $\mathcal{M}$ up to order r. For, no denominator $k\cdot\omega(p)$ can vanish in the domain $\mathcal{W}_{2\delta}$. We now produce explicit quantitative estimates for the transformation.

We start by estimating the generating function. In order to avoid unnecessary complication in the notations, we treat here the complex extensions of the domains as free parameters, that we denote by ϱ,ξ. The relation with the analogous quantities δ,σ in propositions 2 and 3 is $\varrho = 2\delta$, $\xi = \sigma/2$. We prove the following

Lemma 3: *Let $\|H_s\|_{(\varrho,\xi)} \le h^{s-1}\mathcal{E}$ for $s \ge 1$ and some positive h and $\mathcal{E}$. Then, for every positive $d < 1$ one has*

$$\begin{aligned}\|\chi_s\|_{(1-d)(\varrho,\xi)} &\le \frac{\beta^{s-1}\mathcal{E}}{s\alpha}\\ \|Z_s\|_{(1-d)(\varrho,\xi)} \le \|\Psi_s\|_{(1-d)(\varrho,\xi)} &\le \frac{\beta^{s-1}\mathcal{E}}{s}\end{aligned}$$

where

$$\beta = \frac{2^4(r-1)\mathcal{E}}{e\alpha d^2\varrho\xi} + 4h\ . \tag{70}$$

Proof. We first remark that the solution of (67) given above satisfies the estimates

$$\|Z_s\|_{\varrho,\xi} \le \|\Psi_s\|_{\varrho,\xi}\ ,\quad \|\chi_s\|_{\varrho,\xi} \le \frac{1}{\alpha}\|\Psi_s\|_{\varrho,\xi}\ , \tag{71}$$

for whatsoever ϱ and ξ. This is an easy consequence of the definition of non–resonance domain. For $1 \le s \le r$, we define $d_s = d\sqrt{(s-1)/(r-1)}$, and look for sequences

$\{\eta_s\}_{1\le s\le r}$ and $\{\tilde{\vartheta}_{s,m}\}_{0\le s\le r-1,\,1\le m\le r-s}$ such that

$$\|\Psi_s\|_{(1-d_s)(\varrho,\xi)} \le \eta_s\mathcal{E}\ , \quad \|E_s Z_m\|_{(1-d_{s+m})(\varrho,\xi)} \le \tilde{\vartheta}_{s,m}\mathcal{E}\ . \tag{72}$$

In view of (68) we immediately get $\eta_1 = 1$; moreover by (71) and by $E_0 Z_m = Z_m$ we also have $\tilde{\vartheta}_{0,m} = \eta_m$. Now we look for a recursive definition of the sequences η and $\tilde{\vartheta}$. To this end, using (68) and the generalized Cauchy estimates (23) we immediately get

$$\|\Psi_s\|_{(1-d_s)(\varrho,\xi)} \le h^{s-1}\mathcal{E} + \sum_{j=1}^{s-1}\frac{j}{s}\left(\frac{2}{ed_s(d_s-d_j)\varrho\xi\alpha}\eta_j h^{s-1-j}\mathcal{E} + \tilde{\vartheta}_{s-j,j}\right)\mathcal{E}\ .$$

Furthermore, by

$$E_s Z_m = \sum_{j=1}^{s}\frac{j}{s}L_{\chi_j}E_{s-j}Z_m$$

we also get

$$\|E_s Z_m\|_{(1-d_{s+m})(\varrho,\xi)} \le \sum_{j=1}^{s}\frac{j}{s}\cdot\frac{2\mathcal{E}^2}{e(d_{s+m}-d_j)(d_{s+m}-d_{s+m-j})\varrho\xi\alpha}\eta_j\tilde{\vartheta}_{s-j,m}$$

In view of the trivial inequality

$$(\sqrt{s-1}-\sqrt{j-1})(\sqrt{s-1}-\sqrt{s-j-1}) \ge \frac{1}{2} \quad \text{for } 1\le j\le s-1$$

we get

$$\frac{1}{(d_{s+j}-d_m)(d_{s+j}-d_{s+j-m})} = \frac{r-1}{(\sqrt{s+j-1}-\sqrt{m-1})(\sqrt{s+j-1}-\sqrt{s+j-m-1})d^2} \le \frac{2(r-1)}{d^2}$$
$$\frac{1}{d_s(d_s-d_j)} = \frac{r-1}{\sqrt{s-1}(\sqrt{s-1}-\sqrt{j-1})d^2} \le \frac{2(r-1)}{d^2}\ .$$

Thus the sequences can be defined as

$$\eta_s = h^{s-1} + \frac{C_r}{s}\sum_{j=1}^{s-1} j\eta_j h^{s-j-1} + \frac{1}{s}\sum_{j=1}^{s-1} j\tilde{\vartheta}_{s-j,j}$$
$$\tilde{\vartheta}_{s,m} = \frac{C_r}{s}\sum_{j=1}^{s} j\eta_j\tilde{\vartheta}_{s-j,m}\ ,$$

with

$$C_r = \frac{4(r-1)\mathcal{E}}{e\alpha d^2\varrho\xi}\ . \tag{73}$$

Our task now is to estimate the latter sequences, remarking that, in view of (72) and (71) it is actually enough to estimate the sequence η_s. Recalling that $\eta_1 = 1$ and $\tilde{\vartheta}_{0,m} = \eta_m$, one readily gets (by induction) $\tilde{\vartheta}_{s,m} = \eta_m \tilde{\vartheta}_{s,1}$. Thus, it is convenient to introduce $\vartheta_s = \tilde{\vartheta}_{s,1}$, and study the sequences

$$
\begin{aligned}
\eta_s &= h^{s-1} + \frac{C_r}{s} \sum_{j=1}^{s-1} j\eta_j h^{s-j-1} + \frac{1}{s} \sum_{j=1}^{s-1} j\eta_j \vartheta_{s-j} \\
\vartheta_s &= \frac{C_r}{s} \sum_{j=1}^{s} j\eta_j \vartheta_{s-j} \ ,
\end{aligned}
\tag{74}
$$

starting with $\eta_1 = \vartheta_0 = 1$. Multiplying the first one by C_r and subtracting it from the second we get

$$\vartheta_s = 2C_r\eta_s - C_r h^{s-1} - \frac{C_r^2}{s} \sum_{j=1}^{s-1} j\eta_j h^{s-j-1} \ .$$

Furthermore, in view of the fact that the last sum is clearly positive, we also get

$$\vartheta_s < 2C_r\eta_s - C_r h^{s-1} \ .$$

Replacing this inequality in the first of (74) we finally get

$$\eta_s < h^{s-1} + \frac{C_r^2}{s} \sum_{j=1}^{s-1} j\eta_j \eta_{s-j} = h^{s-1} + C_r \sum_{j=1}^{s-1} \eta_j \eta_{s-j} \ .$$

We conclude

$$\eta_s \le (C_r + h)^{s-1} \mu_s \ , \tag{75}$$

where

$$\mu_1 = 1 \ , \quad \mu_s = \sum_{j=1}^{s-1} \mu_j \mu_{s-j} \ ,$$

as it is easily proven by induction. We claim

$$\mu_s = \frac{2^{s-1}(2s-3)!!}{s!} \le 4^{s-1} \ , \tag{76}$$

where the standard notation $(2n+1)!! = 1 \cdot 3 \cdot \ldots \cdot (2n+1)$ has been used. In order to prove this claim, let the function $g(z)$ be defined as $g(z) = \sum_{s \ge 1} \mu_s z^s$, so that $\mu_s = g^{(s)}(0)/s!$. Then it is immediate to check that the recursive definition above is equivalent to the equation $g = z + g^2$. By repeated differentiation one readily finds

$$g' = \frac{1}{1-2g} \ , \ \ldots \ , \ g^{(s)} = \frac{2^{s-1}(2s-3)!!}{(1-2g)^{2s-1}}$$

(check by induction), and this proves (76).

Finally, collecting (76), (75), (73), (72) and (71) the statement immediately follows. *Q.E.D.*

We now come back to the proof of propositions 1 and 2. First, we produce estimates for the coordinate transformation $p = T_\chi^{-1} p'$, $q = T_\chi^{-1} q'$ and for the remainder $\mathcal{R}^{(r)}$. By the hypothesis (a) and by the expansion (25) with the estimate (26) we evaluate the constants h and $\mathcal{E}$ in lemma 3 as

$$h = e^{-K\sigma/4} \ , \quad \mathcal{E} = \varepsilon \left(\frac{1 + e^{-\sigma/4}}{1 - e^{-\sigma/4}} \right)^n \ . \tag{77}$$

However, we remark that in view of the hypotheses of propositions 2 and 3 the estimate (26) applies to $\|H_s\|_{(2\delta,\sigma/2)}$, so that in lemma 3 we must put $\varrho = 2\delta$ and $\xi = \sigma/2$. We also set $d = 1/8$. By lemma 3 we conclude that the generating sequence $\chi^{(r)}$ satisfies $\|\chi_s\|_{(3\delta/2,\sigma/8)} \le \beta^{s-1} G/s$, namely (29), with

$$\beta = \frac{2^{10}(r-1)\mathcal{E}}{e\alpha\delta\sigma} + 4e^{-K\sigma/4} \ , \quad G = \frac{\mathcal{E}}{\alpha} \ , \tag{78}$$

$\mathcal{E}$ being given by (77). The condition (30) for convergence of the canonical transformation generated by T_χ (or by T_χ^{-1}) becomes

$$\left(2^7 e + \frac{2^{10}(r-1)}{e} \right) \frac{\mathcal{E}}{\alpha\delta\sigma} + 4e^{-K\sigma/4} \le \frac{1}{2} \ , \tag{79}$$

which is satisfied in view of (35). Thus, by the general estimates (31), and recalling that δ and σ must be replaced by 2δ and $\sigma/2$, respectively, we get the estimate

$$|p - p'| < \frac{\delta}{4} \ , \quad |q - q'| < \frac{\sigma}{16} \ . \tag{80}$$

Concerning the remainder $\mathcal{R}^{(r)}$, in view of $H^{(r)} = T_\chi^{-1}(H_0 + H_1 + \ldots)$, with easy calculations we get

$$\mathcal{R}^{(r)} = \sum_{s>r} \left(\sum_{j=0}^{s-1} D_j H_{s-j} + D_s H_0 \right) \ .$$

On the other hand, in view of the definition (28) of D_s and of (67) we also have

$$D_s H_0 = -\sum_{j=1}^{s} \frac{j}{s} D_{s-j} L_{\chi_j} H_0 = -\sum_{j=1}^{s} \frac{j}{s} D_{s-j} (\Psi_j - Z_j) \ .$$

We conclude

$$\mathcal{R}^{(r)} = \sum_{s>r} \left(\sum_{j=0}^{s-1} D_j H_{s-j} - \sum_{j=1}^{s} \frac{j}{s} D_{s-j}(\Psi_j - Z_j) \right) .$$

Using (32), recalling that $\|H_s\|_{(2\delta,\sigma/2)} \leq h^{s-1}\mathcal{E}$ by hypothesis in lemma 3, and also that in view of the form of the solution of (67) and of lemma 3 one has $\|\Psi_s - Z_s\|_{(1-2d)(2\delta,\sigma/2)} \leq \|\Psi_s\|_{(1-2d)(2\delta,\sigma/2)} \leq \beta^{s-1}\mathcal{E}/s$, we estimate the remainder as

$$\begin{aligned} \left\|\mathcal{R}^{(r)}\right\|_{(1-2d)(2\delta,\sigma/2)} &\leq \sum_{s>r} \left[\sum_{j=0}^{s-1} \left(\frac{2eG}{d^2\delta\sigma} + \beta \right)^j h^{s-j-1} + \sum_{j=1}^{s} \frac{j}{s^2} \left(\frac{2eG}{d^2\delta\sigma} + \beta \right)^{s-j} \frac{\beta^{j-1}}{j} \mathcal{E} \right] \\ &\leq 2\mathcal{E} \sum_{s>r} \left(\frac{2eG}{d^2\delta\sigma} + \beta \right)^{s-1} \leq 4\mathcal{E} \left(\frac{2eG}{d^2\delta\sigma} + \beta \right)^r . \end{aligned}$$

Proceeding as for the estimate (79) we finally get

$$\|\mathcal{R}^{(r)}\|_{3(2\delta,\sigma/2)/4} \leq 4\mathcal{E}\mu^r \tag{81}$$

We can now complete the proof of proposition 2. Using the elementary inequality

$$|p(t) - p(0)| \leq |p(t) - p'(t)| + |p'(t) - p'(0)| + |p'(0) - p(0)|$$

and recalling (80) we get

$$|p(t) - p'(t)| < \frac{\delta}{4} , \quad |p(0) - p'(0)| < \frac{\delta}{4} \tag{82}$$

Thus, recalling that $|v| \leq \|v\|$, it is enough to prove that

$$\|\Delta p'\| := \|p'(t) - p'(0)\| < \frac{\delta}{2} .$$

Since the proof is rather cumbersome, let us explain in a few words the idea. In the new action variables p' consider the so–called "plane of fast drift" $\Pi_{\mathcal{M}}(p'(0))$ passing through $p'(0)$ and parallel to the resonant module $\mathcal{M}$. The dynamics in the neighbourhood of a resonance can be described as a composition of a fast drift along the plane $\Pi_{\mathcal{M}}(p'(0))$ and a very slow motion transversal to that plane. On the other hand, the conservation of energy together with the convexity of the unperturbed Hamiltonian $h(p)$ force the orbit to move along a curved surface which compels the point to go away from the plane of fast drift. Thus, if we ask $\|p'(t) - p'(0)\| > \delta/2$ with δ large enough, then $p'(t)$ must be far from the plane of fast drift, which is possible only after a very long time.

We give now the formal proof. Considering the subspace $\mathrm{span}(\mathcal{M})$ and $(\mathrm{span}(\mathcal{M}))^{\perp}$ in $\mathbf{R}^n$ we define the corresponding projection operators $\Pi_{\mathcal{M}}$ and $\Pi_{\mathcal{M}}^{\perp}$. Let $\lambda \perp \mathcal{M}$, with $|\lambda| = 1$, and let $I_\lambda = \lambda \cdot p'$. Since $H^{(r)}(p', q') - \mathcal{R}^{(r)}(p', q') = H_0(p') + Z_1(p', q') + \ldots +$

$Z_r(p',q')$ is in normal form with respect to $\mathcal{M}$, we readily get that the time derivative of I_λ is

$$\dot I_\lambda = -\lambda\cdot\frac{\partial\mathcal{R}}{\partial q}\ .$$

By Cauchy's inequality (22) and by (81) we get the estimate

$$|\dot I_\lambda| < \frac{8|\lambda|}{3e\sigma}\|\mathcal{R}^{(r)}\|_{(3\delta/2,3\sigma/8)/4} \le \frac{4\mathcal{E}}{\sigma}\mu^r$$

for all $(p',q')\in\mathcal{G}_{3\delta/2}\times\mathbf{T}^n$. Since $\lambda\perp\mathcal{M}$ is arbitrary, we conclude

$$\|\Pi^\perp_{\mathcal{M}}\Delta p'\| < |t|\frac{4\mathcal{E}}{\sigma}\mu^r \tag{83}$$

which is valid until the orbit escapes the domain $\mathcal{G}_{3\delta/2}$, if ever.

We now use conservation of energy and convexity. By the mean value theorem, if $p'(t)\in B_{\delta/2}(p'(0))$ we have

$$\Delta h := h(p'(t)) - h(p'(0)) = \omega(p'(0))\cdot\Delta p' + \frac{1}{2}A(\overline{p})\Delta p'\cdot\Delta p'\ ,$$

where $\overline{p}$ belongs to the segment $(p'(t)-p'(0))$, which is contained in $B_{\delta/2}(p'(0))$. By the convexity condition the second term is estimated by

$$\left|\frac{1}{2}A(\overline{p})\Delta p'\cdot\Delta p'\right| \ge \frac{1}{2}m\|\Delta p'\|^2$$

so that we get

$$\frac{1}{2}m\|\Delta p'\|^2 \le |\Delta h| + |\omega(p'(0))\cdot\Delta p'|\ . \tag{84}$$

By conservation of energy and using hypothesis (a), we get

$$|\Delta h| \le 2\varepsilon\ .$$

In order to estimate the second term on the r.h.s., we recall that by assumption there is $\omega^*\perp\mathcal{M}$ with $\|\omega(p'(0))-\omega^*\|\le\delta_{\mathcal{M}}$. Thus, we compute

$$\begin{aligned}|\omega(p'(0))\cdot\Delta p'| &\le |\omega(p'(0))\cdot\Pi_{\mathcal{M}}\Delta p'| + |\omega(p'(0))\cdot\Pi^\perp_{\mathcal{M}}\Delta p'|\\ &\le |\omega(p'(0)-\omega^*)\cdot\Pi_{\mathcal{M}}\Delta p'| + \Omega\|\Pi^\perp_{\mathcal{M}}\Delta p'\|\\ &\le \delta_{\mathcal{M}}\|\Delta p'\| + \Omega\|\Pi^\perp_{\mathcal{M}}\Delta p'\|\ .\end{aligned}$$

By substitution in (84) we get

$$\frac{1}{2}m\|\Delta p'\|^2 \le 2\varepsilon + \delta_{\mathcal{M}}\|\Delta p'\| + \Omega\|\Pi^\perp_{\mathcal{M}}\Delta p'\|\ .$$

Let us now denote by T the escape time of the orbit from $B_{\delta/2}(p'(0))$, i.e.

$$T = \sup\{t > 0 \,:\, \|p'(s) - p'(0)\| < \delta/2 \text{ for } 0 \le s \le t\}$$

Thus, for $t = T$ we have $|\Delta p'| = \delta/2$, and so also

$$\frac{m}{8}\delta^2 - 2\varepsilon - \frac{1}{2}\delta_{\mathcal{M}}\delta \le \Omega\|\Pi^{\perp}_{\mathcal{M}}\Delta p'\| \ .$$

For negative t the same argument applies, of course.

In view of condition (34) the l.h.s. is bigger than ε, and we have

$$\|\Pi^{\perp}_{\mathcal{M}}\Delta p'\| \ge \frac{1}{\Omega}\left(\frac{m}{8}\delta^2 - 2\varepsilon - \frac{1}{2}\delta_{\mathcal{M}}\delta\right) \ge \frac{\varepsilon}{\Omega} \ .$$

By (83), the statement of proposition 2 follows.

We finally come to the proof of proposition 3. The estimate (38) has already been proved, being just (80). The form of the transformed Hamiltonian is readily obtained putting

$$h' = H_0 + Z_1 + \ldots + Z_r \ , \quad f' = \mathcal{R}^{(r)}$$

so that $h' = h'(p')$ because $\mathcal{M} = \{0\}$. The estimate (39) is nothing but (81). In order to check (40) and (41) we first compute

$$A'(p') = \frac{\partial^2 h'}{\partial p' \partial p'} = A(p) + \sum_{s=1}^{r} \frac{\partial^2 Z_s}{\partial p \partial p} \ .$$

On the other hand, by Cauchy's estimates we have

$$\left|\frac{\partial^2 Z_s}{\partial p_j \partial p_k}(p')\right| \le \frac{2}{(3\delta/4)^2}|Z_s|_{7\delta/4} \le \frac{4}{\delta^2}\frac{\beta^{s-1}}{s}\mathcal{E} \text{ for all } p' \in \mathcal{G}_\delta \ ,$$

where lemma 3 has been used. Being $\beta < 1/2$ in view of (78) and (79) one has

$$|A_{j,k} - A'_{j,k}| \le \frac{2^3}{\delta^2}\mathcal{E} \ .$$

From this, it is easy to prove that

$$\|A'v\| \le \|Av\| + \|(A' - A)v\| \le \left(M + \frac{2^3 n\mathcal{E}}{\delta^2}\right)\|v\| \ ,$$

which gives (40).

Similarly, one has

$$|A'v \cdot v| \ge |Av \cdot v| - |(A' - A)v \cdot v| \ge m\|v\|^2 - \frac{2^3 n\mathcal{E}}{\delta^2}\|v\|^2$$

which in view of (37) gives (41). In view of condition (37) we also have $m' > m/2$.

In order to check (42) compute

$$\omega'_j(p') = \frac{\partial h'}{\partial p'_j} = \omega_j(p) + \sum_{s=1}^{r} \frac{\partial Z_s}{\partial p_j} \ .$$

By Cauchy's estimate we have

$$\left| \frac{\partial Z_s}{\partial p}(p') \right| \le \frac{4}{3\delta} |Z_s|_{7\delta/4} \le \frac{4}{3\delta} \cdot \frac{\beta^{s-1}}{s} \mathcal{E}$$

for all $p' \in \mathcal{W}_{\varrho/4}$. Using again $\beta < 1/2$ one has

$$|\omega'_j(p') - \omega_j(p)| \le \frac{8}{3\delta} \mathcal{E} \ ,$$

from which (42) immediately follows.

Acknowledgements. This work has been supported by contract CE n. CHRX-CT93-0330/DC and contract CE n. ERB–CHRX–CT94–0460.
Part of this work has been done when one of the authors (A.G.) was guest of the Forschunginstitut für Mathematik, ETH, Zürich. The hospitality of ETH has been greatly appreciated.

References

[1] A. Morbidelli and A. Giorgilli: *Superexponential stability of KAM tori*, J. Stat. Phys., **78**, 1607–1617.

[2] N. N. Nekhoroshev: *Exponential estimates of the stability time of near–integrable Hamiltonian systems*, Russ. Math. Surveys, **32**, 1 (1977).

[3] N. N. Nekhoroshev: *Exponential estimates of the stability time of near–integrable Hamiltonian systems, 2.* Trudy Sem. Petrovs., **5**, 5 (1979).

[4] A. N. Kolmogorov: *On the preservation of conditionally periodic motions.* Dokl. Akad. Nauk SSSR, **98**, 527 (1954).

[5] J. Moser: *On invariant curves of area–preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen Math. Phys. Kl. **2**, 1 (1962).

[6] V. I. Arnold: *Proof of a theorem of A. N. Kolmogorov on the invariance of quasi–periodic motions under small perturbations of the Hamiltonian.* Russ. Math. Surv., **18**, 9 (1963).

[7] A. Morbidelli and A. Giorgilli: *On a connection between KAM and Nekhoroshev's theorems.*, Phys. D., **86**, 514–516, (1995).

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

[9] P. Lochak: *Canonical perturbation theory via simultaneous approximations*, Usp. Math. Nauk. **47**, 59–140 (1992). English transl in Russ. Math. Surv.

[10] P. Lochak and A. I. Neishtadt: *Stability of nearly integrable convex Hamiltonian systems over exponentially long times*, Proc. 1991 Conf. on Dynamical Systems (Boston, Birkhauser).

[11] P. Lochak: *Hamiltonian perturbation theory: periodic orbits, resonances and intermittency*, Nonlinearity **6**, 885–904 (1993).

[12] J. Pöschel: *Nekhoroshev's estimates for quasi–convex Hamiltonian systems*, Math. Z. **213**, 187 (1993).

[13] A. Delshams and P. Gutierrez: *Effective stability and KAM theory.*, J. Stat. Phys., in press.

[14] A. Giorgilli and E. Zehnder: *Exponential stability for time dependent potentials*, ZAMP, **43**, 827–855, (1992)

[15] A. Giorgilli: *Methods of complex analysis in classical perturbation theory*, in D.Benest, C.Froeschlé (eds.), *Chaos and Diffusion in Hamiltonian systems*, proceedings of the fourth workshop in Astronomy and Astrophysics of Chamonix (France) 1994, editions Frontieres.