QUANTITATIVE PERTURBATION THEORY BY SUCCESSIVE ELIMINATIONS OF HARMONICS

ALESSANDRO MORBIDELLI
Département de Mathématique FUNDP
8, Rempart de la Vierge, B–5000 Namur, Belgium

ANTONIO GIORGILLI
Dipartimento di Matematica dell' Universitá di Milano
Via Saldini 50, 20133 Milano, Italy

Abstract. We revisit some results of perturbation theories by a method of successive elimination of harmonics inspired by some ideas of Delaunay. On the one hand, we give a connection between the KAM theorem and the Nekhoroshev theorem. On the other hand, we support in a quantitative fashion a semi–numerical method of analysis of a perturbed system recently introduced by one of the authors.

Key words. Perturbation methods, KAM theorem, Nekhoroshev theorem, action–angle variables.

1. Introduction

In recent times, some interest arose on the possible application of the methods of KAM theory and of the Nekhoroshev like results to problems in celestial mechanics. However, a direct application of these methods in practical problems is generally an hard task, since they are based on iteration schemes well suited for analytical estimates, but not for an explicit computation. The present paper follows this research line, by trying to justify from a quantitative viewpoint a perturbation method inspired by some ideas of Delaunay (1867). The advantage of this method is that it can actually, and effectively, be applied, at least for systems of low dimension, as illustrated in a recent paper by Morbidelli (1991). Here, we give a quantitative analysis of the method, with the aim of giving it a complete theoretical support. Precisely, we show that a suitable adaptation of the method of Delaunay can be used to recover, in a common environment, both the KAM theorem and part of the Nekhoroshev theorem. Although from a strictly mathematical viewpoint the results are not new, we believe that the most interesting aspect of the present paper is that, at least in the present case, we are filling the gap between practical computations and theoretical results in perturbation theory.

Let us recall in some more details the main ideas. A direct use of the perturbation methods underlying the KAM and the Nekhoroshev theories is often impossible, unless one considers very simple cases. The main reason is that in the traditional approach to both the theories one has to deal with different kinds of expansion of a near to integrable Hamiltonian system, and to apply all the known apparatus of canonical transformations. Doing such a work explicitly is usually an hard task. Indeed, the number of terms in the expansion becomes soon so large that it is often almost impossible to perform the computation in a reasonable time. The usual way out of this difficulty is to consider only *local* expansions; for example, one considers a neighbourhood either of a periodic orbit or of an equilibrium point (Giorgilli et al., 1989; Simó, 1989; Celletti and Giorgilli, 1991). The latter method usually allows to reach interesting results, but it is nevertheless hard to obtain a *global* description of the dynamics.

We follow here a different approach. Precisely, we perform a process of successive elimination of harmonics in the Fourier expansion of the Hamiltonian: this is indeed the idea of Delaunay. How to actually perform such an elimination is suggested by the Arnold's construction of the action–angle variables (Arnold, 1963b). First, one ignores all the harmonics in the perturbation but a single one. This is an integrable system, and the introduction of the action–angle variables requires only the evaluation of an integral. Such an operation allows a complete elimination of the considered harmonic without remainder terms. Next, one has to compute the expansion of the till now ignored terms in the new variables, and to iterate the process. How to practically perform such a procedure is illustrated in the quoted paper by Morbidelli. However, from a theoretical viewpoint, it seems not immediately evident that the method is based on a sufficiently established mathematical background. For instance, the choice of the order of elimination of the harmonics is rather arbitrary, and is justified only in a qualitative fashion. Moreover, there is no evidence *a priori* that the already eliminated harmonics do not reappear with bigger and bigger coefficients: this is indeed a major problem in classical expansions.

As already anticipated, we show here that the method can be fully justified. In particular, we obtain the following results.

(i) For a sufficiently small perturbation and for an arbitrary time T, there exists a large open region of the phase space which is foliated in n–dimensional tori; these tori are almost invariant up to time T, in the sense that an orbit starting on a torus remains in a small neighbourhood of it for $|t| < T$. This is stated by our theorem 1 below.

(ii) In the limit of the time T going to infinity, the region above shrinks to a set of invariant tori (which is neither open nor dense). This set could be identified with the invariant tori of the KAM theory; we refer in particular to the formulation of the results in the paper of Arnold (1963a). We do not include an explicit proof of this statement, since it is nothing but the Arnold's result.

(iii) If one requires that the open region above contains open balls of a given (small) radius ε, then the time T must be finite. However, T can be as large as $\exp(1/\varepsilon)^a$, with a suitable exponent a. This result looks as a bridge between the KAM theory and the Nekhoroshev theory, at least for what concerns the nonresonant region. The formal statement is given in theorem 2 below.

The possible relevance of these facts from a practical viewpoint is that we are able to deal with an open set of initial conditions. The unpleasant aspect, is that such a set contains holes, due to resonances, that we are unable to describe.

Understanding these facts requires a more technical discussion. Instead of looking for invariant tori, like in KAM theory, we look for quasi invariant tori which lie on neighbourhoods of the set of KAM tori. To this end, we perform a finite number of perturbation steps. At each step we introduce new action–angle variables which are close to, but do not coincide with, the original unperturbed ones; the difference is usually referred to as "deformation". As a function of the new action angle variables the Hamiltonian has an integrable part, depending only on the actions, and a small remainder depending on all variables. The construction of these action–angle variables above cannot be performed unless one removes from phase space the resonant regions —the holes in the final domain. At each step of the procedure one has the following picture: the new action variables parameterize a set of tori filling up an open region; we call these tori quasi invariant because the time derivative of the corresponding action is small; thus, an orbit stays in a small neighbourhood of a quasi invariant torus for a long time —roughly, at least the inverse of the size of the remainder.

The connection with the KAM theory shows that things are much more complicated. Indeed, we could in principle perform any number of perturbation steps, thus getting a better and better confinement of the orbits for longer and longer times. In the limit of infinite steps one would get the set of KAM tori. This result is implicitly contained in the Arnold's proof of the existence of invariant tori, and we limit ourselves in stressing this usually non considered content of the theorem. However, as is well known, the set of invariant tori has a large measure, but its complement is open, dense and connected. Thus, the KAM result is meaningful from a probabilistic viewpoint, but could hardly be compatible with the experimental errors in determining the initial data.

From the latter viewpoint, the connection with Nekhoroshev theory is perhaps more interesting. To this end, we just stop our perturbation procedure when the set of quasi invariant tori is not guaranteed to contain open balls of a given radius. In such a way, we get a confinement of orbits for times which are exponentially large with the inverse of the radius of the ball. We stress here the following points.

(i) Our statement holds in an open set of large measure. In contrast, Nekhoroshev's theorem holds in the whole phase space.

(ii) The distance of the orbit from a quasi invariant torus is estimated to be exponentially small over an exponentially large time interval.

(iii) The possibility of an exponential bound on the speed of the diffusion process is strictly related to the request that the set of quasi invariant tori contains balls of a suitable size.

Let us add a few comments. On the one hand, point (i) means that our result is weaker than the Nekhoroshev's one. On the other hand, point (ii) gives a statement which looks much stronger than the usual Nekhoroshev's estimate. In fact, the locality of our result, as opposed to the globality of the Nekhoroshev's one, is due to the fact that we are considering only nonresonant regions. In such regions (and only in such regions) we find quasi invariant tori which are close to the unperturbed ones. This explains also why our estimate in (ii) is definitely better than the Nekhoroshev's one. Indeed, our estimate concerns action variables which are constant up to a small term due to the remainder. In the usual Nekhoroshev's approach instead one always refers to the unperturbed action variables describing the unperturbed tori. Thus, the main contribution to his estimate is due to the deformation introduced by the perturbation procedure. The size deformation is in fact rather big, being of order ε^b, with $b < 1$.

A comparison with some previous works is also interesting. A first reference should be made to a very interesting paper by Escande (19xx). In that paper, a similar method of successive elimination of harmonics was used. The idea of Escande was to locate an invariant torus (in his case the so called golden torus) with a method strictly related to the renormalization group in Physics. In particular, he makes essential use of the strict relation existing between the golden number and the Fibonacci's sequence. This allowed him to give the first perturbation steps an explicit form. In contrast, we consider generic nonresonance conditions, and give only estimates on the effect of the transformations. Thus, our treatment is less detailed, but can be applied to any nonresonant region of the phase space. In addition, we are also able to give quantitative estimates, in contrast with the formal approach of Escande.

A second reference concerns a recent work of Lochak (1989); see also Pöschel (1991). His approach is based on consideration of completely resonant unperturbed tori, i.e., tori which are foliated into strictly periodic orbits. In the neighbourhood of such a torus one can give a local stability result, in the spirit of Nekhoroshev theory. The global result is based on a theorem coming from number's theory: roughly, every unperturbed invariant torus can be approximated by a completely resonant one. This can be considered as a clever use of the fact that the completely resonant regions are, in some sense, the most stable ones. Our approach complements the Lochak's one. Indeed, we restrict our attention to regions with nonresonant tori, thus getting a very good confinement of the orbits. The basis of our approach is that the non-resonant regions occupy a large fraction of the phase space. Thus, one is confronted with the paradoxical situation that the most stable regions are the less relevant from a probabilistic viewpoint.

The paper is organized as follows. In section 2 we formulate our results as mathematical theorems. In section 3 we illustrate the method of elimination of a single harmonic. In section 4 we show how the successive elimination of harmonics connects with the quadratic method commonly used in KAM theory. Finally, in section 5, we give the proofs of the theorems. The technical sections 6 and 7 contain the proofs of the main lemmas and a few technical statements.

2. Formulation of the results

We start with some definitions and notations.

In $\mathbf{C}^n$ we introduce the distance $|p - p'| = \max_j |p_j - p'_j|$.

For an open subset $\mathcal{G}$ of $\mathbf{C}^n$ we shall denote by $\mathcal{G} - \delta$, with positive δ, the subset of $\mathcal{G}$ of the points which are contained in $\mathcal{G}$ together with a δ–neighbourhood; a δ–neighbourhood is an open ball of radius δ in the distance above.

Considering the immersion of $\mathbf{T}^n$ in $\mathbf{C}^n$, with coordinates q, we shall denote by $\mathbf{T}^n_\sigma$ the strip $|\operatorname{Im} q| \le \sigma$.

The basic domain of our construction will be $\mathcal{D} = \mathcal{G} \times \mathbf{T}^n_\sigma$. Moreover, we shall denote by $\mathcal{D} - (\delta_1, \delta_2) = (\mathcal{G} - \delta_1) \times \mathbf{T}^n_{\sigma-\delta_2}$. By abuse of notation we shall also denote $\mathcal{D} - \delta = (\mathcal{G} - \delta) \times \mathbf{T}^n_\sigma$.

We are also interested in the volume of the real part of either $\mathcal{D}$ or $\mathcal{G}$. By $\operatorname{Vol}\mathcal{D}$ and $\operatorname{Vol}\mathcal{G}$ we shall denote such a volume. We remark that, from $\mathcal{D} = \mathcal{G} \times \mathbf{T}^n_\sigma$, the relation $\operatorname{Vol}\mathcal{D} = (2\pi)^n \operatorname{Vol}\mathcal{G}$ holds.

Finally, if $f(p,q)$ is an analytic function defined on $\mathcal{D}$, we shall denote with $\|f\|_\mathcal{D}$ the supremum norm, i.e., $\|f\|_\mathcal{D} = \sup_\mathcal{D} |f(p,q)|$. For a vector function $g = (g_1, \ldots, g_n)$, the norm $\|g\|_\mathcal{D}$ will be defined as the maximum of the norms of the components. For vectors $v \in \mathbf{C}^n$ the notation $\|v\|$ will instead denote the Euclidean norm.

With these settings we can state our

Theorem 1: *Let the Hamiltonian*

$$H(p,q) = H_0(p) + \varepsilon H_1(p,q)$$

be real analytic in $\mathcal{D} = \mathcal{G} \times \mathbf{T}^n_\sigma$*, the set* $\mathcal{G}$ *being open in* $\mathbf{C}^n$*; assume that, for some positive constants* Λ, ϑ*,* Θ*,* ε*, and* σ *with* $0 < \vartheta \le \Theta < +\infty$*, the Hamiltonian* H *satisfies:*

a) $\left\|\frac{\partial H_0}{\partial p}\right\|_{\mathcal{D}} \le \Lambda$

b) $\vartheta\|v\| \le \|A(p)v\| \le \Theta\|v\|$ *for any* $v \in \mathbf{C}$, *where* $A(p) = \left(\frac{\partial^2 H_0}{\partial p_i \partial p_j}\right)$

c) $\|H_1\|_{\mathcal{D}} \le \varepsilon$

d) $\left\|\frac{\partial H_1}{\partial p}\right\|_{\mathcal{D}} \le \varepsilon^{3/4}$

Assume moreover that, for some positive constants ϱ, *and* D_V, *the domain* $\mathcal{G}$ *satisfies:*

e) $\mathrm{Vol}(\mathcal{G} \setminus (\mathcal{G} - d)) \le d D_V \,\mathrm{Vol}\,\mathcal{G}$ *for* $0 < d < \varrho$.

Then there exist constants $\overline{\varepsilon}, C^*, C^*_\sigma, V$ *and a positive function* $T_*(\varepsilon)$ *such that the following statement holds true:*

For any $\varepsilon < \overline{\varepsilon}$ *and for any* $T > T_*(\varepsilon)$, *having defined* $\eta = \varepsilon^{1/4}(T_*/T)^{1/3}$, *there exists an open domain* $\mathcal{D}'(\varepsilon, T) \subset \mathcal{D}$ *such that:*

1) $\mathcal{D}'(\varepsilon, T)$ *is made up of* n*–dimensional tori* $\mathcal{T}$ *parameterized by action variables* P *canonically conjugated to angles* Q, *with* $p = P + f(P,Q), q = Q + g(P,Q)$ *and*

$$\|f\|_{\mathcal{D}'} \le 2C^*\varepsilon^{1/2} \ , \quad \|g\|_{\mathcal{D}'} \le 2C^*_\sigma \varepsilon^{1/4} \ ;$$

2) $\mathrm{Vol}(\mathcal{D}' - \eta) \ge \left\{1 - V\varepsilon^{1/16}\left[1 - \left(\frac{T_*}{T}\right)^{1/6}\right]\right\}\mathrm{Vol}\,\mathcal{D}$

3) $\forall (P_0, Q_0) \in \mathcal{D}' - \eta$, *the orbit* $P(t), Q(t)$ *with initial data* P_0, Q_0 *satisfies*

$$|P(t) - P_0| \le \eta \quad \text{for } |t| \le T$$

The existence of quasi invariant tori stated in theorem 1 suggests a possible connection with the Nekhoroshev's theorem. The key point here is that Nekhoroshev's result is global in the action space, while our theorem, like the KAM theorem, gives results on a quite strange subset of the phase space. A minimal connection between Nekhoroshev and KAM theory requires that the open domain $\mathcal{D}'(\varepsilon, T)$ of theorem 1 also does contain balls of a suitable radius. From a practical viewpoint this could be considered as a request of compatibility with the errors in experimental data. As a matter of fact, such a request produces an upper bound on the time T in our statement, which turns out to be essentially the exponential time appearing in Nekhoroshev theory. We remark, on the one hand, that our result is still local, in the sense that it is only valid in a neighbourhood of the set of the invariant tori of KAM theory; on the other hand, we are able to work in action angle variables adapted to the perturbation.

We now state the following

Theorem 2: *In the same hypotheses of theorem 1 there exist a constant* $\overline{\varepsilon}_1 \le \overline{\varepsilon}$, *with* $\overline{\varepsilon}$ *as in theorem 1, and a constant* α *not dependent on* ε *such that for any* $\varepsilon < \overline{\varepsilon}_1$ *the domain*

$$\mathcal{D}'(\varepsilon, T) \quad \text{with } T = \frac{2}{\sigma\varepsilon^{3/4}} \exp\left[\frac{3}{8}\left(\frac{\alpha}{\varepsilon}\right)^{1/(3n-2)}\right]$$

contains balls of radius ε.

3. Elimination of a single harmonic

The basic step in our perturbation method is the elimination of a single harmonic in the Fourier expansion of the Hamiltonian. This is performed via the construction of suitable action–angle variables according to the known framework established by the Arnold–Liouville theorem. Our scheme is inspired by the effective presentation of Henrard (1990); we just show that all the operations can be consistently performed in a complex domain, and that we can produce quantitative estimates.

To start with, we consider the integrable Hamiltonian system

$$H(p,q) = H_0(p) + h_1(p)\cos(k\cdot q)\ , \tag{1}$$

where $p\in\mathcal{G}\subset\mathbf{C}^n$ and $q\in\mathbf{T}^n_\sigma$. We also assume that the domain $\mathcal{G}$ is non–resonant, namely that the quantity $k\cdot\omega(p)$ is bounded away from zero. The first step we make, is to transform the Hamiltonian to the form

$$\tilde H(I,\varphi) = \tilde H_0(I) + \tilde h_1(I)\cos(\lambda\varphi_1)\ ; \tag{2}$$

this is obtained by a linear canonical transformation defined by a unimodular matrix, i.e., a matrix with integer entries and determinant 1. The coefficient λ turns out to be the greatest common divisor among the components of k. This amounts to consider a system which is essentially one–dimensional, which makes the extension to a complex domain straightforward. For such a system, the new action variables $J_1,\ldots,J_n$ turn out to be $J_l = I_l$ for $l=2,\ldots,n$, and

$$J_1 = \frac{1}{2\pi}\oint_\Gamma I_1\,{\rm d}\varphi_1\ , \tag{3}$$

where I_1 has to be expressed as a function of E,φ_1 and the remaining actions $I_2,\ldots,I_n$ by inverting the relation $\tilde H(I,\varphi)=E$. The extension of this definition to the complex domain is straightforward. Indeed: the inversion can be performed also for complex variables and for complex values of E; the cycle Γ on the real torus is nothing but the interval $[0,2\pi]$, and can be continuously deformed into an arbitrary closed curve on a complex neighbourhood of the real torus; the value of the integral (3) clearly turns out to be independent of the curve.

Finally, we introduce the last set of action–angle variables R,r by performing the inverse of the linear canonical transformation above. The variables R,r are close to p,q in the sense that they are defined by the relations

$$p = R + f(R,k\cdot r)\ ,\quad q = r + g(R,k\cdot r)\ , \tag{4}$$

where f and g turn out to be roughly of the same size as h_1.

A precise statement involving quantitative estimates is given by the following

Lemma 1: *Let*

$$H(p,q) = H_0(p) + h_1(p)\cos(k\cdot q)\ , \tag{5}$$

with $0 \neq k \in \mathbf{Z}^n$*, be real analytic in the domain* $\mathcal{D} = \mathcal{G}\times\mathbf{T}^n_\sigma$ *for some positive* σ*, and assume that for positive constants* $\gamma,\Lambda,C,M,\Theta,\vartheta$*, with* $M < \gamma/(|k|\exp(|k|\sigma))$ *and* $0<\vartheta\le\Theta<+\infty$*, one has*

a) $\inf_{p\in\mathcal{G}}\left|k\cdot\frac{\partial H_0}{\partial p}\right| \ge \gamma > 0$;

b) $\left\|\frac{\partial H_0}{\partial p}\right\|_{\mathcal{G}} \le \Lambda$;

c) $\vartheta\|v\| \le \|A(p)v\| \le \Theta\|v\|$ *for any* $v\in\mathbf{C}^n$*, where* $A(p)$ *is the Hessian matrix of* H_0 *with respect to* p*, namely* $A(p) = \left(\frac{\partial^2 H_0}{\partial p_j\partial p_i}\right)$;

d) $\|h_1\|_{\mathcal{G}} \le C\ ,\quad \left\|\frac{\partial h_1}{\partial p}\right\|_{\mathcal{G}} \le M$.

Define

$$\delta_0 = 2(\pi+|k|\sigma)\frac{|k|Ce^{|k|\sigma}}{\gamma-|k|Me^{|k|\sigma}}$$

$$\delta_1 = \frac{4e^{|k|\sigma}(\pi+|k|\sigma)}{\gamma-|k|Me^{|k|\sigma}}\left(\frac{2\sqrt{n}\Theta|k|(\pi+|k|\sigma)C}{\gamma-|k|Me^{|k|\sigma}}+M\right)$$

and assume moreover that $\mathcal{G}-2\delta_0$ *is non–empty and that* $\delta_1<\sigma/2$*. Then, there exists a real analytic canonical transformation* $\mathcal{C}:\mathcal{D}-(\delta_0,\delta_1)\to\mathcal{D}$ *such that the following holds:*

1) $\mathcal{D}-2(\delta_0,\delta_1)\subset\mathcal{C}(\mathcal{D}-(\delta_0,\delta_1))\subset\mathcal{D}$,

2) denoting by R,r *the new variables, the transformation can be given the form*

$$\begin{aligned} p &= R + f(R,k\cdot r)\ ,\\ q &= r + g(R,k\cdot r)\ , \end{aligned} \tag{6}$$

with functions f *and* g *periodic in* $\psi_1 = k\cdot r$ *satisfying*

$$\|f\|_{\mathcal{D}-(\delta_0,\delta_1)} \le \delta_0\ ,\quad \|g\|_{\mathcal{D}-(\delta_0,\delta_1)} \le \delta_1\ ;$$

3) the transformed Hamiltonian $H' = H\circ\mathcal{C}$ *is independent of the new angles* r*, namely one has* $H'=H'(R)$*; moreover, for any positive* δ_2 *such that* $\mathcal{G}-(2\delta_0+\delta_2)$ *is non empty, the transformed Hamiltonian* H' *satisfies (b) and (c) with new constants* Λ',ϑ' *and* Θ' *defined by*

$$\begin{aligned} \Lambda' &= \Lambda + Me^{|k|\sigma}\ ,\\ \vartheta' &= \vartheta - \frac{n\Lambda\delta_0 + Ce^{|k|\sigma}}{\delta_2^2}\ ,\\ \Theta' &= \Theta + \frac{n\Lambda\delta_0 + Ce^{|k|\sigma}}{\delta_2^2}\ . \end{aligned} \tag{7}$$

4. Iterative elimination of harmonics and quadratic step

The aim now is to use the result of Lemma 1 as a base for the successive elimination of harmonics from the Fourier expansion of the Hamiltonian.

We consider the complete Hamiltonian

$$H(p,q) = H_0(p) + H_1(p,q)\ , \tag{8}$$

and expand H_1 in Fourier series of the angles q. For simplicity we assume that the expansion contains only cos terms, namely it has the form

$$H_1(p,q) = \sum_{k\in\mathbf{Z}^n} h_k(p)\cos(k\cdot q)\ . \tag{9}$$

Pick now an arbitrary k, for instance by looking for the largest coefficient h_k, and perform the transformation of lemma 1. To do this it is clearly necessary to reduce the domain $\mathcal{G}$, where the original actions are defined, by removing a strip of a suitable size around the resonant manifold $k\cdot\omega(p)=0$; actually, we remove the strip $|k\cdot\omega(p)|<\gamma$ with a suitable γ. Under that canonical transformation the new Hamiltonian takes the form

$$H'(R,r) = H'_0(R) + H'_1(R,r) + \mathcal{R}(R,r)\ , \tag{10}$$

where H'_1 has the same Fourier expansion as H_1 but without the harmonic k which has been eliminated, and $\mathcal{R}$ contains all the terms generated by the introduction of the new variables in each term of the original Fourier expansion of H_1. The problem here is to estimate the size of $\mathcal{R}$. To this purpose we state the following

Lemma 2: *Consider a real analytic function*

$$\overline{h}(p,q) = \overline{h}_1(p)\cos(m\cdot q)$$

defined on a domain $\mathcal{D}=\mathcal{G}\times\mathbf{T}^n_\sigma$ *and assume that*

a) $\left\|\overline{h}_1\right\|_{\mathcal{G}} \le C_1 < +\infty$,

b) $\left\|\frac{\partial\overline{h}_1}{\partial p}\right\|_{\mathcal{G}} \le M_1 < +\infty$,

where C_1, M_1 *are positive constants. Consider now the canonical transformation* $p(R,r), q(R,r)$ *the existence and the properties of which are proved in Lemma 1. Consider moreover the function* $\overline{h'}(R,r) = \overline{h}(p(R,r),q(R,r))$ *defined on* $\mathcal{D}-(\delta_0,\delta_1)$*; then* $\overline{h'}(R,r)$ *can be written as*

$$\overline{h'}(R,r) = \overline{h}_1(R)\cos(m\cdot r) + \mathcal{R}_{k,m}(R,r)$$

with an analytic function $\mathcal{R}_{k,m}$ *satisfying*

$$\|\mathcal{R}_{k,m}\|_{\mathcal{D}-(\delta_0,\delta_1)} \le n e^{|m|\sigma}(M_1\delta_0 + |m|C_1\delta_1 + n|m|M_1\delta_0\delta_1)\ .$$

This lemma shows on the one hand that if the perturbation H_1 is of size ε, then $\mathcal{R}$ turns out to be roughly of size ε^2 (apart from non–negligible coefficients that have to be taken into account in a quantitative statement); on the other hand the analyticity of the Hamiltonian implies that the coefficients $h_k(p)$ decay exponentially fast, namely that $|h_k| \sim e^{-|k|\sigma}$ for some σ. Thus, we can proceed with the elimination of harmonics disregarding all the new terms in $\mathcal{R}$ generated by the transformation until $|k|$ reaches a value K_0 roughly of size $|\ln\varepsilon|$. At that point we are left with the part of H_1 of size ε^2, namely of the same size as $\mathcal{R}$. Thus, we expand also $\mathcal{R}$ in Fourier series, reorder all the terms, and start the game again with a perturbation of size ε^2. The elimination of all the harmonics satisfying $|k| \le K_0$ is what we call a quadratic step, because of it's evident connection with the quadratic method commonly used in the proof of the KAM theorem.

The main problem here is concerned with the elimination of the resonant strips. As in Arnold's formulation of the KAM theorem, here we look for a region in the phase space of large volume where our results hold. In principle one is interested in regions in the real action space; however, the canonical transformations introduce deformations of the action variables depending on the angles, and this makes the estimates of the volumes of the resonant strips impractical. As suggested by Arnold, the way out of this difficulty is to consider domains in the real part of the whole phase space. Indeed the canonical transformation is known to preserve volumes in phase space, so that the estimate of a volume is independent of the set of canonical variables used.

The quantitative analysis of a single quadratic step leads to the following iterative

Lemma 3: *Let*

$$H(p,q) = H_0(p) + H_1(p,q)$$

with $\int_{\mathbf{T}^n} H_1(p,q)dq = 0$ *be real analytic on a domain* $\mathcal{D} = \mathcal{G} \times \mathbf{T}^n_\sigma$ *for some positive* σ*. Assume that for some positive constants* $\varepsilon, \Lambda, \vartheta, \Theta, D_V$*, with* $0 < \vartheta \le \Theta < +\infty$ *the following hypotheses hold:*

a) $\|H_1\|_{\mathcal{D}} \le \varepsilon$;

b) $\left\| \frac{\partial H_1}{\partial p} \right\|_{\mathcal{D}} \le \varepsilon^{3/4}$;

c) $\left\| \frac{\partial H_0}{\partial p} \right\|_{\mathcal{D}} \le \Lambda < +\infty$;

d) $\vartheta\|v\| \le \|Av\| \le \Theta\|v\|$ *for any* $v \in \mathbf{C}^n$*, where* $A(p) \equiv \left(\frac{\partial^2 H_0}{\partial p_j \partial p_i} \right)$;

e) for some positive ϱ *the domain* $\mathcal{G}$ *has the following property:* $\mathrm{Vol}(\mathcal{G} \setminus (\mathcal{G} - d)) \le dD_V \,\mathrm{Vol}(\mathcal{G})$ *for all* $0 \le d \le \varrho$.

Then, there exist positive constants C_H*,* C_Λ*,* C_ϑ*,* C_σ*,* $C_{\mathcal{G}}$*,* C_Δ*,* A, B*,* $\overline{\varepsilon}$ *dependent on* n, σ*,* $\vartheta, \Theta, \Lambda$ *such that for any* δ_F *satisfying* $\varepsilon^{1/16(n+1)} \le \delta_F < 1/4$*, and for any* $\varepsilon < \overline{\varepsilon}$ *there exist a domain* $\mathcal{D}'$ *and a canonical transformation* $\mathcal{C} : \mathcal{D}' \to \mathcal{D}$ *with the following properties:*

1) $\mathcal{D} \supset \mathcal{D}' - \varepsilon^{1/4} \neq \emptyset$;
2) $W' := \mathrm{Vol}(\mathcal{D} \setminus (\mathcal{D}' - \varepsilon^{1/4})) \leq \left[D_V C_\Delta + A \frac{(B - \ln \varepsilon)^{n-1}}{\delta_F^{n-1}} \right] \frac{\varepsilon^{1/4}}{\delta_F^{n+2}} \mathrm{Vol}(\mathcal{D})$;
3) $\mathcal{D}'$ admits a foliation $\mathcal{G}' \times \mathbf{T}^n_{\sigma'}$ with

$$
\begin{aligned}
&\sigma' \geq \sigma(1 - \delta_F) - C_\sigma \varepsilon^{1/4} > \sigma/2 \ , \\
&\mathrm{Vol}(\mathcal{G}' \setminus (\mathcal{G}' - d)) \leq d D'_V \, \mathrm{Vol}(\mathcal{G}) \quad \text{for all } 0 \leq d \leq \varrho' \quad \text{with} \\
&\varrho' = \varrho - C_\Delta \varepsilon^{1/8} \ , \\
&D'_V = D_V + 2^n C_{\mathcal{G}} \left(\frac{\Theta'}{\vartheta'} \right)^n \left(\frac{B - \ln \varepsilon}{\sigma \delta_F} \right)^{n-1} \ , \\
&\vartheta' = \vartheta - C_\vartheta \varepsilon^{1/8} \ , \\
&\Theta' = \Theta + C_\vartheta \varepsilon^{1/8} \ ;
\end{aligned}
$$

4) $|p - p'| \leq C_\Delta \varepsilon^{1/2}$, $\quad |q - q'| \leq C_\sigma \varepsilon^{1/4}$;
5) the Hamiltonian H' in the variables $p', q' \in \mathcal{D}'$ can be written in the form $H'(p', q') = H'_0(p') + H'_1(p', q')$ with $\int_{\mathbf{T}^n} H'_1(p', q') dq' = 0$, and

$$
\|H'_1\|_{\mathcal{D}'} \leq \varepsilon' \ , \quad \left\| \frac{\partial H'_1}{\partial p'} \right\|_{\mathcal{D}' - \varepsilon^{1/4}} \leq (\varepsilon')^{3/4} \quad \text{with } \varepsilon' = \varepsilon^{5/4} C_H
$$

$$
\left\| \frac{\partial H'_0}{\partial p'} \right\|_{\mathcal{D}' - \varepsilon^{1/4}} \leq \Lambda' \equiv \Lambda + C_\Lambda \varepsilon^{1/2}
$$

moreover, for $p \in \mathcal{G}' - \varepsilon^{1/4}$ *one has*

$$
\vartheta' \|v\| \leq \|A' v\| \leq \Theta' \|v\| \quad \forall v \in \mathbf{C}^n
$$

where Θ' *and* ϑ' *are as above, and*

$$
A' \equiv \left(\frac{\partial^2 H'_0}{\partial p'_j \partial p'_i} \right) \ .
$$

5. Iteration and proofs of the theorems

The proof of theorem 1 requires the iteration of quadratic steps and an estimate of the total loss of volume. The rough argument we use is as follows. According to lemma 3, after a number N of quadratic steps the Hamiltonian has the form

$$
H^{(N)}(p, q) = H_0^{(N)}(p) + H_1^{(N)}(p, q) \ , \tag{11}
$$

where the remainder $H_1^{(N)}$ turns out to be of size $C_N\varepsilon^{2^N}$, with a non negligible constant C_N. As a matter of fact, C_N grows so fast that the actual size of the remainder is ε^N. This suffices to ensure that the diffusion is slower and slower with increasing N. Having fixed the time T, we determine N as a function of T and ε, and compute all the quantities entering the statement of the theorem.

For what concerns the proof of theorem 2, we ask the domain to contain at least a ball of radius ε like in the local version of the Nekhoroshev theorem. This gives an upper bound to N and, in turn, to T.

Proof of theorem 1. We prove the theorem iterating the results of lemma 3. We construct sequences of constants $\varepsilon^{(s)}, \Lambda^{(s)}, \vartheta^{(s)}$, $\Theta^{(s)}, \sigma^{(s)}, \varrho^{(s)}$, $D_V^{(s)}, W^{(s)}$ for $s \geq 1$, to be used in place of the primed constants in lemma 3. A straightforward use of lemma 3 shows that the constants are determined via the recursive system

$$\varepsilon^{(s)} = (\varepsilon^{(s-1)})^{5/4} C_H \ , \tag{12}$$

$$\Lambda^{(s)} = \Lambda^{(s-1)} + C_\Lambda (\varepsilon^{(s-1)})^{1/2} \ , \tag{13}$$

$$\vartheta^{(s)} = \vartheta^{(s-1)} - C_\vartheta (\varepsilon^{(s-1)})^{1/8} \ , \tag{14}$$

$$\Theta^{(s)} = \Theta^{(s-1)} + C_\vartheta (\varepsilon^{(s-1)})^{1/8} \ , \tag{15}$$

$$\sigma^{(s)} = \sigma^{(s-1)} (1 - \delta_F^{(s)}) - C_\sigma (\varepsilon^{(s-1)})^{1/4} \ , \tag{16}$$

$$\varrho^{(s)} = \varrho^{(s-1)} - C_\Delta (\varepsilon^{(s-1)})^{1/8} \ , \tag{17}$$

$$D_V^{(s)} = D_V^{(s-1)} + 2^n \left(\frac{\Theta^{(s)}}{\vartheta^{(s)}} \right)^n C_{\mathcal{G}} \left(\frac{B - \ln \varepsilon^{(s-1)}}{\sigma^{(s-1)} \delta_F^{(s)}} \right)^{n-1} \ , \tag{18}$$

$$W^{(s)} = W^{(s-1)} + \left[D_V^{(s-1)} C_\Delta + A \frac{(B - \ln \varepsilon^{(s-1)})^{n-1}}{(\delta_F^{(s)})^{n-1}} \right] \frac{(\varepsilon^{(s-1)})^{1/4}}{(\delta_F^{(s)})^{n+2}} \operatorname{Vol} \mathcal{D} \ , \tag{19}$$

starting with $\varepsilon^{(0)} = \varepsilon, \ldots, D_V^{(0)} = D_V$, and $W^{(0)} = 0$. Moreover, denoting by $p^{(s)}, q^{(s)}$ the canonical coordinates at step s, from lemma 3 we have

$$|p^{(s)} - p^{(s-1)}| \leq C_\Delta (\varepsilon^{(s-1)})^{1/2} \ , \quad |q^{(s)} - q^{(s-1)}| \leq C_\sigma (\varepsilon^{(s-1)})^{1/4} \ . \tag{20}$$

Here we have to remark that the constants $C_H, C_\Lambda, C_\vartheta, C_\sigma, C_\Delta, A, B$ depend on $\sigma^{(s-1)}, \vartheta^{(s-1)}, \Theta^{(s-1)}, \Lambda^{(s-1)}$. We assume the condition (to be checked later) that, for ε small enough one has

$$\sigma^{(s)} > \sigma/2 \ , \ \vartheta^{(s)} > \vartheta/2 \ , \ \Theta^{(s)} < 2\Theta \ , \ \Lambda^{(s)} < 2\Lambda \quad \forall s \ . \tag{21}$$

Then we can define new constants $C_H^*, C_V^*, C_\vartheta^*, C_\sigma^*, C_\Delta^*, A^*, B^*$ as the maximum of $C_H, C_\Lambda, C_\vartheta, C_\sigma, C_\Delta, A, B$ over the range of allowed variation of $\sigma^{(s)}, \vartheta^{(s)}, \Theta^{(s)}, \Lambda^{(s)}$.

We look now for a closed non recursive expression for $\varepsilon^{(s)}$. This readily follows from the assumption (21) and from the condition on ε

$$\varepsilon < \left(\frac{1}{eC_H^*}\right)^4 \ . \tag{22}$$

Indeed, we remark that equation (12) reads

$$\varepsilon^{(s)} \le (\varepsilon^{(s-1)})^{5/4} C_H^* \ ,$$

and that defining

$$\varepsilon_* = \varepsilon^{1/4} C_H^* \tag{23}$$

we get

$$\varepsilon^{(s)} \le \varepsilon \varepsilon_*^s \ .$$

Now we show that the uniform estimates (21) hold. Assuming that ε satisfies

$$C_V^* \frac{\varepsilon^{1/2}}{1-\varepsilon_*^{1/2}} < \Lambda \ , \quad C_\vartheta^* \frac{\varepsilon^{1/8}}{1-\varepsilon_*^{1/8}} < \vartheta/2 \ , \tag{24}$$

and using equations (13), (14) and (15) we get

$$\begin{aligned}
\Lambda^{(s)} &\le \Lambda + C_V^* \varepsilon^{1/2} \sum_{0\le l\le s} \varepsilon_*^{l/2} \le \Lambda + C_V^* \varepsilon^{1/2} \frac{1}{1-\varepsilon_*^{1/2}} < 2\Lambda \ , \\
\vartheta^{(s)} &\ge \vartheta - C_\vartheta^* \varepsilon^{1/8} \sum_{0\le l\le s} \varepsilon_*^{l/8} \ge \vartheta - C_\vartheta^* \varepsilon^{1/8} \frac{1}{1-\varepsilon_*^{1/8}} > \vartheta/2 \ , \\
\Theta^{(s)} &\le \Theta + C_\vartheta^* \varepsilon^{1/8} \sum_{0\le l\le s} \varepsilon_*^{l/8} \le \Theta + C_\vartheta^* \varepsilon^{1/8} \frac{1}{1-\varepsilon_*^{1/8}} < 2\Theta \ .
\end{aligned}$$

For what concerns $\sigma^{(s)}$, one has to discuss the equation (16). First of all we choose $\delta_F^{(s)}$ such that

$$\delta_F^{(s)} \ge (\varepsilon \varepsilon_*^s)^{1/16(n+1)} \ , \quad \sum_s \delta_F^s < 1/4 \ . \tag{25}$$

This is a non empty condition provided that

$$\varepsilon^{1/16(n+1)} \sum_s \left(\varepsilon_*^{1/16(n+1)}\right)^s \le \frac{\varepsilon^{1/16(n+1)}}{1-\varepsilon_*^{1/16(n+1)}} < 1/4 \ , \tag{26}$$

which gives a further condition on ε. Then we estimate

$$\sigma^{(s)} \ge \sigma \left(1 - \sum_{0\le l\le s} \delta_F^{(s)}\right) - C_\sigma^* \varepsilon^{1/4} \sum_s \varepsilon_*^{s/4} \ ;$$

this is larger than $\sigma/2$ provided the further condition on ε

$$C_\sigma^* \varepsilon^{1/4} \sum_s \varepsilon_*^{s/4} \leq C_\sigma^* \frac{\varepsilon^{1/4}}{1-\varepsilon_*^{1/4}} < \sigma/4 \tag{27}$$

is satisfied.

For what concerns $\varrho^{(s)}$ in (17) one has just to prove $\varrho^{(s)} > 0, \forall s$. This is satisfied for any ε such that

$$C_\Delta^* \varepsilon^{1/8} \sum_s \varepsilon^{s/8} \leq C_\Delta^* \frac{\varepsilon^{1/8}}{1-\varepsilon_*^{1/8}} < \varrho \ . \tag{28}$$

We come now to the estimate of $D_V^{(s)}$ and $W^{(s)}$ in (18) and (19). Here one needs estimates explicitly dependent on s. For $D_V^{(s)}$ we have:

$$\begin{aligned} D_V^{(s)} &\leq D_V + \sum_{0\leq l\leq s-1} \left(\frac{\Theta}{\vartheta}\right)^n 2^{4n} C_{\mathcal{G}} \left(\frac{B-\ln\varepsilon - l\ln\varepsilon_*}{\sigma\delta_F^{(l)}}\right)^{n-1} \\ &\leq D_V + \left(\frac{\Theta}{\vartheta}\right)^n 2^{4n} C_{\mathcal{G}} s \left(\frac{B-\ln\varepsilon - s\ln\varepsilon_*}{\sigma\delta_F^{(s)}}\right)^{n-1} \ . \end{aligned} \tag{29}$$

Concerning $W^{(s)}$, by (19), (25) and (29), we get:

$$W^{(s)} \leq \left\{ \sum_{0\leq l\leq s-1} \left[D_V C_\Delta^* + \beta l (B^* - \ln\varepsilon - l\ln\varepsilon_*)^{n-1}\right] (\varepsilon\varepsilon_*^s)^{1/8} \right\} \mathrm{Vol}\,\mathcal{D} \tag{30}$$

where $\beta = 2\max\Bigl(\left(\frac{\Theta}{\vartheta}\right)^n 2^{4n} C_{\mathcal{G}} C_\Delta^*/\sigma^{n-1}, A^*\Bigr)$. We now use

$$B^* - \ln\varepsilon - l\ln\varepsilon_* = (l+4+B^*+4\ln C_H^*)|\ln\varepsilon_*| \equiv (l+D^*)|\ln\varepsilon_*| \ ; \tag{31}$$

with this we can estimate

$$\sum_{0\leq l\leq s-1} l(B^* - \ln\varepsilon - l\ln\varepsilon_*)^{n-1}\varepsilon_*^{l/8} \leq |\ln\varepsilon_*|^n \sum_{0\leq l\leq s-1} (l+D^*)^n \varepsilon_*^{l/16}\varepsilon_*^{l/16} \ . \tag{32}$$

Using now the inequality

$$(\alpha+s)^n x^{s/2} \leq \alpha^n \quad \text{for } 0 < x \leq e^{-2n/\alpha} \ ,$$

and imposing the new condition on ε

$$\varepsilon_*^{1/16} \leq e^{-2n/D^*} \ , \tag{33}$$

we get from (32) and (30)

$$W^{(s)} \leq \left[D_V C^*_\Delta \frac{1-\varepsilon_*^{s/8}}{1-\varepsilon_*^{1/8}} + \beta (D^*)^n |\ln \varepsilon_*|^n \frac{1-\varepsilon_*^{s/16}}{1-\varepsilon_*^{1/16}} \right] \varepsilon^{1/8} \operatorname{Vol} \mathcal{D}$$
$$\leq \tilde{V} |\ln \varepsilon_*|^n \frac{1-\varepsilon_*^{s/8}}{1-\varepsilon_*^{1/16}} \varepsilon^{1/8} \operatorname{Vol} \mathcal{D} \ ,$$

with $\tilde{V} = 2 \max(D_V C^*_\Delta, \beta (D^*)^n)$. We now impose further conditions on ε

(34) $$\varepsilon_*^{1/16} \leq 1/2 \ , \quad \varepsilon^{1/8} \left| \frac{1}{4} \ln \varepsilon + \ln C^*_H \right|^n < \varepsilon^{1/16}$$

and define $V = 2\tilde{V}$ thus getting the final formula

(35) $$W^{(s)} \leq V \varepsilon^{1/16} (1 - \varepsilon_*^{s/8}) \operatorname{Vol} \mathcal{D} \quad \forall s \ .$$

Clearly, one has $W^{(s)} \leq \operatorname{Vol} \mathcal{D}$ for

(36) $$\varepsilon \leq \left(\frac{1}{V} \right)^{16} \ .$$

Equation (20) gives, for all s:

$$|p^{(s)} - p| \leq C^*_\Delta \frac{\varepsilon^{1/2}}{1-\varepsilon_*^{1/2}} \leq 2 C^*_\Delta \varepsilon^{1/2} \ , \quad |q^{(s)} - q| \leq C^*_\sigma \frac{\varepsilon^{1/4}}{1-\varepsilon_*^{1/4}} \leq 2 C^*_\sigma \varepsilon^{1/4} \ ,$$

namely the statement 1, identifying P and Q with any of the $p^{(s)}, q^{(s)}$.

The last step of this proof is to connect the time of diffusion T to the order of iteration s. Since

(37) $$\left\| H_1^{(s)} \right\|_{\mathcal{D}^{(s)}} \leq \varepsilon^{(s)} \leq \varepsilon \varepsilon_*^s \ ,$$

one has that the diffusion from $\mathcal{D}^{(s)} - \varepsilon \varepsilon_*^{s\,1/4}$ to $\mathcal{D}^{(s)}$ can not occur in a time $|t| \leq T^{(s)}$ where

(38) $$T^{(s)} = \frac{2}{\sigma} \frac{1}{(\varepsilon \varepsilon_*^s)^{3/4}} \ ;$$

here, the factor $2/\sigma$ comes from Cauchy estimates on the angle domain. So we get

$$\varepsilon_*^s = \left(\frac{\sigma}{2T} \right)^{4/3} \frac{1}{\varepsilon} \ ,$$

which means $s = 4/3(\ln\sigma - ln2 - \ln T - 3/4\ln\varepsilon)/\ln\varepsilon_*$. If one defines now $T = T'T_*$ with T_* such that $\ln\sigma - ln2 - 3/4\ln\varepsilon = \ln T_*$, then one has

$$s = -\frac{4}{3\ln\varepsilon_*}\ln\left(\frac{T}{T_*}\right) .$$

In view of (37), we get the statements 2 and 3 with the function η defined as

$$\eta(\varepsilon, T) = (\varepsilon\varepsilon_*^s)^{1/4} = \varepsilon^{1/4}\left(\frac{T_*}{T}\right)^{1/3} ,$$

as stated in the theorem. As a last remark, one could compute an explicit value for $\overline{\varepsilon}$ by collecting all the conditions (22), (24), (26), (27), (28), (33), (34), (36). *Q.E.D.*

Proof of theorem 2. We have just to modify the proof of theorem 1 above. We follow that proof up to formula (29), and at that point we impose that the domain $\mathcal{D}^{(s)}$ contains balls of radius ε. For this purpose, we define ΔV as

$$\Delta V = (1 - V\varepsilon^{1/16})\,\mathrm{Vol}\,\mathcal{D} ;$$

then the condition that $\mathcal{D}^{(s)}$ contains balls of radius ε translates into the inequality

$$D_V^{(s)}\varepsilon \le \Delta V ; \tag{39}$$

indeed, this means that a further restriction of ε can be made without reducing the volume to zero. We choose $\delta_F^{(s)} = 1/(8s^2)$. This definition of $\delta_F^{(s)}$ satisfies $\sum_s \delta_F^{(s)} \le 1/4$; moreover the condition $\delta_F^{(s)} \ge (\varepsilon\varepsilon_*^s)^{1/16(n+1)}$ is clearly satisfied for ε small enough, for example with the implicit condition

$$\varepsilon_* \le \frac{1}{e^{32(n+1)}} . \tag{40}$$

Replacing this in (29), the inequality (39) then gives

$$\left[D_V + \left(\frac{\Theta}{\vartheta}\right)^n \frac{2^{7n}C_{\mathcal{G}}}{\sigma^{n-1}} s^{2n-1}(B^* - \ln\varepsilon - s\ln\varepsilon_*)^{n-1}\right]\varepsilon \le \frac{\Delta V}{\mathrm{Vol}\,\mathcal{D}} . \tag{41}$$

We now use (31) and impose the further condition on ε

$$\Delta V' \equiv \frac{\Delta V}{\mathrm{Vol}\,\mathcal{D}} - D_V\varepsilon > 0 ; \tag{42}$$

then, from (41) we get

$$s \le s_* := \left(\frac{\alpha}{\varepsilon}\right)^{1/3n-2}\frac{1}{|\ln\varepsilon_*|^{(n-1)/(3n-2)}} - D^*$$

with

$$\alpha = \frac{\Delta V' \sigma^{n-1}}{(\Theta/\vartheta)^n 2^{7n} C_{\mathcal{G}}} \ .$$

Use now $s = s_* - 1$ in formula (38), with a last condition on ε, namely

$$|\ln \varepsilon_*|(D^* + 1) \leq \frac{1}{2} \left(\frac{\alpha}{\varepsilon}\right)^{1/(3n-2)} \ ; \tag{43}$$

one gets

$$T \geq \frac{2}{\sigma \varepsilon^{3/4}} \exp\left[\frac{3}{8}\left(\frac{\alpha}{\varepsilon}\right)^{1/(3n-2)}\right] \ ,$$

as claimed. The condition (43) has to be considered together with (40) and (42) and with $\varepsilon < \overline{\varepsilon}$ of theorem 1 for the definition of the upper limit $\overline{\varepsilon}_1$ for ε). This concludes the proof of theorem 2 is completed. *Q.E.D.*

6. Proofs of the main lemmas

Proof of lemma 1. We first perform a linear canonical transformation $(p, q) \to (I, \varphi)$

$$\varphi = Uq \ , \quad I = (U^T)^{-1} p \ ,$$

where U is a unimodular matrix and the upper index T denotes transposition. The matrix U is so defined that $\varphi_1 = \tilde{k} \cdot q$, where $\tilde{k}$ is defined by $k = \lambda \tilde{k}$, with λ the greatest common divisor among the components $(k_1, \ldots, k_n)$ of k. It is well known that such a matrix can be found. With the unimodular transformation U the Hamiltonian takes the form

$$\tilde{H}(I, \varphi) = \tilde{H}_0(I) + \tilde{h}_1(I) \cos(\lambda \varphi_1) \ ,$$

and using hypothesis (d) and the results of Lemma 4 reported in section 7 it is straightforward to get the inequality

$$\left\| \frac{\partial \tilde{h}_1}{\partial I} \right\|_{\tilde{\mathcal{D}}} \leq \frac{|k|}{\lambda} M \leq |k| M \ ,$$

where $\tilde{\mathcal{D}}$ is the transformed of the domain $\mathcal{D}$ under the canonical transformation. We now introduce new action angle variables J, ψ for $\tilde{H}$ in order to remove the dependence on the angles. This means that the new Hamiltonian $\tilde{H}'(J, \psi) = \tilde{H}(I(J, \psi), \varphi_1(J, \psi))$ does not depend on ψ, i.e., $\tilde{H}' = \tilde{H}'(J)$. The Arnold Liouville theorem can be easily applied to the Hamiltonian $\tilde{H}(I, \varphi_1)$ leading directly to the definition of the action variables, which are

$$(44)\qquad \begin{aligned} J_1 =& \frac{1}{2\pi}\oint_\Gamma I_1(I_2,\ldots,I_n,\varphi_1,E)\mathrm{d}\varphi_1\ , \\ J_l =& I_l \qquad (l=2,\ldots,n)\ , \end{aligned}$$

where $I_1(I_2,\ldots,I_n,\varphi_1,E)$ is implicitly defined by the equation

$$(45)\qquad \tilde H(I_1,I_2,\ldots,I_n,\varphi_1)=E$$

and Γ is an arbitrary closed curve on the torus $(I_1(\varphi_1,E),\varphi_1)$ obtained by continuously deforming in the complex space the real cycle $[0,2\pi]$. The existence of the solution of the implicit equation (45) is ensured by the hypotheses (a) and (d), by the condition $M<\gamma/(|k|\exp(|k|\sigma))$, and by the fact that

$$(46)\qquad \inf_{\tilde{\mathcal D}}\left|\frac{\partial\tilde H}{\partial I_1}\right|\ge\gamma-|k|Me^{|k|\sigma}>0\ .$$

We now write the solution of (45) in the form $I_1(\varphi_1,E)=I_1^{(0)}+w(\varphi_1,E)$ where $I_1^{(0)},\varphi_1^{(0)}$ is the starting point of the curve Γ and $w(\varphi_1,E)$ is a function to be determined. However, for our purposes it is enough to find an upper bound for w. To this end, we differentiate (45) with respect to φ_1 and get

$$\frac{\partial\tilde H_0}{\partial I_1}(I_1^{(0)}+w)\frac{\partial w}{\partial\varphi_1}+\frac{\partial\tilde h_1}{\partial I_1}(I_1^{(0)}+w)\frac{\partial w}{\partial\varphi_1}\cos\lambda\varphi_1-\lambda\tilde h_1(I_1^{(0)}+w)\sin\lambda\varphi_1=0$$

which gives

$$(47)\qquad \frac{\partial w}{\partial\varphi_1}=\lambda\tilde h_1\sin\lambda\varphi_1\left(\frac{\partial\tilde H_0}{\partial I_1}+\frac{\partial\tilde h_1}{\partial I_1}\cos\lambda\varphi_1\right)^{-1}\ .$$

In view of (46) and the hypothesis (d) we get directly

$$\left\|\frac{\partial w}{\partial\varphi_1}\right\|_{\tilde{\mathcal D}}\le\frac{\lambda Ce^{|k|\sigma}}{\gamma-|k|Me^{|k|\sigma}}\ ,$$

so that we also get

$$|w(\varphi_1,E)|\le 2(\pi+|k|\sigma)\frac{\lambda Ce^{|k|\sigma}}{\gamma-|k|Me^{|k|\sigma}}$$

which holds for any $\varphi_1\in\mathbf{T}^n_{|k|\sigma/\lambda}$ and for any value $E\in\mathbf{C}$. Therefore, from (44) one has

$$(48)\qquad \left\|J_1-I_1^{(0)}\right\|_{\mathcal D}\le\frac{1}{2\pi}\oint_\Gamma|w|\mathrm{d}\varphi_1\le 2(\pi+|k|\sigma)\frac{\lambda Ce^{|k|\sigma}}{\gamma-|k|Me^{|k|\sigma}}\ .$$

Still using the unimodular matrix U we now introduce the new actions R by $R = U^T J$. From (48) and using also lemma 4 one readily gets the estimate

$$(49)\qquad \|p - R\|_{\mathcal{D}} \le 2(\pi + |k|\sigma)\frac{|k|Ce^{|k|\sigma}}{\gamma - |k|Me^{|k|\sigma}} \equiv \delta_0\ ,$$

namely the first estimate in the statement (2).

We now build up the new angles r conjugated to R and look for an estimate of $|r - q|$. To this end, we first remark that, by Arnold's theorem, there exist angles $\psi_1, \ldots, \psi_n$ conjugated to the actions $J_1, \ldots, J_n$ above. The angles ψ have a constant frequency as $\tilde{H}'$ depends only on J and moreover one has

$$\varphi_l = \psi_l + d_l(J, \psi_1) \quad l = 1, \ldots, n$$

with functions d_l satisfying $d_l(J, \psi_1) = d_l(J, \psi_1 + 2\pi)$. We now complete the unimodular canonical transformation by defining $r = U^{-1}\psi$. Also the new angles r turn out to have constant frequencies, and, moreover, the transformation from q to r is readily seen to be of the form

$$q_l = r_l + g_l(R, k \cdot r)\ .$$

The aim now is to estimate $|g_l|$. By the trivial identity $g_l = r_l + \int \dot{g}_l\,\mathrm{d}t - \int \dot{r}_l\,\mathrm{d}t$, and using Hamilton equations one gets

$$g_l = r_l + \int_0^t \frac{\partial H}{\partial p_l}\mathrm{d}\tau - \int_0^t \frac{\partial H'}{\partial R_l}\mathrm{d}\tau\ ;$$

substituting now $\mathrm{d}\tau = (\partial\tilde{H}'/\partial J_1)^{-1}\mathrm{d}\psi_1$, and using lemma 5 we get the identity

$$q_l = r_l + \left(\frac{\partial\tilde{H}'}{\partial J_1}\right)^{-1}\int_{\Gamma(0,\psi_1)}\frac{\partial H}{\partial p_l}\mathrm{d}\psi_1 - \frac{\psi_1}{2\pi}\left(\frac{\partial\tilde{H}'}{\partial J_1}\right)^{-1}\oint_{\Gamma(0,2\pi)}\frac{\partial H}{\partial p_l}\mathrm{d}\psi_1\ ,$$

where with $\Gamma(a, b)$ we denote any curve on the torus with extremes in $\psi_1 = a, \psi_1 = b$. From this identity we obtain with a little algebra

$$|q_l - r_l| \le 4(\pi + |k|\sigma)\left|\frac{\partial\tilde{H}'}{\partial J_1}\right|^{-1}|B_l|$$

where $B_l = \frac{\partial H}{\partial p_l}(p, q) - \frac{\partial H}{\partial p_l}(R, 0)$. From (49) and hypotheses (c) and (d) we get

$$|B_l| \le \sqrt{n}\Theta\delta_0 + Me^{|k|\sigma}\ .$$

Moreover by lemma 5, the hypotheses (a) and (d), and recalling the condition $M < \gamma/(|k|\exp|k|\sigma)$ one gets

$$\left|k \cdot \frac{\partial H'}{\partial R}\right| \ge \inf_{\mathcal{G}\times\mathbf{T}^n_\sigma}\left|k \cdot \frac{\partial H}{\partial p}\right| \ge \gamma - |k|Me^{|k|\sigma} > 0\ ,$$

so that, using the definition of δ_0 we finally prove the second estimate of statement 2. Statement 1 readily follows from statement 2.

We finally come to the proof of statement 3. From Lemma 5 and hypotheses (b) and (d) one has

$$\left|\frac{\partial H'}{\partial R}\right| \leq \left\|\frac{\partial H}{\partial p_l}\right\|_{\mathcal{D}} \leq \Lambda + Me^{|k|\sigma} ,$$

so that

$$\left\|\frac{\partial H'}{\partial R_l}\right\|_{\mathcal{D}-(\delta_0,\delta_1)} \leq \Lambda' \equiv \Lambda + Me^{|k|\sigma} ,$$

namely the first of (7). To prove the remaining inequalities we first remark that, by the canonical transformation, one has:

$$H'(R) = H_0(R+f(R,k\cdot r)) + h_1(R+f(R,k\cdot r))\cos(r+g(R,k\cdot r)) \equiv H_0(R)+F(R)$$

with a suitable function F. Using hypotheses (b) and (d) and the definition of δ_0, a straightforward estimate gives

$$\|F\|_{\mathcal{D}-(\delta_0,\delta_1)} \leq n\Lambda\delta_0 + Ce^{|k|\sigma} .$$

Consider now the restriction of the domain $\mathcal{D}-(\delta_0,\delta_1)$ to $\mathcal{D}-(\delta_0+\delta_2,\delta_1)$; by the Cauchy estimate one gets

$$\left\|\frac{\partial^2 F}{\partial R_i \partial R_j}\right\|_{\mathcal{D}-(\delta_0+\delta_2,\delta_1)} \leq 2\frac{n\Lambda\delta_0 + Ce^{|k|\sigma}}{\delta_2^2} . \tag{50}$$

Denoting now A' as the hessian matrix of H' with respect to R and $\tilde{A}$ as the hessian matrix of F, one has $A' = A + \tilde{A}$ with (from (50))

$$\left\|\tilde{A}v\right\| \leq 2\frac{n\Lambda\delta_0 + Ce^{|k|\sigma}}{\delta_2^2}\|v\| .$$

The last two inequalities of (7) are then straightforward, and this concludes the proof. *Q.E.D.*

Proof of lemma 2. The thesis is a direct consequence of the fact that

$$\begin{aligned} \overline{h}_1(p) &= \overline{h}_1(R+f(R,r)) = \overline{h}_1(R) + \Delta h_1(R,r) , \\ \cos(m\cdot q) &= \cos[m\cdot(r+g(R,r))] = \cos(m\cdot r) + \Delta c(R,r) , \end{aligned}$$

with

$$\|\Delta h_1\|_{\mathcal{D}-(\delta_0,\delta_1)} \leq nM_1\delta_0 , \quad \|\Delta c\|_{\mathcal{D}-(\delta_0,\delta_1)} \leq n|m|e^{|m|\sigma}\delta_1 .$$

Here, the estimates of $\|f\|$ and $\|g\|$ of lemma 1 have been used. *Q.E.D.*

Proof of lemma 3. Consider the Fourier expansion of the perturbation H_1, namely

$$H_1(p,q) = \sum_{k\in\mathbf{Z}^n} h_k(p)\cos(k\cdot q)$$

(as explained in sect. 4 we restrict for simplicity to the case of expansion with only the cos terms). By the known properties of analytic functions, and in view of hypotheses (a) and (b) one has

$$\|h_k\|_{\mathcal{D}} \le \varepsilon e^{-|k|\sigma} \ , \quad \left\|\frac{\partial h_k}{\partial p}\right\|_{\mathcal{D}} \le \varepsilon^{3/4} e^{-|k|\sigma} \ . \tag{51}$$

We now split H_1 as:

$$H_1(p,q) = H_1^{\le K_0}(p,q) + H_1^{>K_0}(p,q) \ ,$$

where $H_1^{\le K_0}$ and $H_1^{>K_0}$ contain the Fourier components with $|k|\le K_0$ and $|k|>K_0$ respectively. In order to choose K_0 we impose the condition

$$\left\|H_1^{>K_0}\right\|_{\mathcal{D}-(0,\sigma\delta_F)} \le \varepsilon^{5/4} \tag{52}$$

with a positive δ_F satisfying

$$\varepsilon^{1/16(n+1)} \le \delta_F < \frac{1}{4} \ ; \tag{53}$$

this gives a first condition on ε, namely

$$\varepsilon < 4^{-16(n+1)} \ . \tag{54}$$

We also remark that the estimate (52) holds in a domain $\mathcal{G}\times\mathbf{T}^n_{\sigma(1-\delta_F)}$. By (52) and the technical lemma 6 and also using (53), one sees that one can choose

$$K_0 = \frac{B-\ln\varepsilon}{\sigma\delta_F} - 1 \ ,$$

where $B = 2n(2\ln 2 - \ln\sigma) + (n+1)\sigma/4$. The number N of harmonics with $|k|\le K_0$ can be overestimated by

$$N \le 2^n (K_0+1)^{n-1} = 2^n\left(\frac{B-\ln\varepsilon}{\sigma\delta_F}\right)^{n-1} \ . \tag{55}$$

We now start the process of successive elimination of harmonics with $|k|\le K_0$. To this end, we apply N times lemma 1, starting from a domain $\mathcal{G}\times\mathbf{T}^n_{\sigma(1-\delta_F)}$ (that is, everywhere in the statement of lemma 1 we replace σ with $\sigma(1-\delta_F)$). Furthermore we associate to every k a positive γ_k to be used in hypothesis (a) of lemma 1 by choosing

$$\gamma_k = \varepsilon^{1/4} e^{-|k|\sigma\delta_F/4}\left(1+\varepsilon^{1/2}|k|e^{-3|k|\sigma\delta_F/4}\right) \ .$$

Next we remark that, in view of (51), the constants C and M in the hypothesis (d) of lemma 1 can be replaced by

$$C_k = \varepsilon e^{-|k|\sigma} \ , \quad M_k = \varepsilon^{3/4} e^{-|k|\sigma} \ . \tag{56}$$

It is straightforward to check that the condition

$$M_k < \frac{\gamma_k}{|k| e^{|k|\sigma(1-\delta_F)}} \ ,$$

required by lemma 1, is satisfied. Now, the iteration of lemma 1 requires sequences of constants $\Lambda^{(s)}, \vartheta^{(s)}$, $\Theta^{(s)}, \sigma^{(s)}$, $\delta_0^{(s)}, \delta_1^{(s)}$ and $\delta_2^{(s)}$ to be used in hypotheses (b), (c) and in statement 3 of lemma 1. We start with

$$\Lambda^{(1)} = \Lambda \ , \quad \vartheta^{(1)} = \vartheta \ , \quad \Theta^{(1)} = \Theta \ , \quad \sigma^{(1)} = \sigma(1-\delta_F) \ ,$$

and choose $\delta_2^{(s)}$ depending on k as

$$\delta_2^{(s)} = \varepsilon^{1/4} e^{-|k|\sigma\delta_F/4} \ . \tag{57}$$

(Here and in the following we recall that at the step s we eliminate a single harmonic k). Using now statements 2 and 3 and the definition of δ_0, δ_1 in lemma 1, and in view of (56) one finds the recursive formulas:

$$\delta_0^{(s)} = 2(\pi + |k|\sigma)|k| \varepsilon^{3/4} e^{-3|k|\sigma\delta_F/4} \ , \tag{58}$$

$$\delta_1^{(s)} = 4(\pi + |k|\sigma)[2\sqrt{n}\Theta^{(s-1)}|k|(\pi + |k|\sigma) + 1] \varepsilon^{1/2} e^{-|k|\sigma\delta_F/2} \ , \tag{59}$$

$$\Lambda^{(s)} = \Lambda^{(s-1)} + \varepsilon^{3/4} e^{-|k|\sigma\delta_F} \ , \tag{60}$$

$$\vartheta^{(s)} = \vartheta^{(s-1)} - 2\frac{n\Lambda^{(s-1)}\delta_0^{(s)} + \varepsilon e^{-|k|\sigma\delta_F}}{(\delta_2^{(s)})^2} \ , \tag{61}$$

$$\Theta^{(s)} = \Theta^{(s-1)} + 2\frac{n\Lambda^{(s-1)}\delta_0^{(s)} + \varepsilon e^{-|k|\sigma\delta_F}}{(\delta_2^{(s)})^2} \ , \tag{62}$$

$$\sigma^{(s)} = \sigma^{(s-1)} - \delta_1^{(s)} \ . \tag{63}$$

All the quantities $\Lambda', \vartheta', \ldots$ in the statement, are clearly to be identified with

$\Lambda^{(N)}, \vartheta^{(N)}, \ldots$. Considering (60) one readily gets, for $1 \le s \le N$

$$
\begin{aligned}
\Lambda^{(s)} &\le \Lambda + \varepsilon^{3/4} \sum_{k\in\mathbf{Z}^n} e^{-|k|\sigma\delta_F} \\
&\le \Lambda + \varepsilon^{3/4} 2^n \sum_{m\in\mathbf{N}} m^{n-1} e^{-m\sigma\delta_F} \\
&\le \Lambda + \frac{2^{n+1}(n-1)!}{(\sigma\delta_F)^n} \varepsilon^{3/4} \\
&\le \Lambda + \frac{2^{n+1}(n-1)!}{\sigma^n} \varepsilon^{1/2} \ .
\end{aligned}
\tag{64}
$$

Here we have used the fact that the number of vectors $k \in \mathbf{Z}^n$ satisfying $|k| = m$ does not exceed $2^n m^{n-1}$; moreover the technical inequality of lemma 7 and the condition $\delta_F > \varepsilon^{1/16(n+1)} > \varepsilon^{1/4}$ have been used. This gives

$$
\Lambda^{(s)} \le \Lambda + \frac{C_\Lambda}{2} \varepsilon^{1/2}
\tag{65}
$$

for $1 \le s \le N$, with

$$
C_\Lambda = \frac{2^{n+2}(n-1)!}{\sigma^n} \ .
$$

We impose now the condition on ε

$$
\varepsilon < \left(\frac{2\Lambda}{C_\Lambda} \right)^2 \ ,
\tag{66}
$$

which implies in particular $\Lambda^{(s)} \le 2\Lambda$ for $1 \le s \le N$. Making use of this inequality, and proceeding as in the estimate for $\Lambda^{(s)}$ above, we get

$$
\Theta^{(s)} \le \Theta + \frac{C_\vartheta}{2} \varepsilon^{1/8} \ , \quad \vartheta^{(s)} \ge \vartheta - \frac{C_\vartheta}{2} \varepsilon^{1/8}
\tag{67}
$$

for $1 \le s \le N$, with

$$
C_\vartheta = \frac{2^{3n+10}(n+1)!\,\Lambda e^{\pi/16}}{\sigma^{n+1}} \ .
$$

With the further condition on ε

$$
\varepsilon < \left(\frac{\vartheta}{C_\vartheta} \right)^8
\tag{68}
$$

one also gets $\vartheta^{(s)} \ge \vartheta/2,\ \Theta^{(s)} \le 2\Theta$ for $1 \le s \le N$. Using this in (59), we estimate

$$
\sum_{s=1}^{N} \delta_1^{(s)} < C_\sigma \varepsilon^{1/4}
$$

with

$$C_\sigma = \frac{2^{2n+9}\sqrt{n}(n+2)!\,\Theta e^{\pi/8}}{\sigma^{n+1}} \ .$$

This is given as follows: use

$$(\pi + |k|\sigma)[4\Theta\sqrt{n}|k|(\pi+|k|\sigma)+1] < 8\Theta\sqrt{n}\sigma^2\left(\frac{\pi}{\sigma}+|k|\right)^3 \ ;$$

then estimate

$$\sum_{k\in\mathbf{Z}^n}\left(\frac{\pi}{\sigma}+|k|\right)^3 e^{-|k|\sigma\delta_F/2} \le 2^n\sum_{m\in\mathbf{N}}\left(\frac{\pi}{\sigma}+m\right)^{n+2} e^{-m\sigma\delta_F/2} \ ,$$

and use the technical lemma 7 and the inequality (53) to estimate the sum. We now estimate $\sigma^{(s)}$ as given by (63). It is an easy matter to get

$$\sigma^{(s)} \ge \sigma(1-\delta_F) - C_\sigma\varepsilon^{1/4} \quad \text{for } 1\le s\le N$$

and, identifying σ' with $\sigma^{(N)}$, to get the first estimate of statement 3.

We now prove statement 4. To this end, we denote by $p^{(s)}, q^{(s)}$ the canonical variables introduced at step s, so that p', q' are to be identified with $p^{(N)}, q^{(N)}$. The total deformation $|p-p'|$ and $|q-q'|$ is then estimated by adding up all the contributions due to $\delta_0^{(s)}$ and $\delta_1^{(s)}$ respectively, and one gets

$$|p-p'| \le \sum_{s=1}^{N}\delta_0^{(s)} < C_\Delta\varepsilon^{1/2} \ , \quad |q-q'| \le \sum_{s=1}^{N}\delta_1^{(s)} < C_\sigma\varepsilon^{1/4}$$

with

$$C_\Delta = \frac{2^{3n+8}(n+1)!\,e^{\pi/16}}{\sigma^{n+1}} + 1 \ .$$

We now come to the conditions and the statements concerning the domains. We first check that the condition $\delta_1^{(s)} < \sigma^{(s-1)}/2$ of lemma 1 is satisfied; to this end we impose the further condition on ε

$$\varepsilon < \left(\frac{\sigma}{4C_\sigma}\right)^4 \ , \tag{69}$$

and, in view of (53) we get $\sigma^{(s)} > \sigma/2$ for $1\le s\le N$. This, in turn, implies the required condition.

The conditions on $\delta_0^{(s)}$ and $\delta_2^{(s)}$ of lemma 1 at each step trivially follow from statement 2 on the volumes that we are now going to prove. First, we ignore for a moment the resonances inside the domain $\mathcal{D}$, and consider only the loss of volume due to the restrictions at the border of the domain, as seen in lemma 1. At each step, the domain $\mathcal{D}^{(s-1)}$ is restricted to $\mathcal{D}^{(s-1)} - (\delta_0^{(s)}+\delta_2^{(s)}, \delta_1^{(s)})$. The loss of volume can be estimated by observing that the backward image of the domain $\mathcal{D} - (\sum_{l\le s} 2\delta_0^{(l)} + \delta_2^{(l)}, \sum_{l\le s} 2\delta_1^{(l)})$ via the sequence of canonical transformations which lead from p, q to $p^{(s)}, q^{(s)}$, is contained in $\mathcal{D}^{(s-1)} - (\delta_0^{(s)} + \delta_2^{(s)}, \delta_1^{(s)})$ for $1 \le s \le N$. This is readily checked by iterating the statement 1 of lemma 1. This in turn gives

$$\mathrm{Vol}(\mathcal{D} \setminus (\mathcal{D}^{(N)} - \varepsilon^{1/4})) \le \mathrm{Vol}\Big(\mathcal{D} \setminus \Big(\mathcal{D} - \sum_{l\le N}(2\delta_0^{(l)} + \delta_2^{(l)}) - \varepsilon^{1/4}\Big)\Big) ,$$

Here, recalling that at each step we remove a single harmonic and that $\delta_0^{(l)}, \delta_2^{(l)}$ depend on $|k|$ by (58) and (57), we can overestimate

$$\sum_{l\le N}(2\delta_0^{(l)} + \delta_2^{(l)}) + \varepsilon^{1/4} \le \Delta$$

where Δ is defined as the sum of the explicit expressions (58), (57) for k ranging over $\mathbf{Z}^n$; this readily gives, (still using lemma 7)

$$\Delta \le C_\Delta \varepsilon^{1/4} \delta_F^{-(n+2)} . \tag{70}$$

By imposing $\Delta < \varrho$, which is true provided the condition on ε

$$\varepsilon \le \left(\frac{\varrho}{C_\Delta}\right)^8 \tag{71}$$

holds, one gets from (e) that the loss of volume at the border of the domain is bounded by a quantity W_b satisfying

$$W_b \le D_V C_\Delta \varepsilon^{1/4} \delta_F^{-(n+2)} \,\mathrm{Vol}(\mathcal{D}) . \tag{72}$$

We come now to the estimate of the loss of volume due to the elimination of resonant strips. With the notations above for $p^{(s)}, q^{(s)}$, we consider the resonance corresponding to the harmonic k which is eliminated passing from $p^{(s-1)}, q^{(s-1)}$ to $p^{(s)}, q^{(s)}$. First we remove a strip around the resonant surface. The width of the strip is γ_k in the space of frequencies $\omega^{(s-1)}$ associated to $p^{(s-1)}$. Then the strip is enlarged, at every step, passing from $p^{(s-1)}, q^{(s-1)}$ to $p^{(N)}, q^{(N)}$. By the same considerations above, one sees that the projection of the image of this enlarged strip onto $\mathcal{G}^{(s-1)}$ is contained in the strip of width γ_k further enlarged by

$$\sum_{s\le l\le N} (2\delta_0^{(l)} + \delta_2^{(l)}) + \varepsilon^{1/4} \le \Delta .$$

Therefore we can estimate the loss of volume around a resonant surface as the volume of the original strip γ_k enlarged by Δ. From (d) one sees that the image of such a strip in the space of frequencies $\omega^{(s-1)}$ is contained in a larger strip of width

$$\tilde{\gamma}_k = \gamma_k + \sqrt{n}\Theta^{(s-1)}|k|\Delta \le \gamma_k + 2\sqrt{n}\Theta|k|\Delta .$$

Denoting by $\Omega^{(s-1)}$ the image of $\mathcal{G}^{(s-1)}$ in the frequency space, the volume of the latter strip is bounded by

$$W_{\Omega^{(s-1)}} \leq 2\sqrt{n}(\operatorname{diam}\Omega^{(s-1)})^{n-1}\frac{\tilde{\gamma}_k}{|k|} \ .$$

Remarking now that $\operatorname{diam}\Omega^{(s-1)} \leq \Theta^{(s-1)} \operatorname{diam}\mathcal{G}^{(s-1)} \leq \Theta^{(s-1)} \operatorname{diam}\mathcal{G}$, and denoting $C_{\mathcal{G}} = [2\sqrt{n}(\operatorname{diam}\mathcal{G})^{n-1}]/\operatorname{Vol}\mathcal{G}$ one has finally that the volume of this strip in the space $\mathcal{D}^{(s-1)}$ can be estimated by

$$W_k \leq \frac{(\Theta^{(s-1)})^{n-1}}{(\vartheta^{(s-1)})^n} C_{\mathcal{G}} \frac{\tilde{\gamma}_k}{|k|} \operatorname{Vol}\mathcal{D} \ .$$

The last step is to add up the volumes W_k over $|k| \leq K_0$. With easy computations, where we use $\gamma_k \leq 2\varepsilon^{1/4}, \vartheta^{(s-1)} \geq \vartheta', \Theta^{(s-1)} \leq \Theta'$ and (70), we get that the loss of volume due to the resonances is bounded by a quantity W_R defined as

$$W_R = 2^n \left(\frac{\Theta'}{\vartheta'}\right)^n C_{\mathcal{G}}(2\sqrt{n}C_\Delta + 2)\left(\frac{B - \ln\varepsilon}{\sigma\delta_F}\right)^{n-1} \frac{\varepsilon^{1/4}}{\delta_F^{n+2}} \operatorname{Vol}(\mathcal{D}) \ . \tag{73}$$

Defining now

$$A = 2^n \left(\frac{\Theta'}{\vartheta'}\right)^n C_{\mathcal{G}}(2\sqrt{n}C_\Delta + 2)/\sigma^{n-1} \ ,$$

and adding to W_R the value W_b given by (72) the statement 2 is proved. The proof of statement 1 just requires a condition on ε, namely

$$[D_V C_\Delta + A(B - \ln\varepsilon)^{n-1}]\varepsilon^{1/8} \leq 1 \ , \tag{74}$$

where we have used $\delta_F^{(n+2)} \geq \varepsilon^{1/8}$.

We complete now the proof of statement 3. With the same argument applied for the estimation of W_b and W_R, one easily sees that a further restriction of $\mathcal{D}' \equiv \mathcal{D}^{(N)}$ by any quantity $d < \varrho - C_\Delta\Delta \leq \varrho - C_\Delta\varepsilon^{1/8} \equiv \varrho'$ causes a loss of volume estimated by

$$\left[D_V + 2^n\left(\frac{\Theta'}{\vartheta'}\right)^n 2\sqrt{n}C_{\mathcal{G}}\left(\frac{B - \ln\varepsilon}{\sigma\delta_F}\right)^{n-1}\right] d \operatorname{Vol}\mathcal{D} \ ;$$

this gives the values of ϱ' and D'_V used in the statement.

We finally come to the estimates of the statement 5. After the elimination of the harmonics up to K_0, the Hamiltonian has the form

$$\tilde{H}_0(p') + \tilde{H}_1(p',q') + \mathcal{R}(p',q') \ ,$$

where $\tilde{H}_1 = H_1^{\geq K_0}$, and $\mathcal{R}(p',q')$ is generated by the sequence of transformations. Thus, denoting by $\overline{\mathcal{R}}(p')$ the average of $\mathcal{R}(p',q')$ over the angles, we define

$$H'_0(p') = \tilde{H}_0(p') + \overline{\mathcal{R}}(p') \ , \quad H'_1(p',q') = \tilde{H}_1(p',q') + \mathcal{R}(p',q') - \overline{\mathcal{R}}(p') \ .$$

We use now the result of lemma 2. In view of (51) the constants C_1 and M_1 in the hypotheses (a) and (b) of lemma 2 assume now the values

$$C_1 = \varepsilon e^{-|m|\sigma} \ , \quad \varepsilon^{3/4} e^{-|m|\sigma} \ .$$

By substitution of (58), and (60) and using $\Theta^{(s)} \leq 2\Theta$, after a little algebra the result of lemma 2 reads

$$\|\mathcal{R}_{k,m}\|_{\mathcal{D}^{(s)}} \leq 88 n^{5/2} \varepsilon^{3/2} e^{-|m|\sigma\delta_F} e^{-|k|\sigma\delta_F/2} |m| \Theta\sigma^3 \left(|k| + \frac{\pi}{\sigma} + \frac{1}{4\sqrt{n}\Theta\sigma} \right)^5 \ . \tag{75}$$

As the canonical transformations do not change the supremum norms adopted here, the norm of H_1' can be estimated by simply adding to the norm of $H_1^{>K_0}$ the norms of each term $\mathcal{R}_{k,m}$ for any k such that $0 < |k| \leq K_0$ and for any m with $|m| \leq K_0$ related to an harmonic not yet eliminated. A rough estimation of $\|H_1'\|$ is then given by $\|\mathcal{R}\| + \varepsilon^{5/4}$, with

$$\|\mathcal{R}\|_{\mathcal{D}'} \leq \sum_{k\in\mathbf{Z}^n} \sum_{m\in\mathbf{Z}^n} \|\mathcal{R}_{k,m}\|_{\mathcal{D}'} \ .$$

By using twice the technical inequality of lemma 7 and $\delta_F^{2n+6} > \varepsilon^{1/4}$ one easily gets

$$\|H_1'\|_{\mathcal{D}'} \leq C_H \varepsilon \cdot \varepsilon^{1/4}$$

with

$$C_H = \frac{2^{3n+7} A_1 (n+3)!\,(n+1)\,! e^{\sigma(A_2+1/2)/4}}{\sigma^{2n+6}} + 1 \ ,$$

$$A_1 = 88 n^{5/2} \Theta\sigma^3 \ , \quad A_2 = \frac{\pi}{\sigma} + \frac{1}{4\sqrt{n}\Theta\sigma} \ .$$

We now impose the last condition on ε, i.e.,

$$\varepsilon < \left(\frac{1}{2C_H} \right)^4 \ ; \tag{76}$$

by which $\varepsilon^{1/4} C_H < 1/2$; by defining $\varepsilon' = \varepsilon\varepsilon^{1/4} C_H < \varepsilon$ one gets the first inequality of statement (5). The second inequality follows directly from the Cauchy estimates on the derivatives of H_1' in the restricted domain $\mathcal{D}' - \varepsilon^{1/4}$. The third inequality follows from (65), with the remark that $\left\|\overline{\mathcal{R}}\right\|_{\mathcal{D}'} \leq \|\mathcal{R}\|_{\mathcal{D}'} \leq \varepsilon'$, and using the Cauchy inequality to get

$$\left\| \frac{\partial \overline{\mathcal{R}}}{\partial p'} \right\|_{\mathcal{D}' - \varepsilon^{1/4}} \leq \varepsilon'^{3/4} < \varepsilon^{1/2} \ .$$

A similar computation, using (67), gives the estimate for ϑ, Θ'.

Thus, the lemma is true for every $\varepsilon < \overline{\varepsilon}$, where $\overline{\varepsilon}$ could be explicitly determined using all the conditions (54), (66), (68), (69), (71), (74), (76). *Q.E.D.*

7. Technical lemmas

Lemma 4: *Let $k \in \mathbf{Z}^n$ with $\gcd(k_1, \ldots, k_n) = 1$ and with $k_i \neq 0$ for $i = 1, \ldots, n$. Then there exist a unimodular matrix $U = \{u_{i,j}\}$ such that*

i) $u_{1,l} = k_l$

ii) $|u_{j,l}| \leq |k_l|$ *for* $1 \leq j \leq n$, $\quad 1 \leq l \leq n$.

Proof. We prove this lemma by construction. We build U as

$$U = \prod_{i=1,n-1} U_i$$

with U_i such that, denoting by $u^i_{j,l}$ its entries one has $u^i_{i,i} = k_i/\lambda_{i-1}$, $u^i_{i,i+1} = \lambda_i/\lambda_{i-1}$, $u^i_{i+1,i} = a_i$, $u^i_{i+1,i+1} = b_i$, $u^i_{j,j} = 1$ for $j \neq i, i+1$, and 0 wherever else. The coefficients λ are such that $\lambda_0 = 1, \lambda_{n-1} = k_n, \lambda_{i-1} = \text{g.c.d.}(k_i, \lambda_i)$, while a_i and b_i are integers such that $\det U_i = 1$. Evidently, from these settings, U is the unimodular matrix we look for in order to satisfy statement i).

We prove now that one can always choose a_i, b_i in such a way that

$$|a_i| \leq |k_i/\lambda_{i-1}| \ , \quad |b_i| \leq |\lambda_i/\lambda_{i-1}| \ .$$

This is a direct consequence of the fact that k_i/λ_{i-1} and λ_i/λ_{i-1} are prime integers. Indeed, we consider two relative prime integers j_1, j_2 and we look for two integers m_1, m_2 such that

$$m_1 j_2 + 1 = m_2 j_1 \ . \tag{77}$$

The set of integers

$$\{(j_2+1)\mathrm{mod} j_1, \ldots, (j_1 j_2 + 1)\mathrm{mod} j_1\}$$

is composed of j_1 integers the values of which belong to the interval $[0, j_1 - 1]$. These integers are all different as j_1, j_2 are relatively prime. Hence, one of them must be equal to 0, which means that there exist two integers m_1, m_2 such that (77) holds. This implies also that one can always choose m_1 such that $|m_1| \leq |j_1|$. Considering now (77) one has

$$m_2 = \frac{m_1 j_2}{j_1} + \frac{1}{j_1} \ .$$

If $|j_1| \neq 1$ this is enough to prove that $|m_2| \leq |j_2|$. On the contrary, if $|j_1| = 1$ one can always choose $m_1 = \text{sign}(j_2), m_2 = -(|j_2| - 1)$, so that, again, $|m_1| \leq |j_1|, |m_2| \leq |j_2|$.

It is now a matter of a little algebra to write the product of the so defined matrices U_i and verify that U also satisfies statement ii). *Q.E.D.*

Lemma 5: *For the Hamiltonian (5) and with the canonical transformation $\mathcal{C}$ of lemma 1, one has*

$$\frac{\partial H'}{\partial R_l} = \frac{1}{2\pi}\oint_\Gamma \frac{\partial H}{\partial p_l}\mathrm{d}\psi_1 \ ,$$

with $\psi_1 = k \cdot r$ and Γ a cycle on the complex torus parameterized by the angle ψ_1.

Proof. Take $q_l = r_l + g_l(R, \psi_1)$ and differentiate it with respect to time:

$$\dot{q}_l = \dot{r}_l + \frac{\partial g_l}{\partial \psi_1}\dot{\psi}_1 \ .$$

Using now Hamilton equations we have

$$\frac{\partial H}{\partial p_l} = \frac{\partial H'}{\partial R_l} + \left(k \cdot \frac{\partial H'}{\partial R}\right)\frac{\partial g_l}{\partial \psi_1} \ .$$

Integrate now this equation over the cycle Γ on the torus, and get

$$\frac{1}{2\pi}\oint_\Gamma \frac{\partial H}{\partial p_l}\mathrm{d}\psi_1 = \frac{\partial H'}{\partial R_l}\frac{1}{2\pi}\oint_\Gamma \mathrm{d}\psi_1 + \left(k \cdot \frac{\partial H'}{\partial R}\right)\frac{1}{2\pi}\oint_\Gamma \frac{\partial g_l}{\partial \psi_1}\mathrm{d}\psi_1 \ .$$

Recalling now that g is periodic in ψ_1 the last term does vanish and the statement follows. *Q.E.D.*

Lemma 6: *Let*

$$f = \sum_{k \in \mathbf{Z}^n} f_k e^{ik\cdot q}$$

be such that the inequality

$$|f_k| \le \varepsilon e^{-|k|\sigma}$$

is satisfied for some positive ε and σ. Denote

$$f^{\ge K} = \sum_{|k|>K} f_k e^{ik\cdot q} \ .$$

Then, for any positive $\delta < 1$, for any q satisfying $|\operatorname{Im} q| < \sigma(1-\delta)$ and any positive integer K one has

$$\left|f^{\ge K}(q)\right| \le \varepsilon \frac{2^{2n} e^{n\sigma\delta/2}}{(\sigma\delta)^n} e^{-K\sigma\delta/2} \ .$$

Proof. We use the estimate

$$\begin{aligned}\sum_{|k|\geq K} \varepsilon e^{-|k|\sigma\delta} &\leq \varepsilon e^{-K\sigma\delta/2} \sum_{|k|\geq 0} e^{-|k|\sigma\delta/2} \\ &\leq \varepsilon e^{-K\sigma\delta/2} \left(\sum_{j\in\mathbf{Z}} e^{-|j|\sigma\delta/2} \right)^n \\ &= \varepsilon e^{-K\sigma\delta/2} \left(\frac{1+e^{-\sigma\delta/2}}{1-e^{-\sigma\delta/2}} \right)^n \ .\end{aligned}$$

By using now the fact that $1-e^{-x} = e^{-x}(e^x-1) > xe^{-x}$ and $e^{-\sigma\delta/2} < 1$ one gets the thesis. *Q.E.D.*

Lemma 7: *For any $A \geq 0$ and $\alpha > 0$, and for any non negative integer p one has*

$$\sum_{s\geq 1} (s+A)^p e^{-\alpha s} \leq \frac{p!\, e^{\alpha A}}{\alpha^{p+1}} \ .$$

Proof. Taking into account that the function $x^p e^{-\alpha x}$ has a maximum for $x = p/\alpha$, one has

$$\begin{aligned}\sum_{s\geq 1} (s+A)^p e^{-\alpha s} &= e^{\alpha A} \sum_{s\geq 1} (s+A)^p e^{-\alpha(s+A)} \\ &< e^{\alpha A} \left[\int_{A+1}^{p/\alpha+1} x^p e^{-\alpha x} \mathrm{d}x + \int_{p/\alpha-1}^{+\infty} x^p e^{-\alpha x} \mathrm{d}x \right] \\ &< 2e^{\alpha A} \int_0^{+\infty} x^p e^{-\alpha x} \mathrm{d}x \\ &= 2\frac{e^{\alpha A}}{\alpha^{p+1}} \int_0^{+\infty} \varphi^p e^{\varphi} \mathrm{d}\varphi \\ &= 2\frac{p! e^{\alpha A}}{\alpha^{p+1}} \ .\end{aligned}$$

Q.E.D.

References

Arnold, V.,I., (1963a): "Small denominators and problems of stability of motion in classical and celestial mechanics.", *Russ. Math. Surveys*, **18**, 85–191.

Arnold, V., I., (1963b): "On a theorem of Liouville concerning integrable problems of dynamics", *Sib. mathem. zh.*, **4**, 2.

Celletti, A., and Giorgilli, A., (1991): "On the stability of the Lagrangian points in the spatial restricted problem of the three bodies.", *Celest. Mech.*, **50**, 31–58.

Delaunay, C., (1867): "Théorie du mouvement de la Lune", *Mem. Ac. Sc.*, Paris, **29**.

Escande, D., (19xx): "??????????????????".
Giorgilli, A., Delshams, A., Fontich, E., Galgani, L. and Simó, C., (1989): "Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem.", *J. Diff. Eqs.*, **77**, 167–198.
Henrard, J., (1990): " A semi–numerical perturbation method for separable Hamiltonian systems.", *Celest. Mech.*, **49**, 43–67.
Henrard, J., (1991): "The adiabatic invariant in celestial mechanics.", to be published in *Dynamics reported*.
Lochak, P., (1989): "Stabilité en temps exponentiels des systèmes Hamiltoniens proches de systèmes intégrables: résonances et orbites fermèes", preprint.
Morbidelli, A., (1991): "On the successive eliminations of perturbation harmonics.", submitted to *Celest. Mech.*.
Nekhoroshev, N., N., (1977): "Exponential estimates of the stability time of near–integrable Hamiltonian systems.", *Russ. Math. Surveys*, **32**, 1-65.
Nekhoroshev, N., N., (1979): "Exponential estimates of the stability time of near–integrable Hamiltonian systems, 2.", *Trudy Sem. Petrovs.*, **5**, 5-50.
Pöschel, J., (1991): "On Nekhoroshev's estimate for quasi-convex Hamiltonians", preprint ETH, Zürich.
Simó, C., (1989): "Estabilitat de sistemes Hamiltonians.", *Memorias de la Real Academia de Ciencias y Artes de Barcelona*, **48**, 303–348.