A SPECTRAL FORMULATION OF THE NEKHOROSHEV THEOREM AND ITS RELEVANCE FOR NUMERICAL AND EXPERIMENTAL DATA ANALYSIS

Massimiliano Guzzo* and Giancarlo Benettin**

Abstract: In this paper we provide an analytical characterization of the Fourier spectrum of the solutions of quasi–integrable Hamiltonian systems, which completes the Nekhoroshev theorem and looks particularly suitable to describe resonant motions. We also discuss the application of the result to the analysis of numerical and experimental data. The comparison of the rigorous theoretical estimates with numerical results shows a quite good agreement. It turns out that an observation of the spectrum for a relatively short time scale (of order $1/\sqrt{\varepsilon}$, where ε is a natural perturbative parameter) can provide informations on the behavior of the system for the much larger Nekhoroshev times.

1. Introduction

A. Motivation. The Fourier analysis has been used for years as a numerical tool to investigate the motions of quasi–integrable Hamiltonian systems of the form

$$H(I,\varphi) = h(I) + \varepsilon f(I,\varphi) \ , \qquad (I,\varphi) \in B \times \mathbf{T}^n \ , \tag{1.1}$$

where $B \subset \mathbf{R}^n$ is open and $\mathbf{T}^n$ is the n–dimensional torus; h and f are analytic functions and ε is a small parameter. In the last few years relevant results have been obtained, in this perspective, by the use of the so-called "frequency analysis algorithm" [11],[12],[13]. Supported by the theoretical background provided by KAM theorem [10],[16],[1],[2], this algorithm checks if a solution $I(t), \varphi(t)$ of (1.1) is a quasi–periodic function of the time t, and calculates its frequencies with great accuracy. The extensive application of the algorithm to many initial data in the phase–space gives a global picture of the dynamics: it makes evident the domains where the motions appear to be quasi–periodic and those domains where instead irregularities in the frequencies, as calculated by the algorithm,

* CNRS, Observatoire de la Côte d'Azur at Nice. Present address: Dipartimento di Matematica Pura e Applicata dell'Università di Padova, via G. Belzoni 7, 35131 Padova, Italy.
E-mail: guzzo@math.unipd.it

** Dipartimento di Matematica Pura e Applicata dell'Università di Padova, Gruppo Nazionale di Fisica Matematica (CNR), and Istituto Nazionale di Fisica della Materia. Via G. Belzoni 7, 35131 Padova, Italy. E-mail: benettin@math.unipd.it

reveal the lack of the quasi–periodic nature of motions. The location of these domains also permits a reasonable identification with the resonances of (1.1). The frequency analysis algorithm is therefore well suited to detect and represent the quasi–periodic solutions of a quasi–integrable system, as well as to locate its main resonances.

The present paper is instead concerned with those solutions which are not quasi–periodic functions of time, typically the resonant motions. For such solutions, KAM theorem doesn't provide any analytic description. A good analytic framework is instead provided by Nekhoroshev theorem ([17],[18]; see also [4],[6],[14],[19]), which applies if ε is small and h is convex (or more generally *steep*), and describes any motion of the system, including the resonant ones, by means of suitable normal forms adapted to the resonance properties of the initial data. However, at variance with KAM theorem, the available formulations of Nekhoroshev theorem do not provide a satisfactory description of motions, including, besides the long time stability of the actions, informations on the spectrum of observables during resonant motions.[1)] The aim of this paper is to fill this gap: more precisely to provide some insight on the resonant motions, for systems obeying the assumptions of Nekhoroshev theorem, by producing a clear description of the spectrum of a generic observable. Such a theoretical characterization, as we shall see, is easy to be checked numerically or experimentally, and the agreement between rigorous and numerical results is good.

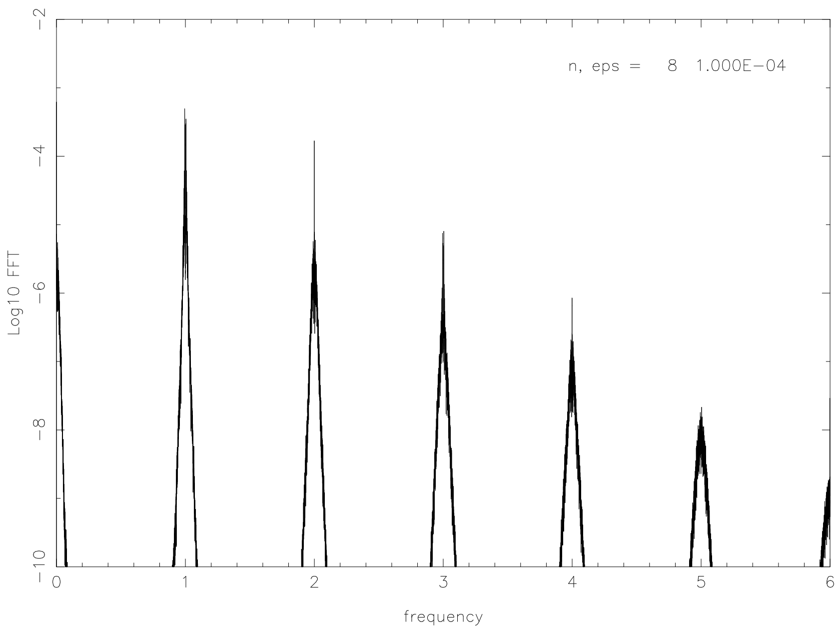

Figure 1. The band structure for the model of eight rotators, for $\varepsilon = 10^{-4}$, in the case of a single frequency (initial data such that $I_1 = 1$, $I_2, \ldots, I_8 = \mathcal{O}(\sqrt{\varepsilon})$). Horizontal axis: the frequency $\alpha/2\pi$. Vertical axis: the modulus squared of the Fourier Transform, in (decimal) logarithmic scale.

[1)] Qualitative descriptions of resonant motions are found in [6], where it is shown that (as far as the remainder can be ignored) some variables in the system are regular, while some other on the same time are chaotic. See also, for a rough first order analysis, [5].

B. A preliminary account of results. Very shortly, the result is as follows: for no matter how small ε, inside resonances (typically, at a distance of order $\sqrt{\varepsilon}$ from the resonant manifolds) the spectrum of a generic observable looks continuous, and (if ε is sufficiently small) it presents narrow peaks around discrete values of the frequency, as shown in figures 1 and 2. The figures refer to a set of $n = 8$ rotators, with Hamiltonian of the form

$$H(I, \varphi) = \frac{1}{2} \sum_{j=1}^{n} I_j^2 + \varepsilon f(\varphi) \ ,$$

f being a trigonometric polynomial (see Section 3 for details). The observable here considered is

$$\mathcal{G}(I, \varphi) = \frac{I_1 + I_3}{1 + (\sin \varphi_1 + \sin \varphi_2)^2} \ , \tag{1.2}$$

and the figures report in semilog scale, versus the frequency α, the square of the Fourier transform (or power spectrum) of

$$G(t) = \Phi(t)\, \mathcal{G}\big(I(t), \varphi(t)\big) \ , \tag{1.3}$$

where $\Phi(t)$ is an analytic "filter function", which is nearly equal to one for a wide time interval, and vanishes at infinity (as is necessary, in order for the Fourier transform to be defined). Our precise choice of the filter is

$$\Phi(t) = \frac{1}{2} \Big[\tanh \Big(\frac{t + 2T}{\sigma} \Big) - \tanh \Big(\frac{t - 2T}{\sigma} \Big) \Big] \ , \tag{1.4}$$

with large T and some $\sigma > 0$ (see Section 3 for details). Figure 1 refers to a maximally resonant motion ($n - 1$ resonances), precisely to $I_1(0) = 1$ and $I_2(0), \ldots, I_8(0)$ of order $\sqrt{\varepsilon}$, $\varepsilon = 10^{-4}$; figure 2 refers instead to an initial datum with two rationally independent frequencies of order one, more precisely to $I_1(0) = 1$, $I_2(0) = \sqrt{2}$, and $I_3(0), \ldots, I_8(0)$ of order $\sqrt{\varepsilon}$, same $\varepsilon = 10^{-4}$. The shape of the peaks suggests an exponential decay of the Fourier transform around lines determined by the initial frequencies, and this is precisely what we are able to prove, including a scaling law for the slopes.

A rough formulation of our theoretical result (for a precise statement see Section 2 below) is as follows: if ε is suitably small, then for any analytic function $\mathcal{G}(I, \varphi)$, and for any chosen time scale T satisfying

$$\sqrt{(\varepsilon_0/\varepsilon)} \le T \le e^{(\varepsilon_0/\varepsilon)^{1/2n}} \ , \tag{1.5}$$

ε_0 being a suitable constant, the function G defined in (1.3) admits a decomposition

$$G(t) = G^0(t) + G^1(t) \ , \tag{1.6}$$

depending on T, such that $G^1(t)$ is negligible for $|t| \le T$, while G^0 is a superposition of modulated plane waves:

$$G^0(t) = \sum_{k \in \mathbf{Z}^n} g_k^0(t)\, e^{ik\cdot\Omega t} \ , \tag{1.7}$$

where $\Omega \in \mathbf{R}^n$ is a suitable vector determined by the initial datum (see below), and g_k^0 has Fourier transform $\tilde{g}_k^0(\alpha)$ satisfying the scaling law

$$\left|\tilde{g}_k^0(\alpha)\right| \le a_0\, T\, e^{-a_1|k|}\, e^{-a_2|\alpha|/\varepsilon^{a_3}} \quad ; \tag{1.8}$$

$a_0, \ldots, a_3$ denote here suitable positive constants, and $|k| = |k_1| + \cdots + |k_n|$. The most important constant is a_3, which gives the scaling law of the slope of the peaks in the figures. We will prove (for convex systems) the uniform bound $a_3 \ge 1/(2n)$, but for a wide class of resonant initial data we get $a_3 = \frac{1}{2}$, and this is indeed the value that we find numerically.

The precise mathematical formulation of these results requires, as usual, many details, and the proof is as annoying as typical proofs in Hamiltonian perturbation theory. It is not difficult, however, to produce a simple heuristic evidence of the scaling law (1.8), with $a_3 = \frac{1}{2}$; see, for this, Section 2C.

C. The location of the peaks in the spectrum. As is clear from (1.7), the quantity $|k \cdot \Omega|$ determines the position of the peaks in the spectrum. The meaning of Ω is easily explained: with reference to the usual Nekhoroshev decomposition of the action space into regions with different resonance properties (see Section 2 for details), if the initial datum $I(0)$ belongs to a suitable neighborhood of a "resonant manifold" $\mathcal{W}_\Lambda$ of multiplicity d, defined as usual by

$$k \cdot \omega(I) = 0 \qquad \text{for any} \quad k \in \Lambda \tag{1.9}$$

where Λ is a d-dimensional sublattice of $\mathbf{Z}^n$ and $\omega = \frac{\partial h}{\partial I}$, then Ω is the projection of the initial frequency $\omega(I(0))$ onto the linear space orthogonal to Λ.

Since Ω is perpendicular to Λ, the terms in the series at the r.h.s. of (1.7) are naturally grouped into classes, such that $k \cdot \Omega$ is constant inside a class: precisely, k and k' are in the same class if $k - k' \in \Lambda$. The classes can be labelled by the vectors of any lattice $\mathcal{L}$ transversal to Λ, $\dim \mathcal{L} = n - d$. Terms in the same class correspond to the same peak in the spectrum, so it is convenient to sum them together into a single coefficient G_k^0, $k \in \mathcal{L}$, thus giving G^0 the more transparent form

$$G^0(t) = \sum_{k \in \mathcal{L}} G_k^0(t)\, e^{ik\cdot\Omega t} \quad ; \tag{1.10}$$

the new coefficients G_k^0 still obey a scaling law like (1.8), namely their Fourier transform $\tilde{G}_k^0(\alpha)$ satisfies

$$\left|\tilde{G}_k^0(\alpha)\right| \le a_0'\, T\, e^{-a_1|k'|}\, e^{-a_2|\alpha|/\varepsilon^{a_3}} \quad ,$$

with suitable $a_0' > 0$, k' being a vector in the class of k with minimal norm. Now the correspondence between peaks in the Fourier spectrum and terms in the sum is one to one (but for the trivial fact that k and the conjugate $-k$ give the same peak). It is also worthwhile to note that there is one term in the series, and thus one peak in the spectrum, for each different small denominator associated to the resonant motion, that is entering the construction of the normal form adapted to the resonant lattice Λ. The series (1.10) can be used to comment figures 1 and 2. Indeed in the case of figure one it is $d = n - 1$, so the lattice Λ is one-dimensional, and the peaks are well separated. More precisely,

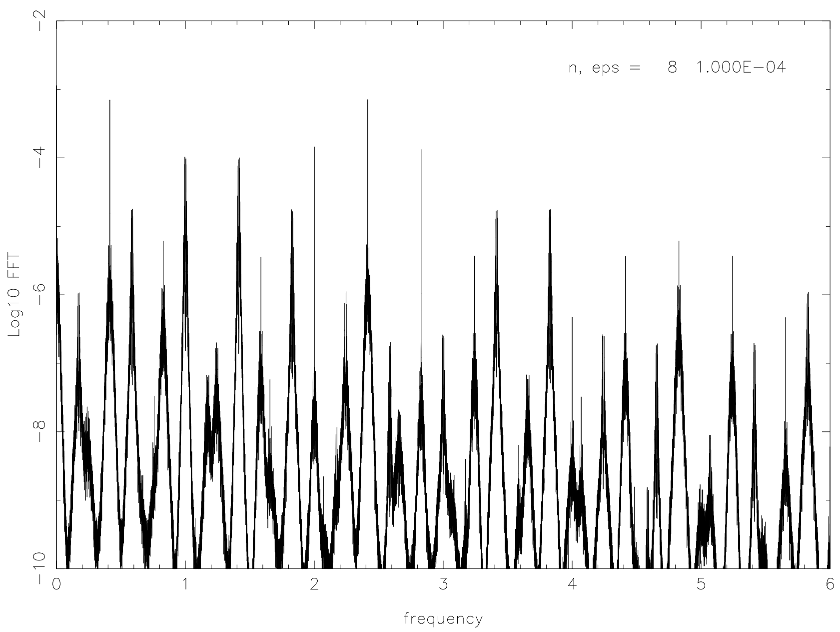

Figure 2. The band structure for the model of eight rotators, for $\varepsilon = 10^{-4}$, in the case of two independent frequencies (initial data such that $I_1 = 1$, $I_2 = \sqrt{2}$, $I_3, \ldots, I_8 = \mathcal{O}(\sqrt{\varepsilon})$.

it is $\Omega = (1, 0, \ldots, 0)$, so the peaks occur at $\alpha/(2\pi) = 1, 2, \ldots$ Instead for figure 2 it is $d = n-2$, so $\mathcal{L}$ is two-dimensional, and the peaks get more abundant (in principle the peaks are dense, but their amplitude decreases so rapidly with $|k|$, that only a finite number of them is visible in the figure).

D. The "band structure" and its concrete use. The novelty of the above decomposition (1.6) consists in distinguishing in the spectrum a quasi–periodic pattern, which is expressed by the plane–waves $e^{ik\cdot\Omega t}$, $k \in \mathcal{L}$, from the chaotic features, which are included in the functions $G^0_k(t)$. Such "amplitude functions" differ from motion to motion, also for nearby initial data, and are certainly complicated and in a sense unpredictable, precisely because of the chaotic character of the resonant motion. Their spectrum however, as far as the system is in the Nekhoroshev regime — that is the assumptions of Nekhoroshev theorem are satisfied, and ε in particular is sufficiently small — must obey the characteristic scaling law (1.8). This "band structure" of the spectrum, including the appropriate scaling law, is easily recognized numerically, see for details Section 3.

The band structure is instead lost if ε is too large, as is clear for example from figure 3, which refers to the same model and the same initial data of figure 2, but to larger $\varepsilon = 10^{-2}$. Quite evidently, the system is no more in the Nekhoroshev regime (in spite of the fact that $\varepsilon = 10^{-2}$ might appear still small), and the actions are in principle allowed to have large variations in short time intervals. Such variations indeed occur, as is clear from figure 4, which shows the behavior of the actions for the same motion as in figure 3, but on a time scale 16 times larger than the one used to produce figure 3. The three curves in the figure represent, top to bottom, $I_2(t)$, $I_1(t)$ and $\frac{1}{6}(I_3(t) + \cdots + I_8(t))$. The

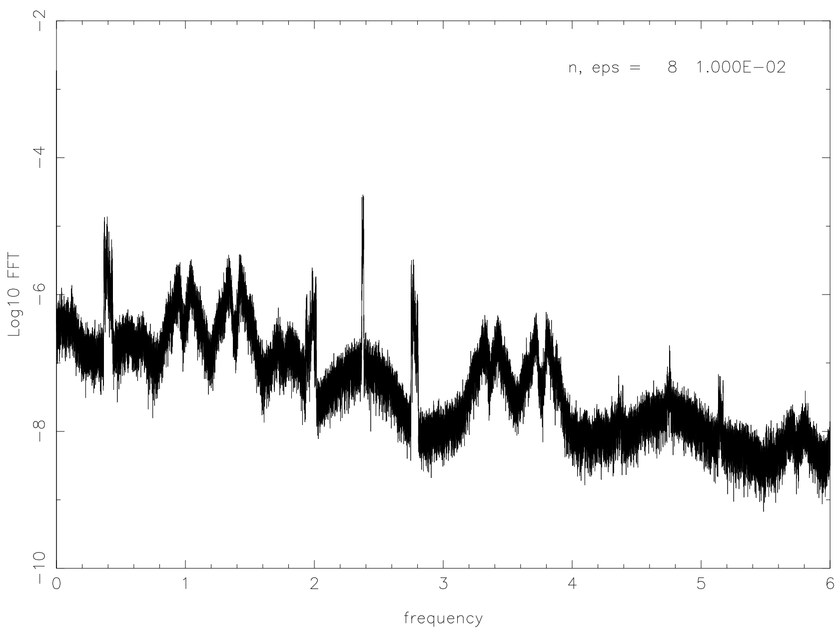

Figure 3. The spectrum for the model of eight rotators, in conditions similar to figure 2, but for larger $\varepsilon = 10^{-2}$.

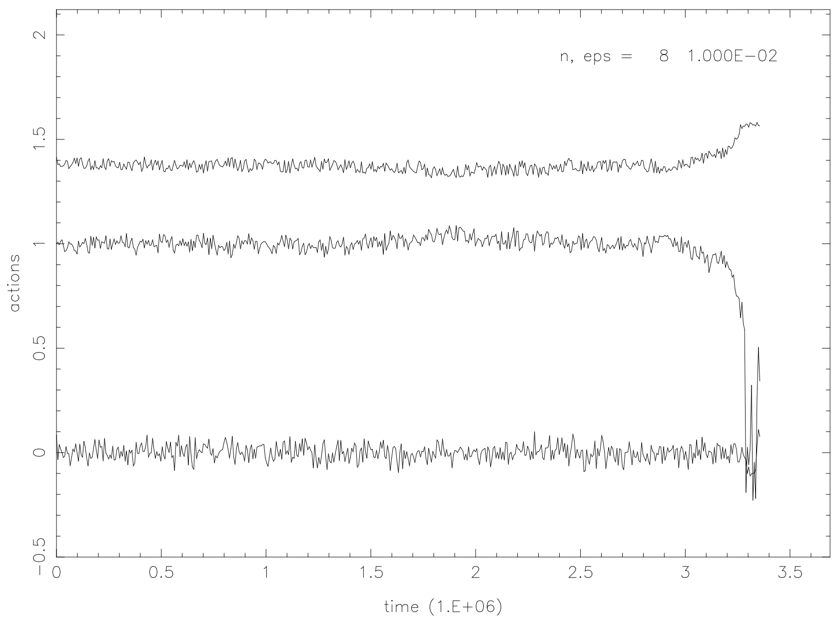

Figure 4. The behavior of the actions, for the same motion as in figure 3. The observation time is 16 times larger.

sudden large variation of the actions after $t = 3 \times 10^6$ is clearly incompatible with the Nekhoroshev regime. But the absence of the Nekhoroshev regime, and thus the possible lack of stability of the actions, was already clear from the lack of the band structure in the spectrum of figure 3, in spite of the fact that the observation time used to produce the spectrum was much shorter (the latter could be made even shorter, see Section 3). On the contrary, observing the band structure for smaller ε, as in figure 1 or 2, indicates that the system is indeed in the Nekhoroshev regime, and thus variations of the actions, if any, can take place only through the much slower mechanism of Arnold diffusion.

Recognizing the presence or absence of the band structure in the spectrum, we stress, requires a relatively short observation time: according to (1.5), any T larger than $\sqrt{\varepsilon_0/\varepsilon}$ is sufficient, and it looks remarkable that a similar observation can produce informations on the motion for the much longer Nekhoroshev time scale. Moreover, although the scaling property of the slope of the peaks with α is an important feature, even the observation of a single motion can provide a relevant insight, and allow a rough distinction between the spectrum of figure 2 (band structure; no sudden evolution of the actions is permitted) and of figure 3 (the structure is lost; a sudden evolution of the actions is no more forbidden). This means that the method in principle is suited to be applied to real experiments, too. Such a strong connection between the theoretical results we are proving, and the practical use one can make of them to understand empirically whether the conclusions of perturbation theory are expected to hold, for a given system and a given value of ε, provides in our opinion the main interest of this work.

The present paper is organized as follows: in Section 2 we give a precise formulation of the main result, as well as a short elementary heuristic proof of it. Section 3 is devoted to the numerical analysis, while Section 4 is devoted to the analytic proof of the statement in Section 2 (some familiarity with the proof of Nekhoroshev theorem is here assumed).

Acknowledgments. We are indebted to A. Morbidelli for useful discussions. This work has been partially supported by the Italian MURST (protocol No. 9802261238_008). M.G. was fully supported by a Marie Curie Training Grant.

2. The spectral formulation of Nekhoroshev theorem

The Nekhoroshev theorem applies to quasi–integrable Hamiltonians of the form (1.1), if the integrable approximation h is quasi-convex (or more generally steep). As is well known, the theorem essentially states that, if ε is sufficiently small, any motion $(I(t), \varphi(t))$ of the system satisfies

$$|I(t) - I(0)| \leq I_0\, \varepsilon^a \qquad \text{for} \quad |t| \leq t_0\, e^{(\varepsilon_0/\varepsilon)^b} \ , \tag{2.1}$$

where $I_0, t_0, \varepsilon_0, a, b$ are suitable positive constants. Possible values of the exponents a, b, see [19], are $a = b = 1/(2n)$. The theorem is proved by constructing in the whole phase–space normal forms adapted to the local resonance properties of the unperturbed Hamiltonian $h(I)$. Besides the stability result (2.1), however, such normal forms can be used to describe in great detail some other relevant features of the motions. Here we will use them to characterize the Fourier spectrum of the solutions.

A. The geometry of resonances. Before formulating our result, we need to give a precise definition of the resonant domains, which is often called the geometry of resonances. We follow here rather closely [19]. We use the notation $\| \, . \, \|$ for the euclidean norm, while for $k \in \mathbf{Z}^n$ we pose $|k| = |k_1| + \cdots + |k_n|$. The geometry of resonances depends on ε and on three parameters ε_0, ω_0 and c_0, and is as follows. Let $\varepsilon \leq \varepsilon_0$, denote

$$N = \left(\frac{\varepsilon_0}{\varepsilon}\right)^{\frac{1}{2n}} , \tag{2.2}$$

and consider all the d-dimensional sublattices Λ of $\mathbf{Z}^n$, $0 \leq d \leq n-1$, called N-lattices, such that: (i) Λ is generated by d vectors $k^{(1)}, \ldots, k^{(d)} \in \mathbf{Z}^n$ with $|k^{(i)}| < N$; (ii) Λ is maximal (it is not properly contained in any sublattice of the same dimension). We denote by $\Lambda = 0$ the zero–dimensional lattice, containing the null vector alone. The cells of any lattice $\Lambda \neq 0$ have a minimal d–dimensional euclidean volume, denoted $|\Lambda|$; conventionally, $|0| = 1$. For each N–lattice Λ we pose

$$\delta_\Lambda = \frac{\omega_0}{(c_0 N)^{n-d} \, |\Lambda|} \ . \tag{2.3}$$

Finally, we define the resonant domain B_Λ by means of the conditions

$$\|\omega(I)\| \geq \omega_0$$

and

$$\|P_\Lambda \omega(I)\| < \delta_\Lambda \ , \qquad \|P_{\Lambda'} \omega(I)\| \geq \delta_{\Lambda'} \quad \text{for any } \Lambda' \text{ of dimension } d+1$$

(only the latter for $d = 0$), where P_Λ denotes the orthogonal projection onto the real space spanned by Λ. The resonant domains cover $B \backslash \{I : \ \|\omega(I)\| < \omega_0\}$.

The parameters ε_0, ω_0 and c_0, and thus the geometry of resonances, depend on the properties of the Hamiltonian at hand, in a way that we do not enter here in detail. We recall however which are the relevant properties characterizing the Hamiltonian, since they are also at the basis of the proposition below (as well as of any perturbative treatment within Nekhoroshev theory).

- The analyticity of H. Given an "extension vector" $\rho = (\rho_I, \rho_\varphi)$ with positive components, we introduce the extended domains

$$\begin{aligned} \Delta_\rho(I) &= \left\{ I' \in \mathbf{C}^n : \|I' - I\| < \rho_I \right\} & \mathcal{B}_\rho &= \bigcup_{I \in B} \Delta_\rho(I) \\ \mathcal{S}_\rho &= \left\{ \varphi \in \mathbf{C}^n : |\mathrm{Im}\, \varphi_j| < \rho_\varphi, \ j = 1, \ldots, n \right\} & \mathcal{D}_\rho &= \mathcal{B}_\rho \times \mathcal{S}_\rho \ , \end{aligned} \tag{2.4}$$

  and require that H is analytic and bounded in $\mathcal{D}_\rho$.
- The convexity of h. Although the natural property would be quasi–convexity, for notational simplicity we restrict here to the convex case, and recall that h is said to be convex in B if there is $m > 0$ such that, for any $I \in B$,

$$\partial^2 h(I) u \cdot u > m \, u \cdot u \qquad \forall \, u \in \mathbf{R}^n \backslash \{0\} \ , \tag{2.5}$$

$\partial^2 h$ denoting the Hessian of h.

- Suitable norms, namely the operator norm of $\partial^2 h$ in $\mathcal{B}_\rho$, that is M such that

$$\|\partial^2 h(I)u\| \le M\,\|u\| \qquad \forall u \in \mathbf{C}^n\ ,\ \forall I \in \mathcal{B}_\rho\ , \tag{2.6}$$

and a convenient norm of f. The sup–norm of f in $\mathcal{D}_\rho$ could work, but good results require a finer choice. We shall be more precise on this point in the course of the proof, where the constants will be computed (the style of the proposition below, for simplicity, is: "there exist constants such that..."). The quantities ρ, m, M and the norm of f determine the constants entering the geometry of resonances, as well as all other constants entering the statement.

B. The statements. We are now able to state our main analytical result.

Proposition 1. *Let H be as in (1.1), and assume:*

i) H is analytic and bounded in the above defined extended set $\mathcal{D}_\rho$, for some ρ;

ii) h is convex in B.

Let $\mathcal{G}(I,\varphi)$ be a real function, extending analytically to $\mathcal{D}_\rho$, and for any real motion $(I(t),\varphi(t))$ let $g(t)=\mathcal{G}(I(t),\varphi(t))$.

There exist constants ε_0, ω_0, c_0 and t_0 such that, with reference to the geometry of resonances induced by ε_0, ω_0 and c_0, and to the free parameter $T\in(T_0,T_1)$,

$$T_0 = t_0\left(\frac{\varepsilon_0}{\varepsilon}\right)^{\frac{1}{2}}\ , \qquad T_1 = t_0\, e^{(\varepsilon_0/\varepsilon)^{\frac{1}{2n}}}\ ,$$

if $\varepsilon<\varepsilon_0$ and the initial datum $I(0)$ belongs to a resonant domain B_Λ, then for $|t|<T_1$ the motion does not escape $\mathcal{D}_\rho$, and for any $|t|<T$, $g(t)$ admits a decomposition

$$g(t)=g^0(t)+g^1(t)\ , \qquad g^0(t)=\sum_{k\in\mathbf{Z}^n} g^0_k(t)e^{ik\cdot\Omega t}\ , \tag{2.7}$$

such that:

a) The functions $g^0_k(t)$ are defined for any $t\in\mathbf{R}$, they extend analytically to the complex strip

$$|\mathrm{Im}\,t| < \beta N^{n-d}\,|\Lambda|\ , \qquad d=\dim\Lambda\ ,$$

with suitable $\beta>0$, and their Fourier transforms $\tilde g^0_k(\alpha)$ do exist.

b) Denoting by $|\mathcal{G}|^\infty_{\mathcal{D}_\rho}$ the sup–norm of $\mathcal{G}$ in $\mathcal{D}_\rho$, the following estimates hold:

$$\sup_{|\mathrm{Re}\,t|\le T\,,\,|\mathrm{Im}\,t|\le\frac{\beta}{2}} |g^0_k(t)| \le a_1 e^{-\frac{1}{48}|k|\rho_\varphi}\,|\mathcal{G}|^\infty_{\mathcal{D}_\rho} \tag{2.8a}$$

$$|\tilde g^0_k(\alpha)| \le a_2\,T\,e^{-\frac{1}{24}|k|\rho_\varphi}\ e^{-\beta|\alpha|N^{n-d}|\Lambda|}\ |\mathcal{G}|^\infty_{\mathcal{D}_\rho}\ . \tag{2.8b}$$

c) The remainder g^1 is analytic for t in the smaller set $|\mathrm{Re}\,t|\le T$, $|\mathrm{Im}\,t|\le\beta/2$, and there satisfies the estimate

$$|g^1(t)| \le \ln(4N^n)\left(a_3\,e^{-\frac{1}{12}N\rho_\varphi} + a_4\,e^{-\frac{2T}{\beta N^{n-d}|\Lambda|}}\right)|\mathcal{G}|^\infty_{\mathcal{D}_\rho}\ . \tag{2.9}$$

In these inequalities, $a_1, \ldots, a_4$ denote suitable positive constants.

Remarks. (i) In the above proposition, the filter (1.4) is not mentioned. From the decomposition (2.7), the decomposition (1.6) of the filtered function $G = \Phi g$, which was considered in the Introduction in connection with the numerical results, trivially follows by posing $G^0 = g^0$; G^1 and g^1 are then negligibly different for $|t| < T$.[2)] (ii) The function $g^0(t)$ is defined for any real t (as is necessary in order for the Fourier trasforms $\tilde{g}^0_k$ to exist), in spite of the fact that g itself is not. This is possible because, as we shall see, g^0 is defined via the truncated normal form.

The above proposition applies to all initial data, in the spirit of the Nekhoroshev theorem. However, the estimate (2.8b), which represents the worst possible scaling law, can be improved for many initial data. This is obvious for initial data on KAM tori, for which g is a quasi–periodic function of time, and also for the other data in the non–resonant domain where, following [8], one can construct normal forms with super–exponentially small remainder. The interesting improvement concerns instead resonant motions, more precisely motions such that the initial action is at a distance of order $\sqrt{\varepsilon}$ from an exact resonance. This is stated in the next proposition (to be thaught of as appended to Proposition 1). The resonant manifold $\mathcal{W}_\Lambda$ here mentioned is defined in the Introduction, see (1.9); $\mathrm{dist}\,(I, \mathcal{W}_\Lambda)$ denotes the inf of $\|I - I'\|$ for $I' \in \mathcal{W}_\Lambda$.

Proposition 2. *There exists $\sigma > 0$ such that, if $I(0) \in B^\Lambda$ and $\mathrm{dist}\,(I(0), \mathcal{W}_\Lambda) < \sigma\sqrt{\varepsilon}$, then (2.8a,b) and (2.9) are replaced by the stronger estimates*

$$\sup_{|\mathrm{Im}\, t| \le b_\Lambda} |g^0_k(t)| \le a_1\, e^{-|k|\frac{1}{48}\rho_\varphi}\, |\mathcal{G}|^\infty_{\mathcal{D}_\rho} \tag{2.10a}$$

$$|\tilde{g}^0_k(\alpha)| \le a_2\, T\, e^{-\frac{1}{4}|k|\rho_\varphi}\, e^{-a_1^\Lambda|\alpha|/\sqrt{\varepsilon}}\, |\mathcal{G}|^\infty_{\mathcal{D}_\rho} \tag{2.10b}$$

and respectively

$$\sup_{|\mathrm{Re}\, t| \le T\,, |\mathrm{Im}\, t| \le b_\Lambda} |g^1(t)| \le \left(a_3 e^{-(a_2^\Lambda/\varepsilon)^{\frac{1}{2(n-d)}}} + a_3^\Lambda e^{-a_4^\Lambda T\sqrt{\varepsilon}} \right) |\mathcal{G}|^\infty_{\mathcal{D}_\rho}\ , \tag{2.11}$$

where $b^\Lambda > \frac{1}{2}\beta$ and $a_1^\Lambda, \ldots, a_4^\Lambda$ are suitable constants.

Propositions 1 and 2 are formulated for perturbations of integrable convex Hamiltonians. However, many physically interesting stability problems, concerning for example the asteroids of the Main Belt (see [9], [15]), the Lagrangian equilibrium points of L4–L5 in the restricted three body problem (see [3]), or the Riemann ellipsoids (see [7]), have been recently shown to be perturbations of *steep* integrable Hamiltonians, rather than convex or quasi–convex ones. Dealing with convexity instead than steepness simplifies many calculations, and allows to use simplified stability arguments like the use of energy conservation. Nevertheless we expect the results to be valid also in the steep case, though the constants and the powers of ε entering the scaling law might worsen.

---

2) The filter, in fact, enters the proof: as we shall see, before defining g^0, the function g is filtered.

C. An heuristic support to Proposition 2. Here we provide an heuristic proof of the scaling law (2.10b) of Proposition 2. We follow the ideas more widely exposed in [6](see also [5]). The argument, very elementary, cannot be turned into a rigorous one, nevertheless it provides some hint to understand the behavior of resonant motions, and is also useful for a better understanding of the numerical work which will be described in the next section. Although the argument is general for convex unperturbed systems, for simplicity we restrict ourselves to the typical case of a system of n weakly coupled rotators,

$$H(I,\varphi) = \frac{1}{2}\sum_{j=1}^{n} I_j^2 + \varepsilon f(I,\varphi) \ . \tag{2.12}$$

We consider the simple case of a d-dimensional resonant lattice of the special form $\Lambda = \{k \in \mathbf{Z}^n : k_1,\ldots,k_m = 0\}$, $m = n-d$, so that the resonant manifold $\mathcal{W}_\Lambda$ has equations $I_{m+1},\ldots I_n = 0$, and $\Omega = (I_1,\ldots,I_m,0,\ldots,0)$. This choice of the resonance is certainly not the most general one, but the argument we shall produce easily adapts to any resonance, through the introduction of the so-called fast and slow variables, see [6].

Nekhoroshev theorem, in its standard version, provides a near to the identity canonical transformation $(I,\varphi) = \mathcal{C}_\Lambda(I',\varphi')$, adapted to the resonance, which conjugates the above Hamiltonian H to H' of the form

$$H'(I',\varphi') = \frac{1}{2}\sum_{j=1}^{n} {I'_j}^2 + \varepsilon f_\Lambda(I',\varphi'_{m+1},\ldots,\varphi'_n) + e^{-(\varepsilon_0/\varepsilon)^a}\,\mathcal{R}_\Lambda(I',\varphi')$$

$$f_\Lambda(I',\varphi'_{m+1},\ldots,\varphi'_n) = \overline{f}_\Lambda(I',\varphi'_{m+1},\ldots,\varphi'_n) + \mathcal{O}(\varepsilon) \ ,$$

where $\overline{f}_\Lambda$ denotes the average of f over the fast angles $\varphi'_1,\ldots,\varphi'_m$, and $\mathcal{R}_\Lambda$ is bounded. As far as the exponentially small remainder can be neglected — this is the part of the argument which by no means can be made rigorous — the behavior of $(I'_{m+1},\ldots,I'_n,\varphi'_{m+1},\ldots,\varphi'_n)$ is governed by the d-degrees of freedom Hamiltonian

$$\mathcal{H}(I'_{m+1},\ldots,I'_n,\varphi'_{m+1},\ldots,\varphi'_n) = \frac{1}{2}\sum_{j=m+1}^{n} {I'_j}^2 + \varepsilon f_\Lambda(I',\varphi'_{m+1},\ldots,\varphi'_n) \ ,$$

in which $I'_1,\ldots,I'_m$, inside f_Λ, play the role of parameters.

The motions considered in Proposition 2 are those with $I_j \simeq I'_j = \mathcal{O}(\sqrt{\varepsilon})$, $m < j \le n$. The trivial rescaling

$$I'_j = \sqrt{J}_j \ , \qquad t = s/\sqrt{\varepsilon} \ , \tag{2.13}$$

reports the actions to order one, and conjugates $\mathcal{H}$ to $\mathcal{K} = \varepsilon^{-1}\mathcal{H}$ of the form

$$\mathcal{K} = \mathcal{K}_0 + \mathcal{O}(\sqrt{\varepsilon}) \ , \qquad \mathcal{K}_0 = \tfrac{1}{2}\sum_{j=m+1}^{n} J_j^2 + g(\varphi'_{m+1},\ldots,\varphi'_n) \ , \tag{2.14}$$

where $g(\varphi'_{m+1},\ldots,\varphi'_n) = \overline{f}(I'_1,\ldots,I'_m,0,\ldots,0,\varphi'_{m+1},\ldots,\varphi'_n)$. Such an Hamiltonian is strongly coupled, and (if $d \ge 2$) typically exhibits chaotic motions, too.[3)] Such motions

[3)] As commented in [6], it is easy to find systems of this form such that, for example, the truncated system $\frac{1}{2}\sum_j J_j^2 + g$ has an homoclinic intersection; such an intersection then persists in $\mathcal{K}$ for small ε.

in turn appear as chaotic resonant motions of the original Hamiltonian (2.12), which, according to (2.13), take place in a neighborhood of order $\sqrt{\varepsilon}$ of the resonant manifold, and on a time scale of order $1/\sqrt{\varepsilon}$. For a generic observable $\mathcal{G}$, depending both on the regular and on the chaotic variables, a scaling law of the form (2.10b) easily follows.

3. Numerical results.

A sketch of the numerical results has been already presented in the Introduction. Here we propose a more detailed and systematic exposition of our numerical work, with emphasis on the scaling property (2.10b) of the band structure. All details which are necessary to reproduce our numerical results will be included.

A. Technical remarks. All trajectories have been simulated numerically by means of the "leap frog" algorithm, with typical time step $\tau_0 = 10^{-2}$; energy was conserved up to errors of order 10^{-7}.

The function $\mathcal{G}$ was generally sampled at time intervals $\tau = 0.2$, and each spectrum was obtained, via the standard Fast Fourier Transform, on the basis of $\mathcal{N} = 2^{20}$ points. The typical length of a trajectory is thus $\mathcal{T} = \tau\mathcal{N} \simeq 2 \times 10^5$. Occasionally we introduced a long transient $\mathcal{T}_0$, typically $\mathcal{T}_0 = 10^6$, before computing the spectrum, with the result that the spectrum is perfectly stable. Concerning the filter, we used Φ of the form (1.4), of course after shifting the time origin of the motion at $\mathcal{T}/2$. We used typically $\sigma = T = 0.15\,\mathcal{T}$, so the filter is quite smooth.[4] As the last technical remark, let us say that the choice $\tau = 0.2$ introduces an ultraviolet cut-off in the spectrum, at $\alpha_{\max} = 2\pi/\tau \simeq 30$. This is fairly enough to observe the lowest part of the spectrum, namely the one we are practically interested in. Anyhow, none of these choices of T, σ, τ_0, τ and $\mathcal{N}$ is really critical. Looking at continuous spectra, rather than at discrete ones, appears on the whole easier: in particular, the delicate problem of the exact location of the lines in the spectrum, which is a major reason why the sophisticated frequency analysis algorithms are necessary in connection with discrete spectra, here is not as crucial, and ordinary Fourier analysis, after filtering, turns out to be fairly sufficient.

The choice (1.2) of the observable $\mathcal{G}$ is widely arbitrary. The only attention is that $\mathcal{G}$ is not too simple (otherwise some bands can disappear or get too small). The elementary analytic structure of $\mathcal{G}$ in (1.2) is useful to get, from the beginning, a nice decreasing of the hight of the peaks, as in the analytic estimates, but is not mandatory. Finally, the choice of the potential f is also by no means critical; according to the heuristic argument reported

---

[4] It is worthwhile to note that the Fourier transform of the filter alone has the shape of a peak around $\alpha = 0$, of width of order $1/\sigma$; if applied to a quasi–periodic function, the filter turns lines into negligibly thin bands of approximate width $1/\sigma$.

Table I

$k^{(i)}$	$\mathcal{A}_i$	ψ_i	$k^{(i)}$	$\mathcal{A}_i$	ψ_i
(1,0,0)	0.5	0	(1,-1,0)	1	0
(0,1,0)	1	0	(0,1,-1)	1	0
(0,0,1)	1	0	(-1,0,1)	1	0

The table of data defining f for the three rotators model.

in Section 2C, it is only important that $\overline{f}$ contains at least two independent combinations of the angles. Practically, we assign f in the form of a Fourier polynomial

$$f(\varphi) = \sum_{i=1}^{l} \mathcal{A}_i \cos(k^{(i)} \cdot \varphi + \psi_i) \ ,$$

and specify $k^{(i)} \in \mathbf{Z}^n$ and $\mathcal{A}_i, \psi_i \in \mathbf{R}$ through a table of data.

B. Results on the band structure. We made several series of numerical computations. A first series was performed with only three rotators, f being defined by Table I. The initial actions were $I_1(0) = 1$ and $I_2(0), I_3(0) = \mathcal{O}(\sqrt{\varepsilon})$, so the resonant lattice is $\Lambda = \{k \in \mathbf{Z}^3 : k_1 = 0\}$, while $\mathcal{W}_\Lambda$ has equations $I_2 = I_3 = 0$. The exact choice of the initial datum for $(I_2, I_3, \varphi_2, \varphi_3)$ was made with the aid of a Poincaré section of the Hamiltonian $\mathcal{K}_0(J_2, J_3, \varphi_2, \varphi_3)$ introduced in Section 2C. We used the section $\varphi_3 = 0$, and observed that for energy $\mathcal{K}_0 = 1$ most motions are chaotic. Our choice was $J_2(0) = 0.5$, $\varphi_2(0) = 0$; this gives $J_3(0) \simeq 2.398$. Then we posed $I_j(0) = \sqrt{\varepsilon} J_j(0)$, $j = 2, 3$. $\varphi_1(0)$ is totally irrelevant, and we took $\varphi_1(0) = 0.543$.

For the above chosen Λ it is $\Omega = (1, 0, 0)$. According to Proposition 2, the spectrum is then expected to consist of a discrete sequence of bands, centered around $\alpha/(2\pi) = |k \cdot \Omega| = 1, 2, 3, \ldots$, with slopes proportional to $1/\sqrt{\varepsilon}$. Three spectra are shown in figure 5; the values of ε, top to bottom, are scaled by factors 4 ($\varepsilon = 1.6 \times 10^{-3}$, 4×10^{-4}, 10^{-4}); the details on the right reproduce, for each spectrum, the band around $\alpha/(2\pi) = 1$, on a frequency scale magnified, top to bottom, by factors two. The similarity of the drawings at different ε is quite evident. A quantitative check of the scaling law is not so easy, since it is not easy to determine with great precision the slope of of the peaks, who look too irregular. After a few attempts to introduce some smoothing in the spectra, and so to determine in some automatic way the slope, we concluded that the best procedure is to draw by hand, in each figure, a line similar to the one in the top right frame, and then measure its slope. The result, namely the slope versus $\sqrt{\varepsilon}$ for several values of ε, in log–log scale, is reported in figure 6 (triangles). The solid line, drawn by comparison, has slope 1/2.

A second series of numerical experiments was performed with $n = 8$, potential f specified by Table II, and initial data such that $I_1(0) = 1$ and $I(2), \ldots, I(8) = \mathcal{O}(\sqrt{\varepsilon})$, so that $d = 7$, $\Lambda = \{k \in \mathbf{Z}^8 : k_1 = 0\}$ and $\Omega = (1, 0, \ldots, 0)$. Here too, there is only one large frequency, and $|k \cdot \Omega|$ assumes discrete values. The result is that, in spite of the larger n,

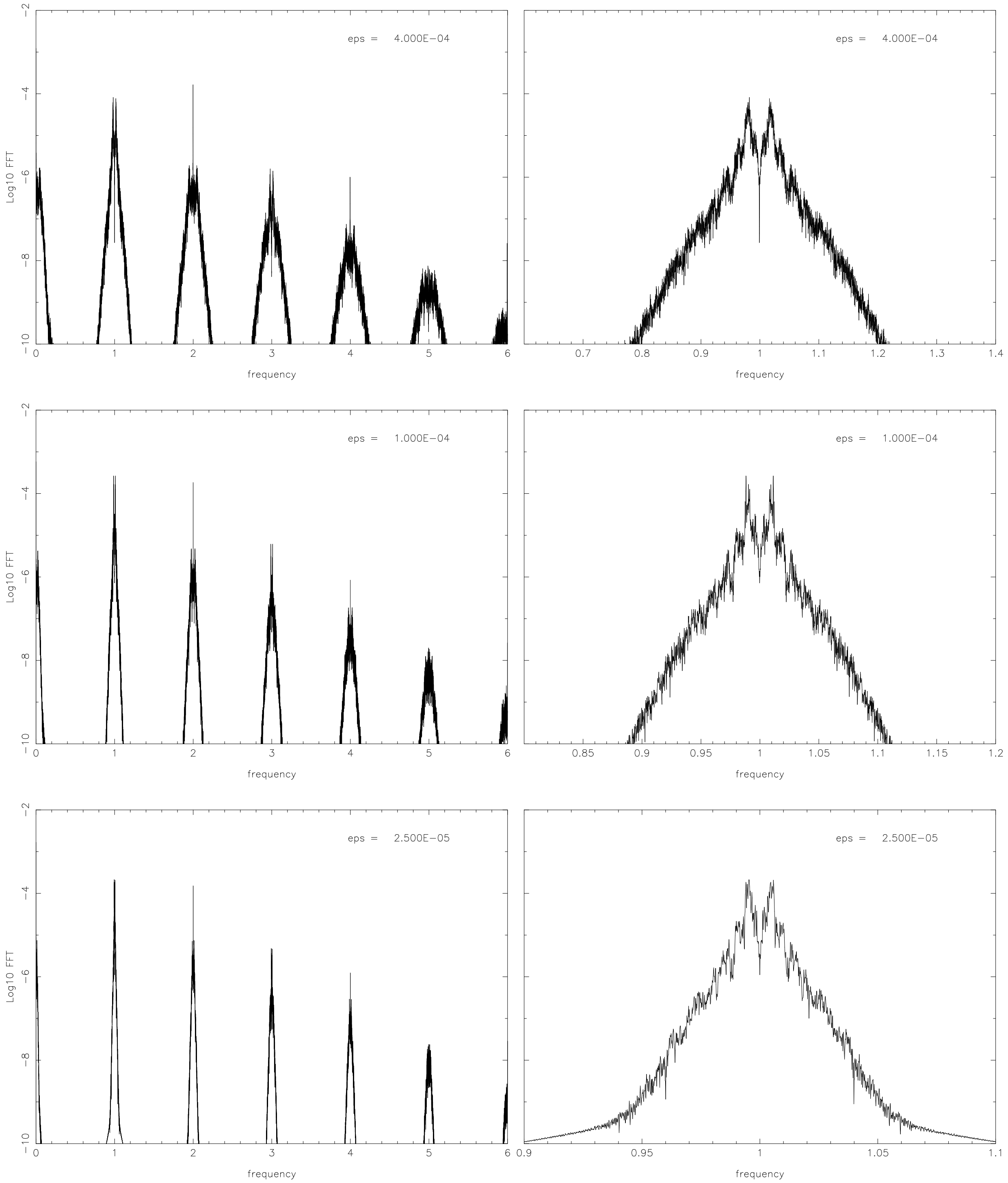

Figure 5. Three spectra for the model of three rotators. Top to bottom: $\varepsilon = 4 \times 10^{-4}$, 10^{-4}, 2.5×10^{-5}. Right part: detail of the band around $\alpha/(2\pi) = 1$ (horizontal scale magnified, top to bottom, by factors two).

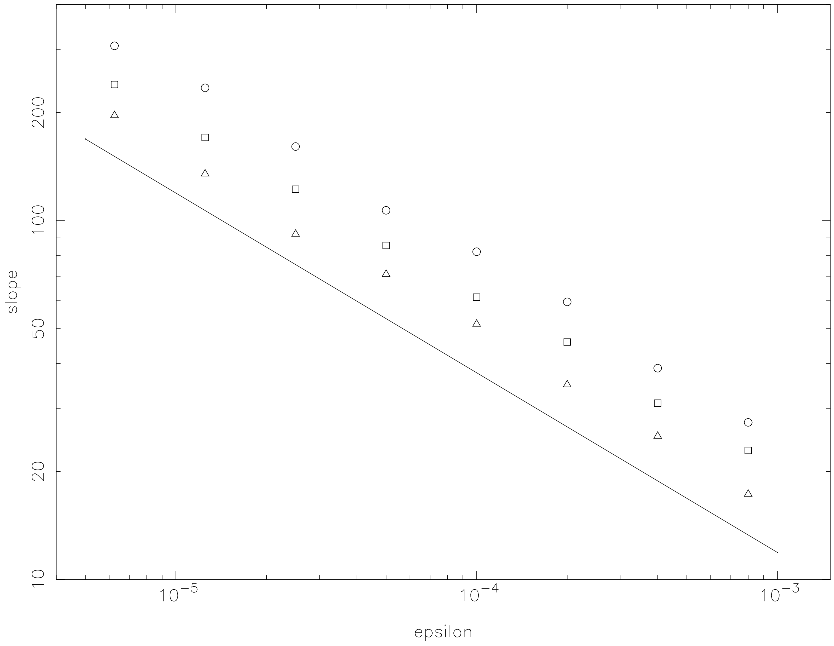

Figure 6. The slope of the peaks versus ε, log-log scale. Triangles: model of three rotators, with initially $I_1 = 1$, $I_2, I_3 = \mathcal{O}(\sqrt{\varepsilon})$. Squares: model of eight rotators, with one fast frequency (initially $I_1 = 1$, $I_2, \ldots, I_8 = \sqrt{\varepsilon}$). Circles: model of eight rotators, with two fast frequencies (initially $I_1 = 1$, $I_2 = \sqrt{2}$, $I_3, \ldots, I_8 = \mathcal{O}(\sqrt{\varepsilon})$). The solid line, drawn by comparison, has slope 1/2.

Table II

$k^{(i)}$	$\mathcal{A}_i$	ψ_i	$k^{(i)}$	$\mathcal{A}_i$	ψ_i
(1,0,0,0,0,0,0,0)	0.5	0	(0,1,-1,0,0,0,0,0)	1	0
(0,0,1,0,0,0,0,0)	0.5	0	(0,0,1,-1,0,0,0,0)	1	0
(0,0,0,0,1,0,0,0)	0.5	0	(0,0,0,1,-1,0,0,0)	1	0
(0,0,0,0,0,0,1,0)	0.5	0	(0,0,0,0,1,-1,0,0)	1	0
(1,0,0,-1,0,0,0,0)	1	0	(0,0,0,0,0,1,-1,0)	1	0
(1,0,0,0,0,-1,0,0)	1	0	(0,0,0,0,0,0,1,-1)	1	0
(-1,0,0,0,0,0,0,1)	1	0			

The table of data defining f for the eight rotators model

the spectra are nearly identical to the previous ones. An example was already produced in the Introduction, see figure 1. As before, we checked numerically the scaling law (2.10b). The result is reported in figure 6 (squares).

A more interesting third series of computations was made on the above model of eight rotators, but with different initial data, namely with $I_1(0) = 1$, $I_2(0) = \sqrt{2}$, and $I_3(0), \ldots, I_8(0)$ of order $\sqrt{\varepsilon}$. Now we have two independent frequencies, namely $\Omega = (1, \sqrt{2}, 0, \ldots, 0)$, and so the set of the possible values of $|k \cdot \Omega|$, corresponding to the centers

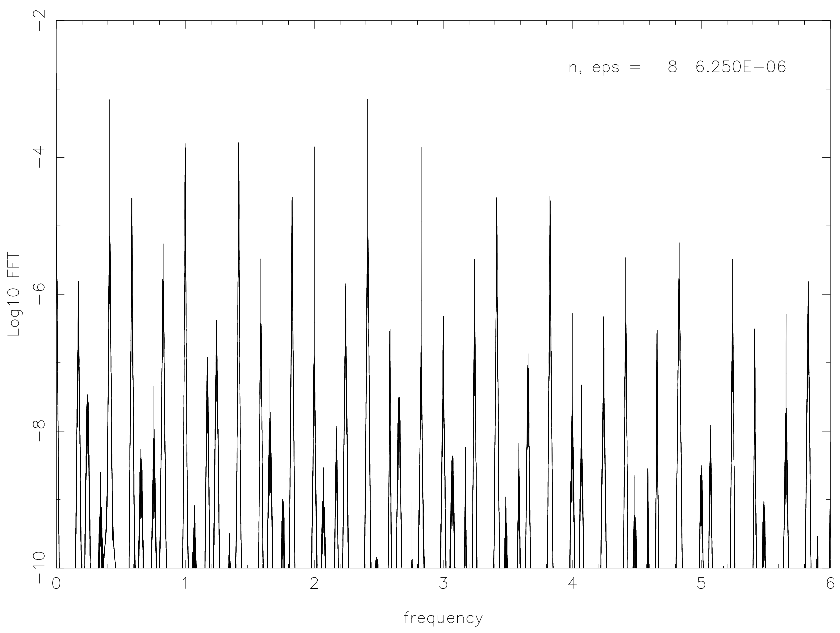

Figure 7. The spectrum for the model of eight rotators, in conditions similar to figure 2, but for ε reduced by a factor 16 ($\varepsilon = 6.25 \times 10^{-6}$).

of the bands, is larger. An example of spectrum was provided in the Introduction, see figure 2; another example, for ε reduced by a factor 16, is produced in figure 7. In principle the bands in these spectra are dense, but this requires considering large k, and according to (2.10b) the amplitude of the peaks gets negligible. The scaling law (2.10b) is illustrated in figure 6 (circles).

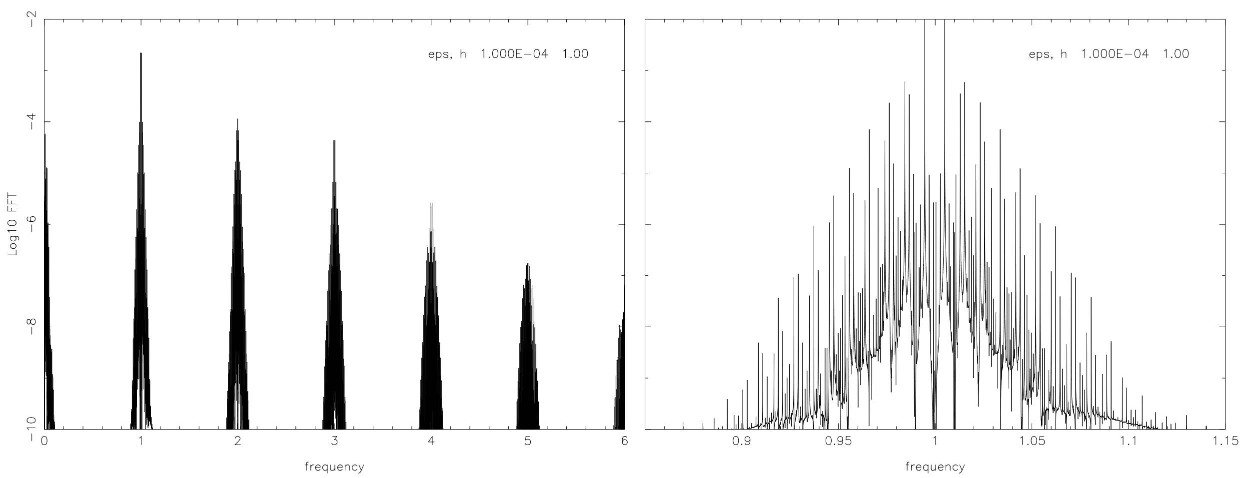

Figure 8. The spectrum for the model of three rotators, for the ordered motions which take place in the inner part of a resonance, at $\varepsilon = 10^{-4}$. The detail on the right shows that bands are now resolved into lines.

C. Motions in the inner part of resonances. For the sake of completeness, we present here some further results concerning the spectrum of resonant motions which take place in the inner part of the resonant domains. Consider again the Hamiltonian $\mathcal{K}_0$ introduced in (2.14), and let $(\varphi^*_{m+1},\ldots,\varphi^*_n)$ be a minimum of the potential $g(\varphi_{m+1},\ldots,\varphi_n)$. Quite clearly, the Hamiltonian $\mathcal{K}_0$ has an elliptic equilibrium in $(0,\ldots,0,\varphi^*_{m+1},\ldots,\varphi^*_n)$, and in a sufficiently small neighborhood of this point most motions (for KAM theorem, which is known to hold in general assumptions around elliptic equilibria) are quasi periodic. As far as the heuristic argument sketched in Section 2C is reliable, the original Hamiltonian (2.12) is also expected to exhibit quasi-periodic motions. These regular motions take place in the inner part of a resonance — the scaled actions J stay close to zero, and correspondingly the actions I perform small oscillations around the resonant manifold $\mathcal{W}_\Lambda$ — and are qualitatively different from the motions on KAM tori outside resonances. On the one hand, for quasi-periodicity, such motions must produce discrete spectra. On the other hand, Proposition 2 for them still holds, so the band structure cannot be lost. The result is shown in figure 8: the band structure persists, but now the bands are resolved into lines, as is clear from the detail at the right.[5)] The figure refers to the model of three rotators, with the same potential f as before; the initial datum was $I_1(0)=1$ and $\varphi_1(0)=0.543$, while $I_2(0), I_3(0), \varphi_2(0), \varphi_3(0)$ were chosen as before with the aid of the Poincaré section of $\mathcal{K}_0$, but at the smaller energy -0.5 (for this value of the energy it turns out that most motions are regular). The situation is pretty the same for the model of eight rotators, both in the case of one and of two independent frequencies: if the initial data are chosen in the inner part of the resonance (this requires some care), then bands are resolved into lines.

D. Spectra without the band structure. As remarked in the Introduction, see in particular figure 3, if ε is too large then the band structure is lost. This is a clear indication that the assumptions of Nekhoroshev theorem are not satisfied, and large variations of the actions in a short time are no more forbidden. As is shown by the already commented figure 4, such sudden variations do take place, although possibly much later. The model and the initial data, for figures 3 and 4, are the same as for figure 2; only ε was changed from 10^{-4} to 10^{-2}. Figure 9a provides the spectrum for the same motion as in figure 3 and 4, but on a time scale shortened by a factor 16, namely for $\mathcal{N}=2^{16}$, $\mathcal{T}\simeq 1.3\times 10^4$. As is remarkable, the shorter time is still sufficient to appreciate the lack of the band structure, and thus the possible instability of the actions, while a direct observation of the actions on the same time scale, see figure 9b, gives no information at all.

In this connection it is worthwhile to remark that the presence or absence of the band structure in the spectrum can be appreciated only by plotting the spectra in a logarithmic vertical scale. The information instead is nearly lost, if a linear scale is used for the vertical axis, too. Figure 10 reports the same spectrum as in figure 4, but in a linear scale. The presence of well defined peaks can give the impression that the motion is essentially regular, and that Nekhoroshev theorem applies. The sudden variation of the actions in Figure 4 shows instead that this is not the case. By further increasing ε, the peaks are going to disappear. This is a kind of second transition which takes place in the model, from a

[5)] To determine with precision the structure of lines inside bands, more sophisticated algorithms of frequency analysis would be necessary. But to know if a band structure obeying proposition 2 exists or not, and thus if the system fits the Nekhoroshev regime, our simple Fourier analysis looks sufficient.

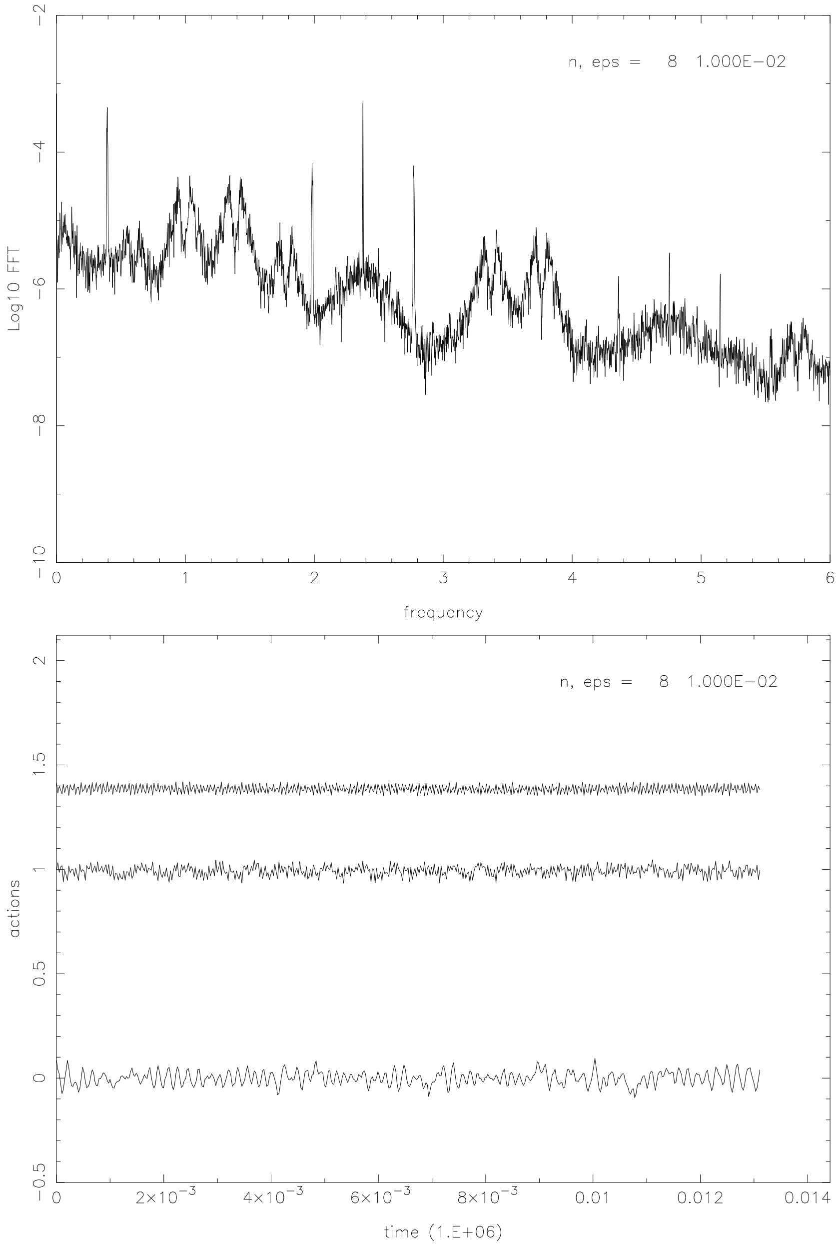

Figure 9. Upper: the spectrum for the model of eight rotators, for the same motion as in figure 3, but on a time scale shortened by a factor 16. Lower: the behavior of the actions on the same time-scale (256 times shorter than the time scale of figure 4).

moderately chaotic situation to a fully chaotic situation (very likely, this is the transition observed in [5]). But according to our analysis, already in the moderately chaotic situation the system is no more in the Nekhoroshev regime, and the actions cannot be guaranteed to be stable for very long times.

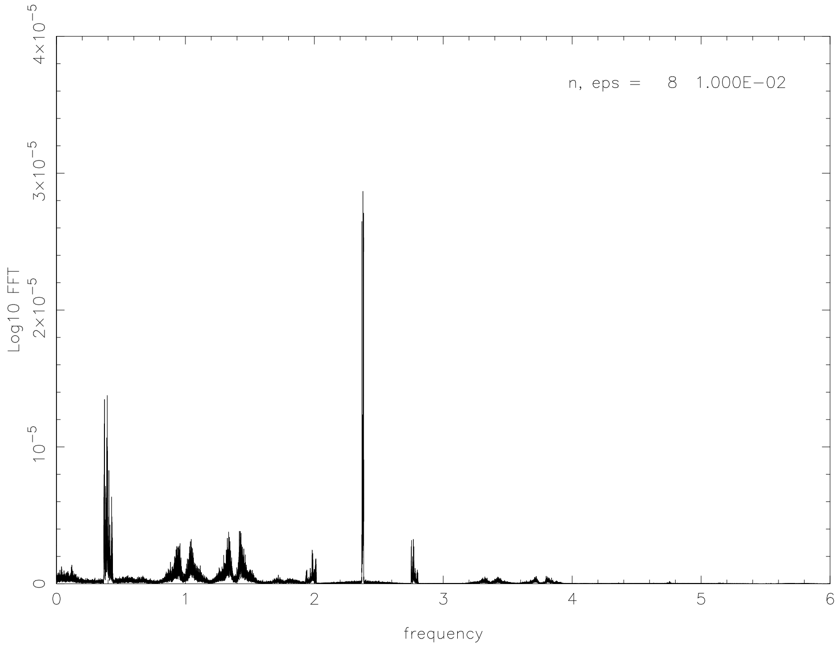

Figure 10. The spectrum already reported in figure 3, but on a linear vertical scale.

4. The proof of Propositions 1 and 2

A. Preliminaries. Before entering the proof, we must fix some further notation. For any real analytic function $u : \mathcal{D}_\rho \to \mathbf{C}$ (the complex domain $\mathcal{D}_\rho$ being defined in 2.4), with Fourier components $\hat{u}_k$, we introduce the "Fourier norm"

$$|u|_\rho = \sum_{k\in\mathbf{Z}^n} |\hat{u}_k|^\infty_{\mathcal{B}_\rho} \ e^{|k|\rho_\varphi} \ , \tag{4.1}$$

where $|\,.\,|^\infty_{\mathcal{B}}$ denotes the sup–norm in $\mathcal{B}$. Besides the already introduced extended domains $\mathcal{B}_\rho$ and $\mathcal{D}_\rho$, we shall also consider, for any resonant domain B_Λ and any extension vector $\rho' = (\rho'_I, \rho'_\varphi) \le \rho$ (the inequality is intended to work separately on the two entries), the "extended resonant domains"

$$\mathcal{B}^\Lambda_{\rho'} = \bigcup_{I\in B^\Lambda} \Delta_{\rho'}(I) \ , \qquad \mathcal{D}^\Lambda_{\rho'} = \mathcal{B}^\Lambda_{\rho'} \times \mathcal{S}_{\rho'} \ .$$

For functions $u : \mathcal{D}^\Lambda_{\rho'} \to \mathbf{C}$, in analogy with (4.1), we define

$$|u|^\Lambda_{\rho'} = \sum_{k\in\mathbf{Z}^n} |\hat{u}_k|^\infty_{\mathcal{B}^\Lambda_{\rho'}} \ e^{|k|\rho'_\varphi} \ .$$

We follow the ordinary proof of the Nekhoroshev theorem, more precisely ref. [19], and construct in each resonant domain a canonical map, giving the Hamiltonian a normal form adapted to the local resonance properties. In the Lemma below we refer to the geometry of the resonances introduced in Section 2, as determined by the following constants:

$$\omega_0 = \frac{1}{8} m \rho_I \ , \qquad c_0 = 11 \frac{M}{m} \ , \qquad \varepsilon_0 = 2^{-10} \, \frac{m \rho_I^2}{c_0^{2n} \, |f|_\rho}$$

(m, M are defined in (2.5) and (2.6) respectively).

Lemma 1. *Let $\varepsilon \le \varepsilon_0$, and consider any N-lattice Λ. Denote*

$$\rho_I^\Lambda = 8 \, \frac{\delta_\Lambda}{m} < \rho_I \ , \qquad \rho^\Lambda = (\rho_I^\Lambda, \rho_\varphi) \ , \qquad \rho_*^\Lambda = (\tfrac{1}{2}\rho_I^\Lambda, \, \tfrac{1}{6}\rho_\varphi) \ .$$

There exists a real analytic canonical map $(I,\varphi) = \mathcal{C}_\Lambda(I',\varphi')$, $\mathcal{C}_\Lambda : \mathcal{D}^\Lambda_{\rho_*^\Lambda} \to \mathcal{D}^\Lambda_{\rho^\Lambda}$, such that:

(i) H is conjugated to $H_\Lambda = H \circ \mathcal{C}_\Lambda$ of the form

$$H_\Lambda(I',\varphi') = h(I') + \varepsilon f_\Lambda(I',\varphi') + \varepsilon e^{-\frac{1}{6}N\rho_\varphi} \mathcal{R}_\Lambda(I',\varphi') \ , \tag{4.2}$$

where f_Λ has Fourier components only in Λ, i.e.

$$f_\Lambda(I',\varphi') = \sum_{k\in\Lambda} \hat f_{\Lambda,k}(I') e^{ik\cdot\varphi'} \ , \tag{4.3}$$

and f_Λ, $\mathcal{R}_\Lambda$ are bounded by

$$|f_\Lambda|^\Lambda_{\rho_*^\Lambda} \ , \qquad |\mathcal{R}_\Lambda|^\Lambda_{\rho_*^\Lambda} \le 2\,|f|_\rho \ . \tag{4.4}$$

(ii) The transformation $\mathcal{C}_\Lambda$ is near to the identity for small ε, and in particular it is

$$||I' - I|| < \tfrac{1}{64}\rho_I^\Lambda \ .$$

(iii) For any real analytic $u : \mathcal{D}^\Lambda_{\rho'} \to \mathbf{C}$, with $\rho' = (\rho'_I, \rho'_\varphi)$ such that

$$\tfrac{1}{16}\rho_I^\Lambda \le \rho'_I \le \tfrac{1}{2}\rho_I^\Lambda \ , \qquad \tfrac{1}{16}\rho_\varphi \le \rho'_\varphi \le \tfrac{1}{6}\rho_\varphi \ ,$$

it is

$$|u \circ \mathcal{C}_\Lambda|^\Lambda_{\rho' - \frac{1}{16}\rho^\Lambda} \ , \ \left|u \circ \mathcal{C}_\Lambda^{-1}\right|^\Lambda_{\rho' - \frac{1}{16}\rho^\Lambda} \ < \ 2\ |u|^\Lambda_{\rho'} \ . \tag{4.5}$$

The proof can be found in [19]. The long–time stability of the actions follows from this lemma and from the convexity of h. A possible statement, actually Nekhoroshev theorem, is the following:

Nekhoroshev theorem. *If $\varepsilon < \varepsilon_0$, there exists $t_* > 0$ such that for*

$$|t| \le T_* = t_* \, e^{\left(\frac{\varepsilon_0}{\varepsilon}\right)^{\frac{1}{2n}} \frac{\rho_\varphi}{6}}$$

it is $I'(t) \in \mathcal{B}^{\Lambda}_{\rho^{\Lambda}_*}$, *and*

$$\|I'(t) - I'(0)\| \le \tfrac{1}{3}\,\rho^{\Lambda}_I \ , \qquad \|I'(t) - I(0)\| \le \tfrac{7}{20}\,\rho^{\Lambda}_I \ , \qquad \|I(t) - I(0)\| \le \tfrac{1}{2}\,\rho^{\Lambda}_I \ . \qquad (4.6)$$

For the proof see for example [19]. (For later use, the second of (4.6), not appearing in the statement of [19], was made here explicit; moreover the constants have been slightly improved.)

B. The truncated normal form. Let us consider a N-lattice Λ such that $I(0) \in \mathcal{B}^{\Lambda}$ (the lattice Λ, from now on, should be considered as fixed). Given $\mathcal{G}$ real analytic in $\mathcal{D}_{\rho}$, we introduce its pull-back $\mathcal{F} = \mathcal{G} \circ \mathcal{C}_{\Lambda}$ (for the purpose of a light notation, in $\mathcal{F}$, as well as in other functions that we shall introduce, the dependence on Λ is omitted), which simply represents $\mathcal{G}$ in the new coordinates (I', φ'), and is analytic in $\mathcal{D}^{\Lambda}_{\rho^{\Lambda}_*}$. Denote by Φ^t and Φ^t_{Λ} the Hamiltonian flow respectively in the old and new coordinates, and let

$$(I(t), \varphi(t)) = \Phi^t(I(0), \varphi(0)) \ , \qquad (I'(t), \varphi'(t)) = \Phi^t_{\Lambda}(I'(0), \varphi'(0)) \ ,$$

with of course $(I'(0), \varphi'(0)) = \mathcal{C}^{-1}_{\Lambda}(I(0), \varphi(0))$. The function

$$g(t) = \mathcal{G}(I(t), \varphi(t)) = \mathcal{F}(I'(t), \varphi'(t)) \ ,$$

according to Nekhoroshev theorem, is well defined at least for real t not exceeding the Nekhoroshev time T_*. This function is precisely the function considered in Propositions 1 and 2, and also, after filtering, the function for which is numerically computed the spectrum. Consider now the Fourier decomposition

$$\mathcal{F}(I', \varphi') = \sum_{k \in \mathbf{Z}^n} \mathcal{F}_k(I', \varphi') \ , \qquad \mathcal{F}_k(I', \varphi') = \hat{\mathcal{F}}(I')\, e^{ik\cdot\varphi'} \ .$$

Quite trivially it is

$$g(t) = \sum_{k \in \mathbf{Z}^n} g_k(t)\, e^{ik\cdot\Omega t} \ , \qquad \text{with} \quad g_k(t) = \mathcal{F}_k(I'(t), \varphi'(t)) e^{-ik\cdot\Omega t} \ . \qquad (4.7)$$

For such a function we must eventually produce the decomposition (1.6) claimed in Proposition 1. The definition of g^0_k, however, is delicate, and cannot be produced immediately. Before, we need to exploit the analyticity properties of the truncated normal form

$$\tilde{H}_{\Lambda}(I', \varphi') = h(I') + \varepsilon f_{\Lambda}(I', \varphi') \ ,$$

and for any real parameter τ, $|\tau| \le T_*$, we introduce a preliminary decomposition

$$g_k(t) = g^{\tau,0}_k(t) + g^{\tau,1}_k(t) \ , \qquad (4.8)$$

where

$$g_k^{\tau,0}(t) = \mathcal{F}_k(\tilde\Phi_\Lambda^{t-\tau}(\Phi_\Lambda^\tau(I'(0),\varphi'(0)))e^{-ik\cdot\Omega t} \ ,$$

and $g_k^{\tau,1}(t)$ is defined by difference; $\tilde\Phi_\Lambda^t$ denotes the flow of $\tilde H_\Lambda$. It is worthwhile to remark that, due to the absence of the remainder in the truncated normal form, $g_k^{\tau,0}(t)$ is defined for any $t\in\mathbf{R}$. The properties of $g_k^{\tau,0}$ and $g_k^{\tau,1}(t)$ are stated in the lemma below.

Lemma 2. *If $\varepsilon<\varepsilon_0$ and $|\tau|\le T_*$, there exist constants b_1 and A_1 such that $g_k^{\tau,0}$ and $g_k^{\tau,1}$ are analytic respectively in*

$$\mathcal{T}_0 = \bigl\{t\in\mathbf{C} : |{\rm Im}\,t| < b_1 N^{n-d}|\Lambda|\bigr\} \ , \qquad \mathcal{T}_1 = \bigl\{t\in\mathbf{C} : |{\rm Re}\,t| < T_* \ , \ |{\rm Im}\,t| < b_1\bigr\} \ ,$$

and moreover

$$\sup_{t\in\mathcal{T}_1} |g_k| \le 2\, e^{-\frac{1}{48}|k|\rho_\varphi}\, |\mathcal{G}|_\rho \tag{4.9a}$$

$$\sup_{t\in\mathcal{T}_0} |g_k^{\tau,0}| \le 2\, e^{-\frac{1}{24}|k|\rho_\varphi}\, |\mathcal{G}|_\rho \tag{4.9b}$$

$$\sup_{t\in\mathcal{T}_1} |g_k^{\tau,1}| \le 4\, e^{-\frac{1}{48}|k|\rho_\varphi}\, |\mathcal{G}|_\rho \tag{4.9c}$$

$$\sup_{|t-\tau|<b_1} |g_k^{\tau,1}| \le \sqrt{\varepsilon}\, A_1\, e^{-\frac{1}{48}|k|\rho_\varphi}\, e^{-\frac{1}{6}N\rho_\varphi}\, |\mathcal{G}|_\rho \ . \tag{4.9d}$$

Proof: From (4.5) of Lemma 1, with a further reduction of the analyticity domain of the angles by $\frac{1}{24}\rho_\varphi$, it is not difficult to get the estimate

$$|\mathcal{F}_k|^\Lambda_{(\frac{7}{16}\rho_I^\Lambda,\frac{1}{16}\rho_\varphi)} \le 2\, e^{-\frac{1}{24}|k|\rho_\varphi}\, |\mathcal{G}|_\rho \ . \tag{4.10}$$

Standard estimates based on (4.2), (4.4) ensure that $(I'(t),\varphi'(t))$ does not escape the domain $\mathcal{D}^\Lambda_{(\frac{7}{16}\rho_I^\Lambda,\frac{1}{16}\rho_\varphi)}$ for complex t such that

$$|{\rm Re}\,t| \le T_* \quad , \quad |{\rm Im}\,t| \le t_1 = \tfrac{\rho_\varphi}{16}\left(3\|\omega(I(0)\| + 5\omega_0\frac{M}{m}\right)^{-1} \ .$$

Choosing $b_1\le t_1$ it follows:

$$\sup_{t\in\mathcal{T}_1} |g_k(t)| \le |\mathcal{F}_k|^\Lambda_{(\frac{7}{16}\rho_I^\Lambda,\frac{1}{16}\rho_\varphi)}\ e^{|k\cdot\Omega|t_1} \le 2\, e^{-\frac{1}{48}|k|\rho_\varphi}\, |\mathcal{G}|_\rho \ . \tag{4.11}$$

We can now estimate $g_k^{\tau,0}(t)$. To this purpose we use a trick, namely we observe that

$$g_k^{\tau,0}(t) = \mathcal{F}_k(\bar\Phi_\Lambda^{t-\tau}(I'(\tau),\varphi'(\tau)-\Omega\tau)) \ , \tag{4.12}$$

where $\bar\Phi_\Lambda^t$ is the flow of

$$\bar H_\Lambda(I',\varphi') = \tilde H_\Lambda(I',\varphi') - \Omega I \ ; \tag{4.13}$$

the essential point is that the evolution produced by the normal form Hamiltonian $\tilde{H}$, thanks to (4.3), commutes with a shift of the angles $\varphi' \mapsto \varphi' + \Omega s$, for any s. Elementary computations show that $\bar{\Phi}^s_\Lambda(I'(t), \varphi'(t) - \Omega t)$ does not escape $\mathcal{D}^\Lambda_{(\frac{7}{16}\rho^\Lambda_I, \frac{1}{16}\rho_\varphi)}$ for $t \in \mathbf{R}$ with $|\mathrm{Re}\, t| \le T_*$ and $s \in \mathbf{C}$ with $|s| \le t_1 N^{n-d} |\Lambda|$; the estimate (4.9b) immediately follows.

The next estimate (4.9c) trivially follows from (4.9b) and (4.11). The local estimate (4.9d) requires instead some work, that we just sketch. From the very definition of $g^{\tau,1}_k$ it is

$$g^{\tau,1}_k(t) = e^{-ik\cdot\Omega t} \left[\mathcal{F}_k(I^{(1)}(t), \varphi^{(1)}(t)) - \mathcal{F}_k(I^{(2)}(t), \varphi^{(2)}(t))\right] , \tag{4.14}$$

where

$$(I^{(1)}(t), \varphi^{(1)}(t)) = \Phi^{t-\tau}_\Lambda(I'(\tau), \varphi'(\tau)) , \qquad (I^{(2)}(t), \varphi^{(2)}(t)) = \mathcal{F}(\tilde{\Phi}^{t-\tau}_\Lambda(I'(\tau), \varphi'(\tau)) .$$

The flows $\Phi^{t-\tau}_\Lambda$ and $\tilde{\Phi}^{t-\tau}_\Lambda$ are very similar, since the respective Hamiltonians H_Λ and $\tilde{H}_\Lambda$ differ only by the exponentially small remainder defined and estimated in Lemma 1. More precisely, the distance

$$\Delta(t) = \sum_{i=1}^n \left(\frac{\left| I^{(1)}_i(t) - I^{(2)}_i(t)\right|}{\rho^\Lambda_I} + \left|\varphi^{(1)}_i(t) - \varphi^{(2)}_i(t)\right| \right)$$

has a growth bounded by

$$\dot{\Delta} \le \lambda_1 \int_0^{t-\tau} \Delta(s) ds + \lambda_2 \sqrt{\varepsilon}\, e^{-\frac{1}{6}N\rho_\varphi}$$

and consequently (Gronwall lemma)

$$\Delta(t) \le \lambda_2 \sqrt{\varepsilon}\, e^{-\frac{1}{6}N\rho_\varphi}\, e^{\lambda_1|t-\tau|} ,$$

where λ_1 is a Lipshitz constant of the Hamiltonian vector field of $\tilde{H}_\Lambda$, and λ_2 is a bound on the Hamiltonian vector field of the remainder. The estimate of λ_1, λ_2 are easy but annoying: they are based on Cauchy estimates of the second derivatives of $\tilde{H}_\Lambda$ and, respectively, of the first derivatives of $\mathcal{R}_\Lambda$. Possible values of these constants are:

$$\begin{aligned} \lambda_1 &= n \left(\frac{4}{\rho_\varphi} \left|h\right|_\rho + \frac{1}{2} \max\left(1, \frac{1}{\rho^2_\varphi}\right)\right) \\ \lambda_2 &= 32 n \left(1 + \frac{1}{\rho_\varphi}\right) \sqrt{m \left|f\right|_\rho} \end{aligned} . \tag{4.15}$$

From (4.14) it is now easy to obtain:

$$\sup_{|t-\tau|\le\frac{t_1}{2}} |g^{\tau,1}_k| \le \sqrt{\varepsilon}\, A_1\, e^{-\frac{1}{48}|k|\rho_\varphi}\, e^{-\frac{1}{6}N\rho_\varphi} \left|\mathcal{G}\right|_\rho$$

with the following possible value for the constant A:

$$A = 2^{12} n \left(1 + \frac{1}{\rho_\varphi}\right) e^{\lambda_1 \frac{t_1}{2}} \sqrt{m \left|f\right|_\rho} \quad .$$

A possible value for the constant b_1 is $b_1 = t_1/2$.

C. Definition of the g_k^0 functions Let $T \in (T_0, T_*/3)$ and $0 < \beta \leq b_1/2$ free parameters, and let Φ be a filter function defined as in (1.4):

$$\Phi(t) = \frac{1}{2}\Big[\tanh\Big(\frac{t+2T}{\sigma}\Big) - \tanh\Big(\frac{t-2T}{\sigma}\Big)\Big] \ , \tag{4.16}$$

with $\sigma = \beta N^{n-d} \left|\Lambda\right|$, and $\sigma < T$. For any $k \in \mathbf{Z}^n$ we define g_k^0 in the following way:

$$\begin{aligned} g_k^0(t) =& \frac{1}{2\pi} \sum_{j=0}^{[12T/\beta]} \int_0^{+\infty} d\alpha \int_{\tau_j}^{\tau_{j+1}} dx \Phi(x+i\sigma) g_k^{\tau_j,0}(x+i\sigma) e^{(i(x-t)-\sigma)\alpha} + \\ & \frac{1}{2\pi} \sum_{j=0}^{[12T/\beta]} \int_{-\infty}^{0} d\alpha \int_{\tau_j}^{\tau_{j+1}} dx \Phi(x-i\sigma) g_k^{\tau_j,0}(x-i\sigma) e^{(i(x-t)+\sigma)\alpha} \quad , \end{aligned} \tag{4.17}$$

where $\tau_j = -3T + j\frac{\beta}{2}$ if $j \leq [12T/\beta] - 1$ and $\tau_{[12T/\beta]} = 3T$. It is easy to see that all integrations in (4.17) are convergent for $\left|\mathrm{Im}\, t\right| < \sigma$, and moreover the Fourier transform of g_k^0,

$$\tilde{g}_k^0(\alpha) = \begin{cases} e^{-\alpha\sigma} \sum_{j=0}^{[12T/\beta]} \int_{\tau_j}^{\tau_{j+1}} dx \Phi(x+i\sigma) g_k^{\tau_j,0}(x+i\sigma) e^{i\alpha x} & \text{if } \alpha \geq 0 \\ e^{\alpha\sigma} \sum_{j=0}^{[12T/\beta]} \int_{\tau_j}^{\tau_{j+1}} dx \Phi(x-i\sigma) g_k^{\tau_j,0}(x-i\sigma) e^{i\alpha x} & \text{if } \alpha < 0 \end{cases} \ ,$$

satisfies the estimate

$$\left|\tilde{g}_k^0(\alpha)\right| \leq 9T e^{-\left|\alpha\right|\sigma} e^{-\left|k\right|\frac{\rho_\varphi}{24}} \left|\mathcal{G}\right|_\rho^V$$

for any $\alpha \in \mathbf{R}$ (the trivial estimate $\sup_{\left|\mathrm{Im}\, t\right| \leq \sigma} \left|\Phi(t)\right| \leq 3/2$ has been used). Therefore, all g_k^0 satisfy statement *a)* and inequality (2.8b) of Proposition 1. In the next subsections *D*,*E* we prove that the function g^1 defined by difference

$$g^1(t) = g(t) - \sum_{k \in \mathbf{Z}^n} g_k^0(t) e^{ik\cdot\Omega t} \ , \tag{4.18}$$

satisfies statement *b)* of Proposition 1, and consequently g_k^0 satisfies (2.8a).

D. Proof of Proposition 1. We show that g^1 can be written as the sum of four functions $u^i(t)$, $i = 1, \ldots, 4$, all of which will be proven to be exponentially small (subsection *E*) for t satisfying $\left|\mathrm{Re}\, t\right| \leq T$ and $\left|\mathrm{Im}\, t\right| \leq \beta/2$.

We start by multiplying g by the filter function Φ. In the time interval $t \in (-T, T)$ the filtered function differs from the unfiltered one only for the exponentially small term

$u^1 = (1-\Phi)g$ (see (4.26)). Then, for any $t \in (-3T, 3T)$ we represent the filtered function $\Phi(t)g(t)$ by means of the following Fourier representation:

$$\Phi(t)g(t) = \frac{1}{2\pi}\int d\alpha e^{-i\alpha t}\int_{-3T}^{3T} dx\Phi(x)g(x)e^{i\alpha x} \ . \tag{4.19}$$

We can replace g in (4.19) with its expansion (4.7), thus obtaining:

$$\Phi(t)g(t) = \frac{1}{2\pi}\int d\alpha e^{-i\alpha t}\int_{-3T}^{3T} dx \sum_{k\in\mathbf{Z}^n}\Phi(x)g_k(x)e^{ik\cdot\Omega x}e^{i\alpha x} = \sum_{k\in\mathbf{Z}^n}\hat{g}_k(t)e^{ik\cdot\Omega t} \tag{4.20}$$

where $\hat{g}_k$ is defined (after the trivial change of variable $\alpha \mapsto \alpha + k\cdot\Omega$) by

$$\hat{g}_k(t) = \frac{1}{2\pi}\int d\alpha e^{-i\alpha t}\int_{-3T}^{3T} dx\Phi(x)g_k(x)e^{i\alpha x} \ . \tag{4.21}$$

Because of the analyticity of the $g_k(x)$ in $\mathcal{T}_1$, we can lift the integration path of the x variable from the real path $\{s \in (-3T, 3T)\}$ to the complex path $\{s + i\beta\,,\ s \in (-3T, 3T)\}$, if $\alpha \geq 0$, or to $\{s - i\beta\,,\ s \in (-3T, 3T)\}$, if $\alpha < 0$. Precisely, it is:

$$\begin{aligned}\hat{g}_k(t) = &-\frac{1}{2\pi}\int_0^{\infty} d\alpha e^{-i\alpha t}e^{-|\alpha|\beta}\int_{-3T}^{3T} dx\Phi(x+i\beta)g_k(x+i\beta)e^{i\alpha x}\\ &-\frac{1}{2\pi}\int_{-\infty}^{0} d\alpha e^{-i\alpha t}e^{-|\alpha|\beta}\int_{-3T}^{3T} dx\Phi(x-i\beta)g_k(x-i\beta)e^{i\alpha x} + u_k^2(t)\end{aligned} \tag{4.22}$$

where u_k^2 is given by the integrations of x on the complex paths $\{\pm 3T \pm is, s \in (0,\beta)\}$. Setting $u^2 = \sum_k u_k^2 e^{ik\cdot\Omega t}$, it turns out that u^2 for $|\mathrm{Im}\, t| < \beta$ is exponentially small (see (4.28)).

In view of the evaluation of the difference $g_k - g_k^0$, we replace g_k in (4.22) with its representation $g_k^{\tau_j,0} + g_k^{\tau_j,1}$ in any interval $x \in [\tau_j, \tau_{j+1}]$. Thus, we obtain:

$$\begin{aligned}\hat{g}_k(t) = &-\frac{1}{2\pi}\sum_{j=0}^{[12T/\beta]}\int_0^{\infty} d\alpha e^{-i\alpha t}e^{-|\alpha|\beta}\int_{\tau_j}^{\tau_{j+1}} dx\Phi(x+i\beta)g_k^{\tau_j,0}(x+i\beta)e^{i\alpha x} -\\ &\frac{1}{2\pi}\sum_{j=0}^{[12T/\beta]}\int_{-\infty}^{0} d\alpha e^{-i\alpha t}e^{-|\alpha|\beta}\int_{\tau_j}^{\tau_{j+1}} dx\Phi(x-i\beta)g_k^{\tau_j,0}(x-i\beta)e^{i\alpha x} + u_k^2(t) + u_k^3(t)\ ,\end{aligned} \tag{4.23}$$

where u_k^3 corresponds to the integrations of the $g_k^{\tau_j,1}$, which are exponentially small in $\{s \pm i\beta, s \in [\tau_j, \tau_{j+1}]\}$, and therefore, setting $u^3 = \sum_k u_k^3 e^{ik\cdot\Omega t}$, it turns out that u^3 is exponentially small (see (4.29)).

Because of the analyticity of the $g_k^{\tau_j,0}$ in $\mathcal{T}_0$ we can further lift the integration path of x to $\{s + i\sigma, s \in (-3T, 3T)\}$, if $\alpha \geq 0$, or to $\{s - i\sigma, s \in (-3T, 3T)\}$, if $\alpha < 0$. Precisely, it is:

$$
\begin{aligned}
\hat{g}_k(t) =& \frac{1}{2\pi}\sum_{j=0}^{[12T/\beta]}\int_0^{\infty} d\alpha e^{-i\alpha t}e^{-|\alpha|\sigma}\int_{\tau_j}^{\tau_{j+1}} dx\Phi(x+i\sigma)g_k^{\tau_j,0}(x+i\sigma)e^{i\alpha x}- \\
& \frac{1}{2\pi}\sum_{j=0}^{[12T/\beta]}\int_{-\infty}^{0} d\alpha e^{-i\alpha t}e^{-|\alpha|\sigma}\int_{\tau_j}^{\tau_{j+1}} dx\Phi(x-i\sigma)g_k^{\tau_j,0}(x-i\sigma)e^{i\alpha x}+ \\
& u_k^2(t)+u_k^3(t)+u_k^4(t) = g_k^0(t)+u_k^2(t)+u_k^3(t)+u_k^4(t) \quad ,
\end{aligned}
\tag{4.24}
$$

where u_k^4 is given by the integrations in the complex paths $\{\tau_j \pm is, s \in (0,\sigma)\}$, $j = 0,\ldots,[12T/\beta]$. Setting $u^4 = \sum_k u_k^4 e^{ik\cdot\Omega t}$, it turns out that u^4 for $|\mathrm{Im}\, t| \leq \beta/2$ is exponentially small (see (4.34)).

E. Technical estimates. Before estimating $u^1,\ldots,u^4$ we recall some trivial estimates on the filter Φ. Indeed, working directly on definition (1.4), it is easy to prove the following inequalities:

$$
\sup_{|\mathrm{Re}\, t|\leq 3T\,,\,|\mathrm{Im}\, t|\leq\sigma} |\Phi(t)| \leq \frac{3}{2} \tag{4.25a}
$$

$$
\sup_{|\mathrm{Re}\, t|\leq T\,,\,|\mathrm{Im}\, t|\leq\sigma} |\Phi(t)-1| \leq 2e^{-2\frac{T}{\sigma}} \tag{4.25b}
$$

$$
\sup_{|\mathrm{Re}\, t|\geq T'\,,\,|\mathrm{Im}\, t|\leq\sigma} |\Phi(t)| \leq 2e^{-2\frac{T'}{\sigma}}e^{4\frac{T}{\sigma}} \tag{4.25c}
$$

for any $T' \geq 3T$.

(i) ESTIMATE OF u^1. Since $(I(t),\varphi(t))$ does not leave $\mathcal{B}_\rho$ within a time t such that $|\mathrm{Re}\, t| \leq T_*$ and $|\mathrm{Im}\, t| \leq b_1$ (see the proof of Lemma 2), it is $|g(t)| \leq |\mathcal{G}|_\rho$. Therefore, from (4.25b) it follows

$$
\sup_{|\mathrm{Re}\, t|\leq T\,,\,|\mathrm{Im}\, t|\leq b_1} \left|u^1(t)\right| = \sup_{|\mathrm{Re}\, t|\leq T\,,\,|\mathrm{Im}\, t|\leq b_1} |g(t)(1-\Phi(t))| \leq 16e^{\frac{-2T}{\sigma}}\,|\mathcal{G}|_\rho^V \quad . \tag{4.26}
$$

(ii) ESTIMATE OF u^2. The integrals defining $u_k^2(t)$, namely

$$
\begin{aligned}
u_k^2(t) =& \frac{1}{2\pi}\int_0^{\infty} d\alpha e^{-i\alpha t}\int_0^{\beta} ds\Phi(3T+is)g_k(3T+is)e^{i3T\alpha}e^{-|\alpha|s}- \\
& \frac{1}{2\pi}\int_0^{\infty} d\alpha e^{-i\alpha t}\int_0^{\beta} ds\Phi(-3T+is)g_k(-3T+is)e^{-i3T\alpha}e^{-|\alpha|s}+ \\
& \frac{1}{2\pi}\int_{-\infty}^{0} d\alpha e^{-i\alpha t}\int_0^{\beta} ds\Phi(-3T-is)g_k(-3T-is)e^{-i3T\alpha}e^{-|\alpha|s}- \\
& \frac{1}{2\pi}\int_{-\infty}^{0} d\alpha e^{-i\alpha t}\int_0^{\beta} ds\Phi(3T-is)g_k(3T-is)e^{i3T\alpha}e^{-|\alpha|s} \quad ,
\end{aligned}
\tag{4.27}
$$

are absolutely convergent if $|\mathrm{Im}\, t| < \beta$. We can then exchange the order of the integrations and perform first the integration over α. So for example, the first integral becomes

$$\int_0^\beta ds \frac{\Phi(3T+is)g_k(3T+is)}{i(\mathrm{Re}\, t - 3T) + s - \mathrm{Im}\, t} \quad ,$$

and with with the help of (4.9b) and (4.25c) one can easily show that its abs–value is smaller than $4\beta e^{-2\frac{T}{\sigma}} e^{-|k|\frac{\rho\varphi}{48}} |\mathcal{G}|_\rho/(T\pi)$, provided that $|\mathrm{Re}\, t| \leq T$ and $|\mathrm{Im}\, t| < \beta$. The other integrals in (4.27) can be estimated in exactly the same way. The result is:

$$\sup_{|\mathrm{Re}\, t|\leq T\ ,\ |\mathrm{Im}\, t|<\beta} \left|u^2(t)\right| \leq \frac{16\beta}{\pi T}\left(1 - e^{-\frac{\rho\varphi}{64}}\right)^{-1} e^{-2\frac{T}{\sigma}} |\mathcal{G}|_\rho \quad . \tag{4.28}$$

(iii) ESTIMATE OF u^3. The integrals defining u_k^3, namely

$$\begin{aligned} u^3(t) = &-\frac{1}{2\pi}\sum_{j=0}^{[12T/\beta]}\int_{-\infty}^0 d\alpha e^{-i\alpha t}e^{-|\alpha|\beta}\int_{\tau_j}^{\tau_{j+1}} dx\Phi(x+i\beta)g_k^{\tau_j,1}(x+i\beta)e^{i\alpha x} \\ &-\frac{1}{2\pi}\sum_{j=0}^{[12T/\beta]}\int_0^\infty d\alpha e^{-i\alpha t}e^{-|\alpha|\beta}\int_{\tau_j}^{\tau_{j+1}} dx\Phi(x-i\beta)g_k^{\tau_j,1}(x-i\beta)e^{i\alpha x}\ , \end{aligned}$$

are absolutely convergent if $|\mathrm{Im}\, t| < \beta$. These integrals are easily estimated with the help of inequality (4.9d). On the complex domain defined by $|\mathrm{Im}\, t| \leq \beta/2$ one gets

$$\sup_{|\mathrm{Im}\, t|\leq\frac{\beta}{2}} \left|u^3(t)\right| \leq \frac{18\beta}{\pi}\sqrt{\epsilon}A_1\frac{t_*}{\beta}\left(1 - e^{-\frac{\rho\varphi}{64}}\right)^{-1} e^{-N\frac{\rho\varphi}{12}} |\mathcal{G}|_\rho \quad . \tag{4.29}$$

(iv) ESTIMATE OF u^4. The integrals defining $u_k^4(t)$ are absolutely convergent if $|\mathrm{Im}\, t| < \beta$. Therefore, it is:

$$\sup_{|\mathrm{Im}\, t|\leq\frac{\beta}{2}} \left|u_k^4(t)\right| \leq \frac{1}{2\pi}(C_1 + C_2)\int_0^\infty d\alpha \int_\beta^\sigma ds\ e^{-|\alpha|\left(s-\frac{\beta}{2}\right)} \tag{4.30}$$

where

$$C_1 = 4\sup_{|s|\leq\sigma}\left|\Phi(\pm 3T \pm is)g_k^{\pm 3T,0}(\pm 3T \pm is)\right|$$

$$C_2 = 2\sup_{|s|\leq\sigma}\left(\sum_{j=0}^{[12T/\beta]-2}\left|\Phi(\tau_{j+1}\pm is)\right|\left|(g_k^{\tau_j,0} - g_k^{\tau_{j+1},0})(\tau_{j+1}\pm is)\right|\right) \quad .$$

The integral in (4.30) is easily evaluated:

$$\int_0^\infty d\alpha\int_\beta^\sigma e^{-|\alpha|\left(s-\frac{\beta}{2}\right)} = \ln\left(\frac{2\sigma-\beta}{\beta}\right) \leq \ln\left(\frac{2\sigma}{\beta}\right) \leq \ln(2N^n)$$

and with the help of (4.25c) and (4.9b), the constant C_1 is easily estimated as well:

$$C_1 \le 16 e^{-2\frac{T}{\sigma}} e^{-|k|\frac{\rho_\varphi}{24}} \left|\mathcal{G}\right|_\rho \quad .$$

Instead, the estimate of the constant C_2 requires a little attention. We first rewrite $(g_k^{\tau_j,0} - g_k^{\tau_{j+1},0})(\tau_{j+1} \pm is)$ using (4.12):

$$\begin{aligned}(g_k^{\tau_j,0} - g_k^{\tau_{j+1},0})(\tau_{j+1} \pm is) =& \mathcal{F}_k(\bar{\Phi}_\Lambda^{\tau_{j+1}-\tau_j+is}(I'(\tau_j), \varphi'(\tau_j) - \Omega\tau_j)) \\ &- \mathcal{F}_k(\bar{\Phi}_\Lambda^{is}(I'(\tau_{j+1}), \varphi'(\tau_{j+1}) - \Omega\tau_{j+1}))\end{aligned} \tag{4.31}$$

where $\bar{\Phi}_\Lambda^t$ denotes the Hamiltonian flow of (4.13). Then, since $\bar{\Phi}_\Lambda^t$ and $\tilde{\Phi}_\Lambda^t$ both commute with a shift of the angles $\mathcal{C}^t : \varphi' \mapsto \varphi' + \Omega t$ and $\bar{\Phi}_\Lambda^t = \mathcal{C}^{-t}\tilde{\Phi}_\Lambda^t$, (4.31) can be rewritten as:

$$\begin{aligned}(g_k^{\tau_j,0} - g_k^{\tau_{j+1},0})(\tau_{j+1} \pm is) =& \mathcal{F}_k(\mathcal{C}^{-\tau_{j+1}}\bar{\Phi}_\Lambda^{is}\tilde{\Phi}_\Lambda^{\tau_{j+1}-\tau_j}(I'(\tau_j), \varphi'(\tau_j))) \\ &- \mathcal{F}_k(\mathcal{C}^{-\tau_{j+1}}\bar{\Phi}_\Lambda^{is}\Phi_\Lambda^{\tau_{j+1}-\tau_j}(I'(\tau_j), \varphi'(\tau_j))) \quad .\end{aligned} \tag{4.32}$$

The difference between Φ_Λ^t and $\tilde{\Phi}_\Lambda^t$ was already estimated in the proof of Lemma 2. More precisely, one proves that the distance

$$\Delta = \sum_{i=1}^n \left(\frac{\left|I_i^{(1)} - I_i^{(2)}\right|}{\rho_I^\Lambda} + \left|\varphi_i^{(1)} - \varphi_i^{(2)}\right| \right) \quad ,$$

where $(I^{(1)}, \varphi^{(1)}) = \Phi_\Lambda^{\tau_{j+1}-\tau_j}(I'(\tau_{j+1}), \varphi'(\tau_{j+1}))$ and $(I^{(2)}, \varphi^{(2)}) = \tilde{\Phi}_\Lambda^{\tau_{j+1}-\tau_j}(I'(\tau_j), \varphi'(\tau_j))$, is bounded by:

$$\Delta \le \lambda_2 \sqrt{\varepsilon}\, e^{-\frac{1}{6}N\rho_\varphi}\, e^{\lambda_1 \frac{b}{2}} \ , \tag{4.33}$$

where λ_1 and λ_2 are the constants given in (4.15). Moreover, elementary computations show that both $\mathcal{C}^{-\tau_{j+1}}\bar{\Phi}_\Lambda^{is}(I^{(1)}, \varphi^{(1)})$ and $\mathcal{C}^{-\tau_{j+1}}\bar{\Phi}_\Lambda^{is}(I^{(2)}, \varphi^{(2)})$ do not escape $\mathcal{D}^\Lambda_{(\frac{2}{5}\rho_I^\Lambda, \frac{1}{32}\rho_\varphi)}$, if $|s| \le t_1 N^{n-d} |\Lambda|/2$, and so, from (4.10) and (4.33), one deduces:

$$\left|(g_k^{\tau_j,0} - g_k^{\tau_{j+1},0})(\tau_{j+1} \pm is)\right| \le 64\lambda_2 \left(1 + \frac{1}{\rho_\varphi}\right) \sqrt{\epsilon} e^{\lambda_1 \frac{\beta}{2}} e^{-|k|\frac{\rho_\varphi}{24}} e^{-N\frac{\rho_\varphi}{6}} \left|\mathcal{G}\right|_\rho \quad .$$

By collecting all these estimates we then obtain:

$$\sup_{|\mathrm{Im}\, t| \le \frac{\beta}{2}} \left|u^4(t)\right| \le \frac{8}{\pi} \frac{\ln(2N^n)}{1 - e^{-\frac{\rho_\varphi}{64}}} \left(e^{-2\frac{T}{\sigma}} + 3\lambda_2 \left(1 + \frac{1}{\rho_\varphi}\right) \sqrt{\epsilon} e^{\lambda_1 \frac{\beta}{2}} e^{-N\frac{\rho_\varphi}{6}} \right) \left|\mathcal{G}\right|_\rho \quad . \tag{4.34}$$

We now conclude the proof by observing that, from inequalities (4.26), (4.28), (4.29) and (4.34) follows inequality (2.9) with

$$\begin{aligned} a_3 &= 3\sqrt{\epsilon_0} \left(1 - e^{-\frac{\rho_\varphi}{64}}\right)^{-1} \left(\frac{6t_*}{\pi} A_1 + \lambda_2 \left(1 + \frac{1}{\rho_\varphi}\right) e^{\lambda_1 \frac{b}{2}} \right) \frac{\left|\mathcal{G}\right|_{\mathcal{D}_\rho}^\infty}{\left|\mathcal{G}\right|_\rho} \\ a_4 &= 16 + \frac{8}{\pi} \left(1 - e^{-\frac{\rho_\varphi}{64}}\right)^{-1} \left(2\frac{\beta}{T} + 1\right) \frac{\left|\mathcal{G}\right|_{\mathcal{D}_\rho}^\infty}{\left|\mathcal{G}\right|_\rho} \quad . \end{aligned}$$

Finally, (2.8a) can be obtained, quite trivially, by estimating g_k^0 as the difference between Φg_k and $u_k^2 + u_k^3 + u_k^4$. For constants β, t_0 we can choose $\beta = b_1/2$ and $t_0 = \min(t_*/3, \beta)$. The proof of Proposition 1 is therefore concluded.

We conclude this section with a remark on the proof of Proposition 2. We do not report here the details of such a proof because it can be done with minor changes to the proof of Proposition 1, taking into account a result proven by Pöschel on the exponential stability near resonances in perturbed convex systems. Specifically, we refer to Lemma 1 of [19], from which one can prove that if $I(0) \in B_\Lambda$ and $\|P_\Lambda I(0)\| \le \sigma_0\sqrt{\varepsilon}$, then for an exponentially long time t it is $\|P_\Lambda I(t)\| \le \sigma_1\sqrt{\varepsilon}$, where σ_1 is a constant greater than σ_0, but independent of ε. Therefore, one can follow the lines of the proof of Proposition 1, replacing δ_Λ and ρ_I^Λ with $\sqrt{\varepsilon}$.

References

[1] Arnol'd V.I. (1963): "Proof of A. N. Kolmogorov's theorem on the conservation of conditionally periodic motions with a small variation in the Hamiltonian." *Russian Math. Surv.*, **18**, No. 5, 9-36.

[2] Arnol'd V.I. (1963): "Small denominators and problems of stability of motion in classical and celestial mechanics." *Russ. Math. Surveys*, **18**, 85–191.

[3] Benettin G., Fassò F. and Guzzo M. (1998): "Nekhoroshev–stability of L4 and L5 in the spatial restricted three–body problem." *Regular and chaotic dynamics*, **3**, 3.

[4] Benettin G., Galgani L. and Giorgilli A. (1985): "A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems." *Cel. Mech.*, **37**, 1.

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

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

[7] Fassò F. (2000): "Stability properties of the Riemann ellipsoids." Preprint.

[8] Giorgilli A. and Morbidelli A. (1997): "Invariant KAM tori and global stability for Hamiltonian systems." *ZAMP*, **48**, n. 1, 102–134.

[9] Guzzo M. and Morbidelli A. (1997): "Construction of a Nekhoroshev like result for the asteroid belt dynamical system." *Celest. Mech. Dyn. Astron.*, **66**, 255–292.

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

[11] Laskar J. (1990): “The chaotic motion of the Solar System: a numerical estimate of the size of the chaotic zones.” *Icarus*, **88**, 299-291.

[12] Laskar J., Froeschlé C. and Celletti A. (1992): “The measure of chaos by the numerical nalysis of the fundamental frequencies. Application to the standard mapping.” *Physica D*, **56**, 253-269.

[13] Laskar J. (1993): “Frequency analysis for multi–dimensional systems. Global dynamics and diffusion.” *Physica D*, **67**, 257-281.

[14] Lochak P. (1992): “Canonical perturbation theory via simultaneous approximation.” *Russ. Math. Surv.*, **47**, 57–133.

[15] Morbidelli A. and Guzzo M. (1997): “The Nekhoroshev theorem and the asteroid belt dynamical system.” *Celest. Mech. Dyn. Astron.*, **65**, 107-136.

[16] Moser J. (1958): “New Aspects in the Theory of Stability of Hamiltonian Systems.” *Comm. on Pure and Appl. Math.*, **11**, 81–114.

[17] Nekhoroshev N.N. (1977): “Exponential estimates of the stability time of near–integrable Hamiltonian systems.” *Russ. Math. Surveys*, **32**, 1-65.

[18] Nekhoroshev N.N. (1979): “Exponential estimates of the stability time of near–integrable Hamiltonian systems, 2.” *Trudy Sem. Petrovs.*, **5**, 5-50.

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