

Giovanni Gallavotti, Guido Gentile <sup>1</sup>

**Abstract:** A selfcontained proof of the KAM theorem in the Thirring model is discussed, completely relaxing the “strong diophantine property” hypothesis used in previous papers.

Keywords: KAM, invariant tori, classical mechanics, perturbation theory, chaos

## 1. Introduction

In [G] a selfcontained proof of the KAM theorem in the Thirring model is discussed, under the hypothesis that the rotation vectors  $\vec{\omega}_0$  verify a *strong diophantine property*. At the end of the same paper a heuristic argument is given to show that in fact such a hypothesis can be relaxed. In the present work we develop the heuristic argument into an extension of the KAM theorem proof described in [G]; the extension applies to rotations vectors verifying only the usual diophantine condition. It is a proof again based on Eliasson’s method, [E].

In our opinion this shows that a hypothesis like the strong diophantine property of [G], or something similar to it, is very natural as it simplifies the structure of the proof by separating from it the analysis of a simple arithmetic property, whose untimely analysis would obscure the proof.

For an introductory discussion of the model and a more organic exposition of the problem, we refer to [G], [G1], and to the references there reported. In the remaining part of this section we confine ourselves to define the model, to introduce the basic notation, and to give the result we have obtained.

The Thirring model, [T], is described by the hamiltonian:

$$\frac{1}{2}J^{-1}\vec{A} \cdot \vec{A} + \varepsilon f(\vec{\alpha}) \quad (1.1)$$

where  $J$  is the (diagonal) matrix of the inertia moments,  $\vec{A} = (A_1, \dots, A_l) \in R^l$  are their angular momenta and  $\vec{\alpha} = (\alpha_1, \dots, \alpha_l) \in T^l$  are the angles describing their positions: the matrix  $J$  will be supposed non singular; but we only suppose that  $\min_{j=1, \dots, l} J_j = J_0 > 0$ , and no assumption is made on the size of the *twist rate*  $T = \min J_j^{-1}$ : the results will be uniform in  $T$  (hence they can be called “twistless results”). We suppose  $f$  to be an even trigonometric polynomial of degree  $N$ :

$$f(\vec{\alpha}) = \sum_{0 < |\vec{\nu}| \leq N} f_{\vec{\nu}} \cos \vec{\nu} \cdot \vec{\alpha}, \quad f_{\vec{\nu}} = f_{-\vec{\nu}}, \quad |\vec{\nu}| = \sum_{j=1}^l |\nu_j| \quad (1.2)$$

We shall consider a “rotation vector”  $\vec{\omega}_0 = (\omega_1, \dots, \omega_l) \in R^l$  verifying the *diophantine condition*:

$$\bar{C}_0 |\vec{\omega}_0 \cdot \vec{\nu}| \geq |\vec{\nu}|^{-\tau}, \quad \vec{0} \neq \vec{\nu} \in Z^l \quad (1.3)$$

with diophantine constants  $\bar{C}_0, \tau$ . The *diophantine vectors* have full measure in  $R^l$  if  $\tau$  is fixed  $\tau > l - 1$ . We shall set  $\vec{A}_0 = J\vec{\omega}_0$ .

As in [G], we prove the following result.

**Theorem:** *The system described by the Hamiltonian (1.1) admits an  $\varepsilon$ -analytic family of motions starting at  $\vec{\alpha} = \vec{0}$  and having the form:*

$$\vec{A} = \vec{A}_0 + \vec{H}(\vec{\omega}_0 t; \varepsilon), \quad \vec{\alpha} = \vec{\omega}_0 t + \vec{h}(\vec{\omega}_0 t; \varepsilon) \quad (1.4)$$

\* Archived in [mp\\_arc@math.utexas.edu](mailto:mp_arc@math.utexas.edu), #93-229, with the title *Non recursive proof of the KAM theorem*.

<sup>1</sup> Dipartimento di Fisica, Università di Roma, “La Sapienza”, P.le Moro 2, 00185 Roma, Italia. E-mail: [gallavotti@40221.hepnet1bl1.gov](mailto:gallavotti@40221.hepnet1bl1.gov), [gentileg@39943.hepnet1bl1.gov](mailto:gentileg@39943.hepnet1bl1.gov).

with  $\vec{H}(\vec{\psi}; \varepsilon), \vec{h}(\vec{\psi}; \varepsilon)$  analytic, divisible by  $\varepsilon$ , for  $|\operatorname{Im} \psi_j| < \xi$ ,  $\vec{\psi} \in T^l$ , and for  $|\varepsilon| < \varepsilon_0$  with:

$$\varepsilon_0^{-1} = bJ_0^{-1}(2^\tau \bar{C}_0)^2 f_0 N^{2+l} e^{cN} e^{\xi N} \quad (1.5)$$

where  $b, c$  are  $l$ -dependent positive constants,  $f_0 = \max_{\vec{\nu}} |f_{\vec{\nu}}|$ .

This means that the set  $\vec{A} = \vec{A}_0 + \vec{H}(\vec{\psi}; \varepsilon)$ ,  $\vec{\alpha} = \vec{\psi} + \vec{h}(\vec{\psi}; \varepsilon)$  described as  $\vec{\psi}$  varies in  $T^l$  is, for  $\varepsilon$  small enough, an invariant torus for (1.1), which is run quasi periodically with angular velocity vector  $\vec{\omega}_0$ . It is a family of invariant tori coinciding, for  $\varepsilon = 0$ , with the torus  $\vec{A} = \vec{A}_0$ ,  $\vec{\alpha} = \vec{\psi} \in T^l$ . The presence of the factor  $2^\tau$  marks the only difference from the analogous result in [G].

Calling  $\vec{H}^{(k)}(\vec{\psi}), \vec{h}^{(k)}(\vec{\psi})$  the  $k$ -th order coefficients of the Taylor expansion of  $\vec{H}, \vec{h}$  in powers of  $\varepsilon$  and writing the equation of motion as  $\dot{\vec{\alpha}} = J^{-1} \vec{A}$  and  $\dot{\vec{A}} = -\varepsilon \partial_{\vec{\alpha}} f(\vec{\alpha})$  we get immediately recursion relations for  $\vec{H}^{(k)}, \vec{h}^{(k)}$ , namely, for  $k > 1$ :

$$\begin{aligned} \vec{\omega}_0 \cdot \vec{\partial} h_j^{(k)} &= J_j^{-1} H_j^{(k)} \\ \vec{\omega}_0 \cdot \vec{\partial} H_j^{(k)} &= - \sum_{\substack{m_1, \dots, m_l \\ |\vec{m}| > 0}} \frac{1}{\prod_{s=1}^l m_s!} \partial_{\alpha_j} \partial_{\alpha_1}^{m_1} \dots \partial_{\alpha_l}^{m_l} f(\vec{\omega}_0 t) \cdot \sum^* \prod_{s=1}^l \prod_{j=1}^{m_s} h_s^{(k_j^s)}(\vec{\omega}_0 t) \end{aligned} \quad (1.6)$$

where the  $\sum^*$  denotes summation over the integers  $k_j^s \geq 1$  with:  $\sum_{s=1}^l \sum_{j=1}^{m_s} k_j^s = k - 1$ .

The trigonometric polynomial  $\vec{h}^{(k)}(\vec{\psi})$  will be completely determined (if possible at all) by requiring it to have  $\vec{0}$  average over  $\vec{\psi}$ , (note that  $\vec{H}^{(k)}$  has to have zero average over  $\vec{\psi}$ ). For  $k = 1$  it is:

$$\vec{h}^{(1)}(\vec{\psi}) = - \sum_{\vec{\nu} \neq \vec{0}} \frac{iJ^{-1} \vec{\nu}}{(i\vec{\omega}_0 \cdot \vec{\nu})^2} f_{\vec{\nu}} e^{i\vec{\nu} \cdot \vec{\psi}} \quad (1.7)$$

One easily finds that the equation for  $\vec{h}^{(k)}$  can be solved and its solution is a trigonometric polynomial in  $\vec{\psi}$ , of degree  $\leq kN$ , odd if  $\vec{h}^{(k)}$  is determined by imposing that its average over  $\vec{\psi}$  vanishes.

The remaining part of the paper is structured as follows: in section 2 we set a diagrammatic expansion of  $\vec{h}^{(k)}$ , as in [G]. In section 3 we discuss a proposition which leads to the original result of this paper, and in section 4 we prove the theorem, repeating the discussion in [G], with some minor changes due to the weakening of the strong diophantine property hypothesis.

The above theorem fully reproduces, in the model (1.1), the theorem of Eliasson: for another alternative proof of the same theorem with no assumption of parity or of finite degree on the trigonometric polynomial  $f$ , see [CF].

## 2. Diagrammatic expansion

Let  $\vartheta$  be a tree diagram: it will consist of a family of “lines” (*i.e.* segments) numbered from 1 to  $k$  arranged to form a (rooted) tree diagram as in the figure:

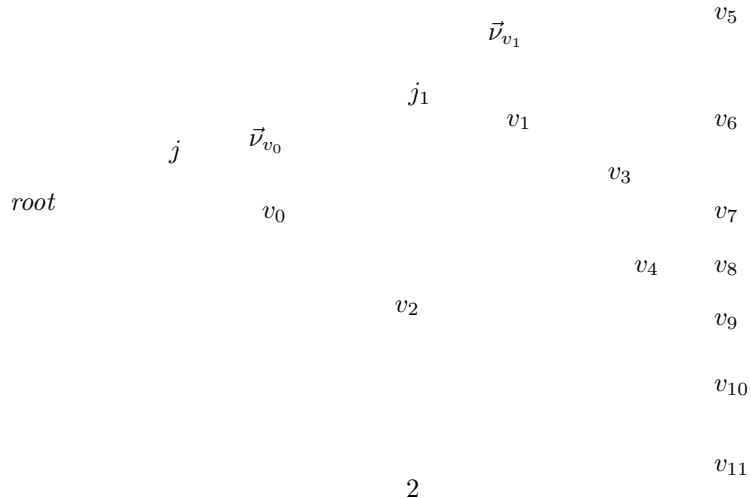


fig. 1: A tree diagram  $\vartheta$  with  $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2, m_{v_4} = 2$  and  $m = 12, \prod m_v! = 2^4 \cdot 6$ , and some decorations. The line numbers, distinguishing the lines, are not shown.

To each vertex  $v$  we attach a “mode label”  $\vec{v}_v \in Z^l, |\vec{v}_v| \leq N$  and to each branch leading to  $v$  we attach a “branch label”  $j_v = 1, \dots, l$ . The order of the diagram will be  $k = \text{number of vertices} = \text{number of branches}$  (the tree root will not be regarded as a vertex).

We imagine that all the diagram lines have the same length (even though they are drawn with arbitrary length in fig.1). A group acts on the set of diagrams, generated by the permutations of the subdiagrams having the same vertex as root. Two diagrams that can be superposed by the action of a transformation of the group will be regarded as identical (recall however that the diagram lines are numbered, *i.e.* are regarded as distinct, and the superposition has to be such that all the decorations of the diagram match). Tree diagrams are regarded as partially ordered sets of vertices (or lines) with a minimal element given by the root (or the root line). We shall imagine that each branch carries also an arrow pointing to the root (“gravity” direction, opposite to the order).

We define the “momentum” entering  $v$  as  $\vec{v}(v) = \sum_{w \geq v} \vec{v}_w$ : therefore the momentum entering a vertex  $v$  is given by the sum of the momenta entering the immediately following vertices plus the “momentum emitted” by  $v$  (*i.e.* the mode  $\vec{v}_v$ ). If from a vertex  $v$  emerge  $m_1$  lines carrying a label  $j = 1, m_2$  lines carrying  $j = 2, \dots$ , it follows that (1.6) can be rewritten:

$$h_{\vec{v}j}^{(k)} = \frac{1}{k!} \sum^* \prod_{v \in \vartheta} \frac{(-iJ^{-1}\vec{v}_v)_{j_v} f_{\vec{v}_v} \prod_{s=1}^l (i\vec{v}_v)_s^{m_s}}{(i\vec{\omega}_0 \cdot \vec{v}(v))^2} \quad (2.1)$$

with the sum running over the diagrams  $\vartheta$  of order  $k$  and with  $\vec{v}(v_0) = \vec{v}$  and  $j_{v_0} = j$ ; and the combinatorics can be checked from (1.6), by taking into account that we regard the diagram lines as all different (to fix the factorials). The  $*$  recalls that the diagram  $\vartheta$  can and will be supposed such that  $\vec{v}(v) \neq \vec{0}$  for all  $v \in \vartheta$  (by the remarked parity properties of  $\vec{h}^{(k)}$ ).

As in [G], according to Eliasson’s terminology, we can define the *resonant diagrams* as the diagrams with vertices  $v', v$ , with  $v' < v$ , not necessarily nearest neighbours, such that  $\vec{v}(v) = \vec{v}(v')$ . If there were no resonant diagrams, it would be straightforward to obtain a bound like (1.5). However *there are* resonant diagrams: the key remark is that they cancel *almost exactly*. Here we expand the heuristic discussion on the two above remarks of [G].

To see that convergence problems can arise only if resonant diagrams are considered, we shall first restrict the sum in (2.1) to a sum over diagrams such that: i)  $\vec{v}(v) \neq \vec{0}$  if  $v$  is any vertex, and: ii)  $\vec{v}(v) \neq \vec{v}(v')$  for all pairs of comparable vertices  $v', v$ , (not necessarily next to each other in the diagram order, however), with  $v \geq v_0$ .

There are at most  $2^{2k} k!$  diagrams and the labels  $j$  can be at most  $l^k$  while the  $\vec{v}_v$  can be chosen in a number of ways bounded by  $(2N + 1)^{lk} < (3N)^{lk}$ . Therefore, if  $f_0 = \max_{\vec{v}} |f_{\vec{v}}|$ :

$$|h_{\vec{v}j}^{(k)}| \leq (3N)^{lk} 2^{2k} l^k \frac{f_0^k C_0^{2k}}{J_0^k} N^{2k-1} \max_{\vartheta} \prod_{v \in \vartheta} (C_0 \vec{\omega}_0 \cdot \vec{v}(v))^{-2} \leq (f_0 C_0^2 J_0^{-1})^k N^{(l+2)k-1} (4l3^l)^k M$$

where the maximum is over the diagrams  $\vartheta$  verifying the i),ii) above. Hence the whole problem is reduced to estimating the maximum with  $M$ .

Let  $\zeta(q) \equiv (3N^\tau q^\tau)^{-1}$ , where  $\tau$  is the diophantine constant in (1.3): then  $C_0 |\vec{\omega}_0 \cdot \vec{v}| \geq 3\zeta(q)$  if  $0 < |\vec{v}| \leq q$ ; we can say that  $\vec{v} \in Z^l$  is “ $q$ -singular” if  $C_0 |\vec{\omega}_0 \cdot \vec{v}| < \zeta(q)$ . Then the following (extension) of a lemma by Brjuno, see [P], holds for diagrams of degree  $k$  verifying i),ii) above. Fixed  $q \geq 1$  let  $N(k)$  be the number of  $q$ -singular harmonics, among the vertex momenta  $\vec{v}(v)$  of the vertices  $v \in \vartheta$ ; then  $N(k) \leq 2 \frac{k}{q}$ . Assuming the above claim true (we shall not use it outside the present heuristic argument), we fix an exponentially decreasing sequence  $\gamma^n$ , with  $\gamma > 1$ . The number of  $\gamma^{-n}$ -singular harmonics which are not  $q = \gamma^{-(n-1)}$ -singular is bounded by  $2k\gamma^n$ , (being trivially bounded by the number of  $\gamma^{-n}$ -singular harmonics!). Hence:

$$\prod_{v \in \vartheta} \frac{1}{(C_0 \vec{\omega}_0 \cdot \vec{v}(v))^2} \leq \prod_{n \leq n_0}^{-\infty} \zeta(\gamma^{-(n-1)})^{-4k\gamma^n} \leq \exp 4k \sum_{n \geq 1} 2^{-n} \log(3 \cdot 2^{\tau(n-1)} N^\tau) = M \quad (10)$$

where  $n_0 = 1 + \lceil \log(N3^{-\tau-1}) / \log \gamma \rceil$ . Therefore the series for the approximation to  $\vec{h}_{\vec{v}}^{(k)}$ , that we are considering because of the extra restriction in the sum Eq.(2.1), has radius of convergence in  $\varepsilon$  bounded below.

However *there are* resonant diagrams. The reason why they cancel almost exactly is very simple: imagine to detach from the diagram  $\vartheta$  with first node  $v'$  the subdiagram  $\vartheta_2$  with first node  $v$ . Then attach it to all the *remaining* vertices  $w \geq v', w \in \vartheta/\vartheta_2$  (i.e. the graph  $\vartheta$ , with the subgraph  $\vartheta_2$  removed). We obtain a family of diagrams whose contributions to  $h^{(k)}$  differ because:

- 1) some of the branches above  $v'$  have changed total momentum by the amount  $\vec{v}(v)$ : this means that some of the denominators  $(\vec{\omega} \cdot \vec{v}(w))^{-2}$  have become  $(\vec{\omega} \cdot \vec{v}(w) \pm \varepsilon)^{-2}$  if  $\varepsilon \equiv \vec{\omega}_0 \cdot \vec{v}(v)$ ; and:
- 2) because there is one of the vertex factors which changes by taking successively the values  $i\nu_{wj}$ ,  $j$  being the line label of the line leading to  $v$ , and  $w \in \vartheta/\vartheta_2$  is the vertex to which such line is reattached.

Hence if  $\vec{\omega} \cdot \vec{v} = \varepsilon = 0$  we would build in this *resummation* a quantity proportional to:  $\sum \vec{v}_w = \vec{v}(v') - \vec{v}(v)$  which is zero, because  $\vec{v}(v) = \vec{v}(v')$  means that the sum of the  $\vec{v}_w$ 's vanishes. Since  $\vec{\omega} \cdot \vec{v} = \varepsilon \neq 0$  we can expect to see a sum of order  $\varepsilon^2$ , if we sum as well on a overall change of sign of the  $\vec{v}_w$  values, (which sum up to  $\vec{0}$ ), and take into account explicitly the dependence on the mode labels of (2.1), (up to a possible change of some denominators, the product in (2.1) does not change, provided  $f$  is an even function, when the signs of the mode labels corresponding to the nodes in  $\vartheta \setminus \vartheta_2$  are *simultaneously* changed).

But this can be true only if  $\varepsilon \ll \vec{\omega} \cdot \vec{v}'$ , for any line momentum  $\vec{v}'$  of a line in  $\vartheta/\vartheta_2$ . If the latter property is not true this means that  $\vec{\omega} \cdot \vec{v}'$  is small and that there are many vertices in  $\vartheta/\vartheta_2$  of order of the amount needed to create a momentum with small divisors of order  $\varepsilon$ .

The tree diagrams will play the role of *Feynman diagrams* in field theory; and they will be plagued by overlapping divergences. They will therefore be collected into another family of graphs, on which the bounds are easy. Such graphs in field theory are called *trees*, (see [G4]); we perform the same operation but we shall not introduce explicitly the new notion of trees (related to our clusters defined below) to avoid obvious confusion with the trees just introduced above. The  $(\vec{\omega} \cdot \vec{v})^{-2}$  are the *propagators*, in our analogy.

In [G], a *scaling* parameter  $\gamma$  (which could be taken  $\gamma = 2$ ) was fixed. Then, defining the adimensional frequency  $\vec{\omega} \equiv \vec{C}_0 \vec{\omega}_0$ , a propagator  $(\vec{\omega} \cdot \vec{v})^{-2}$  was said to be *on scale*  $n$  if  $2^{n-1} < |\vec{\omega} \cdot \vec{v}| \leq 2^n$ , for  $n \leq 0$ , and it was set  $n = 1$  if  $1 < |\vec{\omega} \cdot \vec{v}|$ . Nevertheless, if we want to eliminate the strong diophantine property, we need to change the decomposition of the propagator. We define a new vector  $\vec{\omega} = 2^\tau \vec{C}_0 \vec{\omega}_0$ , and we say that a propagator  $(\vec{\omega} \cdot \vec{v})^{-2}$  is *on scale*  $n$  if  $\gamma_{n-1} < |\vec{\omega} \cdot \vec{v}| \leq \gamma_n$ , for  $n \leq 0$ , and we set  $n = 1$  if  $\gamma_0 < |\vec{\omega} \cdot \vec{v}|$ . Here  $\{\gamma_n\}$  is a sequence such that  $1/2 \leq \gamma_n 2^{-n} \leq 1$ , which will be suitably chosen: how to fix such a sequence will be explained in section 3. Here we outline that the fixing of the sequence depends on the rotation vector we have chosen: the conceptual advantage we have, with respect to the result obtained under the strong diophantine property hypothesis, is that the sequence of scales  $\gamma_n$  is not “prescribed a priori” but it is determined by the arithmetic properties of the rotation vector  $\vec{\omega}_0$  (see the ending comments in [G]).

Proceeding as in quantum field theory, see [G3], given a diagram  $\vartheta$ , we can attach a *scale label* to each line  $v'v$  in (8) (with  $v'$  being the vertex preceding  $v$ ): it is equal to  $n$  if  $n$  is the scale of the line propagator. Note that the labels thus attached to a diagram are uniquely determined by the diagram: they will have only the function of helping to visualize the orders of magnitude of the various diagram lines.

Looking at such labels we identify the connected clusters  $T$  of vertices that are linked by a continuous path of lines with the same scale label  $n_T$  or a higher one and that are maximal. We shall say that *the cluster*  $T$  *has scale*  $n_T$ . As far as we are concerned, we can visualize a tree as a collection of clusters between which there exists a relation of partial ordering.

Among the clusters we consider the ones with the property that there is only one diagram line entering them and only one exiting and both carry the same momentum. Here we use that the diagram lines carry an arrow pointing to the root: this gives a meaning to the words “incoming” and “outgoing”. If  $V$  is one such cluster we denote  $\lambda_V$  the incoming line: the line scale  $n = n_{\lambda_V}$  is smaller than the smallest scale  $n' = n_V$  of the lines inside  $V$ . We call  $w_1$  the vertex into which the line  $\lambda_V$  ends, inside  $V$ . *We say that such a*  $V$  *is a resonance if the number of lines contained in*  $V$  *is*  $\leq E 2^{-n\varepsilon}$ , where  $n = n_{\lambda_V}$ , and  $E, \varepsilon$  are defined by:  $E \equiv 2^{-3\varepsilon} N^{-1}$ ,  $\varepsilon = \tau^{-1}$ . We call  $n_{\lambda_V}$  *the resonance-scale*, and  $\lambda_V$  *a resonant line*.

To avoid confusion we note explicitly that, since a resonance is also a cluster, it will have, as such, a scale (its scale as a cluster) which is different (lower) than the “resonance-scale”. When referring to the resonance-scale, we shall always use the notation with the *trait-d’union*. Likewise we call scale of a resonant cluster the scale of a cluster which is also a resonance.

### 3. Multiscale decomposition of the propagator

Given  $\vec{\omega}_0$  verifying (1.3) with some constant  $\vec{C}_0$  and some  $\tau > 0$  we define  $C_0 \equiv 2^\tau \vec{C}_0$ : this leaves (1.3) still valid. Recalling that  $\vec{\omega} \equiv C_0 \vec{\omega}_0$ , (see §2), we define the set  $B_n$  of the values of  $|\vec{\omega} \cdot \vec{v}|$  as  $\vec{v}$  varies in the set  $0 \leq |\vec{v}| < (2^{n+3})^{-1/\tau}$ , with  $n = 0, -1, -2, \dots$ . The sets  $B_n$  verify the inclusion relation  $B_n \subset B_m$  if  $m < n$ . The main property of the sets  $B_n$  is that the spacing between their elements is at least  $2^\tau (2(2^{n+3})^{-1/\tau})^{-\tau} \geq 2^{n+3}$ , by the diophantine property (1.3); also  $x \in B_n$  is such that  $x > 2^{n+3}$ , if  $x \neq 0$ .

More abstractly let  $B_n$ ,  $n = 0, -1, \dots$ , be a sequence of sets such that i)  $0 \in B_n$ , ii)  $B_n \subset B_m$  if  $m < n$  and iii) the spacing between the points in  $B_n$  is at least  $2^{n+3}$  (the latter will be the *spacing property*). Then we can prove the following lemma:

**Lemma:** *There exists a sequence  $\gamma_0, \gamma_{-1}, \dots$  with  $\gamma_p \in [2^{p-1}, 2^p]$  such that:*

$$|x - \gamma_p| \geq 2^{n+1}, \quad \text{if } n \leq p \leq 0 \quad \text{and } x \in B_n \quad (3.1)$$

for all  $n \leq 0$ .

*Remark:* hence if  $x \in \cup_n B_n$  with  $|x| \leq \gamma_p$  then  $|x| < \gamma_p$ .

*Proof:* Note that if  $\gamma_p \in [2^{p-1}, 2^p] \equiv I_p$  the (3.1) are obviously verified for  $n > p - 3$ , hence we can suppose  $p \geq n + 3$ .

Fix  $p \leq 0$  and let  $G = [a, b]$  be an interval verifying what we shall call below the *property*  $\mathcal{P}_n$ :

$$\mathcal{P}_n : \quad |x - \gamma| \geq 2^{m+1} \quad \text{for all } n \leq m \leq -3, \quad \gamma \in G, \quad |G| \geq 2^{n+1} \quad (3.2)$$

Let  $G_{p-3} \equiv [2^{p-1}, b_{p-3}]$ , with  $b_{p-3} \in [2^{p-1}, 2^p]$ , be a *maximal* interval verifying property  $\mathcal{P}_{p-3}$  (note that  $G_{p-3}$  exists because  $x \in B_{p-3}$ ,  $x \neq 0$  implies  $x \geq 2^p$  by the spacing property, and it is  $b_{p-3} \geq 2^{p-1} + 2^{p-2}$ ).

Assume inductively that the intervals  $[a_n, b_n] = G_n$  can be so chosen that  $G_{n'} \subseteq G_{n''}$  if  $n' < n''$  and  $G_n$  is maximal among the intervals contained in  $G_{n+1}$  and verifying the property  $\mathcal{P}_n$ .

If we can check that the hypothesis implies the existence of an interval  $G \subseteq G_n$  verifying  $\mathcal{P}_{n-1}$  we shall be able to define  $G_{n-1}$  to be a maximal interval among the ones contained in  $G_n$  and with the property  $\mathcal{P}_{n-1}$ : in case of ambiguity we shall take  $G_{n-1}$  to coincide with the rightmost possible choice.

Clearly we shall be able to define  $\gamma_p = \lim_{n \rightarrow -\infty} b_n$ , which will verify (3.1).

To check the existence of  $G \subseteq G_n$  verifying  $\mathcal{P}_{n-1}$  we consider first the case in which  $B_{n-1}$  has one and only one point  $x$  in  $G_n$ . If  $|G_n| \geq 2^{n+2}$  and  $x$  is in the first half of  $G_n$  we can take, by the spacing property,  $G = [x + 2^n, \min\{b_n, x + 2^{n+2} - 2^n\}]$ ; if it is in the second half we take  $G = [\max\{a_n, x - 2^{n+2} + 2^n\}, x - 2^n]$ .

If, on the other hand,  $|G_n| < 2^{n+2}$  it is  $G_n \subset G_{n+1}$  strictly and, furthermore, the interval  $(x - 2^{n+2}, x + 2^{n+2})$  does not contain points of  $B_{n-1}$  other than  $x$  itself (by the spacing property). The strict inclusion implies that there is a point  $y \in B_n$  at distance exactly  $2^{n+1}$  from  $G_n$  (recall the maximality of  $G_n$ ).

Suppose that  $x$  is in the first half of  $G_n$  and  $y < a_n$ , i.e.  $y = a_n - 2^{n+1}$ ; then  $x - y < 2^{n+1} + 2^{n+1}$  contradicting the spacing property. Hence  $y > b_n$ , i.e.  $y = b_n + 2^{n+1}$ : in such case it cannot be, again by the spacing property, that  $x + 2^{n+2} > y = b_n + 2^{n+1}$ , so that  $b_n - x \geq 2^{n+1}$  and we can take  $G = [b_n - 2^n, b_n]$ . If  $x$  is in the second half the roles of left and right are exchanged.

This completes the analysis of the case in which only one point of  $B_{n-1}$  falls in  $G_n$ . The cases in which either no point or at least two points of  $B_n$  fall in  $G_n$  are analogous but easier. If two consecutive points  $x < y$  of  $B_{n-1}$  fall inside  $G_n$  we must have  $y - x \geq 2^{n+2}$  by the spacing property: hence  $G = [x + 2^n, y - 2^n] \subset G_n$  enjoys the property  $\mathcal{P}_{n-1}$ . If no point of  $B_{n-1}$  falls in  $G_n$  let  $y \in B_{n-1}$  be the closest point to  $G_n$ ; if its distance to  $G_n$  exceeds  $2^n$  we take  $G = G_n$ . Otherwise suppose that  $y > b_n$ : the spacing property implies that the interval  $(y - 2^{n+2}, y)$  is free of points of  $B_{n-1}$ . Hence if  $a = \max\{a_n, y - 2^{n+2} + 2^n\}$ ,  $b = y - 2^n$  then  $G = [a, b]$  has the property  $\mathcal{P}_{n-1}$ . If, instead,  $y < a_n$  we set  $a = y + 2^n > a_n$  and  $b = \min\{b_n, y + 2^{n+2} - 2^n\}$  and  $G = [a, b]$  enjoys property  $\mathcal{P}_{n-1}$ .

*Remark:* note that the above proof is constructive.

Consider the special case in which the sets  $B_n$  are the ones defined at the beginning of the section. The above lemma can then be translated into the following arithmetic proposition.

**Proposition:** *Given a diophantine vector  $\vec{\omega}_0$ , i.e. a vector verifying (1.3), let  $C_0 = 2^\tau \vec{C}_0$  and  $\vec{\omega} = C_0 \vec{\omega}_0$ ; then it is possible to find a sequence  $\gamma_p \in [2^{p-1}, 2^p]$ , and:*

$$|\vec{\omega} \cdot \vec{v}| - \gamma_p \geq 2^{n+1}, \quad \text{if } 0 < |\vec{v}| \leq (2^{n+3})^{-\tau^{-1}} \quad (3.3)$$

for all  $n \leq 0$  and for all  $p \geq n$ . Furthermore  $|\vec{\omega} \cdot \vec{v}| \neq \gamma_n$ , for all  $n \leq 0$ .

*Remark:*

1) The sequence  $\gamma_p$  is constructively defined by the proof above.

2) The (3.3) is very similar to the condition added in [G] to the (1.3) to define the *strong diophantine property*. The point of the present paper is that all that is really needed (see the following §4) to prove the theorem in §1 are (1.3) and (3.3): the latter is a simple arithmetic property which is in fact a consequence of (1.3).

3) As remarked in [G] almost all  $\vec{\omega}_0$  verify (1.3) for some  $\bar{C}_0$  and some  $\tau > l-1$  with a sequence  $\gamma_p$  that can be prescribed *a priori* as  $\gamma_p = 2^p$ . This, however, leaves out important cases like  $l = 2$  and  $\vec{\omega}_0$  with a quadratic irrational as rotation number. And it has the very unfortunate drawback of being non constructive, as the set of full measure of the  $\vec{\omega}$  verifying the strong diophantine property is constructed by abstract nonsense arguments (*e.g.* the Borel Cantelli lemma). Nevertheless considering strongly diophantine vectors is natural as it leads to a simplified proof, [G], of the KAM theorem, with Eliasson's method, eliminating a "secondary" difficulty.

4) For the purpose of comparison note that the final comment of ref [G] conjectures the above proposition: however the constant  $C_0$  introduced there is  $2C_0$  in the above notations and, therefore, the quantities called there  $\gamma_p$  are 2 times the ones in (3.3) (in other words the first inequality in (3.3) has to be multiplied side by side by 2 to become the statement of [G]).

#### 4. Proof of the theorem

Let us consider a diagram  $\vartheta$  and its clusters. We wish to estimate the number  $N_n$  of lines with scale  $n \leq 0$  in it, assuming  $N_n > 0$  (we remind that a line is on scale  $n$ , if the line propagator  $(\vec{\omega} \cdot \vec{v})^{-2}$  is such that  $\gamma_{n-1} < |\vec{\omega} \cdot \vec{v}| < \gamma_n$ , see also the remark after (3.1)).

Denoting  $T$  a cluster of scale  $n$  let  $m_T$  be the number of resonances of resonance-scale  $n$  contained in  $T$  (*i.e.* with incoming lines of scale  $n$ ); we have the following inequality, valid for any diagram  $\vartheta$ :

$$N_n \leq \frac{4k}{E 2^{-\varepsilon n}} + \sum_{T, n_T=n} (-1 + m_T) \quad (4.1)$$

with  $E = N^{-1} 2^{-3\varepsilon}$ ,  $\varepsilon = \tau^{-1}$ . This is an extension of Brjuno's lemma, [B], [P], called in [G] "resonant Siegel-Brjuno bound": the proof, extracted from [G], can be found in appendix A.

Consider a diagram  $\vartheta^1$ ; we define the family  $\mathcal{F}(\vartheta^1)$  generated by  $\vartheta^1$  as follows. Given a resonance  $V$  of  $\vartheta^1$  we detach the part of  $\vartheta^1$  above  $\lambda_V$  and attach it successively to the points  $w \in \tilde{V}$ , where  $\tilde{V}$  is the set of vertices of  $V$  (including the endpoint  $w_1$  of  $\lambda_V$  contained in  $V$ ) outside the resonances contained in  $V$ . Note that "above" is with respect to the trees partial ordering which is *opposite* to that of the arrows. We say that a line is in  $\tilde{V}$ , if it is contained in  $V$  and has at least one point in  $\tilde{V}$ . Note that all the lines  $\lambda$  in  $\tilde{V}$  have a scale  $n_\lambda \geq n_V$ , (as *all* the lines  $\lambda$  in  $V$ ).

For each resonance  $V$  of  $\vartheta^1$  we shall call  $M_V$  the number of vertices in  $\tilde{V}$ . To the just defined set of diagrams we add the diagrams obtained by reversing simultaneously the signs of the vertex modes  $\vec{v}_w$ , for  $w \in \tilde{V}$ , (this can be done without breaking the relationship which has to exist between the lines, as it can be easily checked by observing that  $\sum_{w \in \tilde{V}} \vec{v}_w = \vec{0}$ , since, for any resonance  $V$ ,  $\sum_{v \in V} \vec{v}_v = \vec{0}$ ): the change of sign is performed independently for the various resonant clusters. This defines a family of  $\prod 2M_V$  diagrams that we call  $\mathcal{F}(\vartheta^1)$ . The number  $\prod 2M_V$  will be bounded by  $\exp \sum 2M_V \leq e^{2k}$ .

Let  $\lambda$  be a line, in a cluster  $T$ , contained inside the resonances  $V = V_1 \subset V_2 \subset \dots$  of resonance-scales  $n = n_1 > n_2 > \dots$ ; then the shifting of the lines  $\lambda_{V_i}$  can cause a change in the size of the propagator of  $\lambda$  by at most  $\gamma_{n_1} + \gamma_{n_2} + \dots < 2^{n_1} + 2^{n_2} + \dots < 2^{n+1}$ .

Since the number of lines inside  $V$  is smaller than  $\bar{N}_n \equiv E 2^{-n\tau^{-1}}$ , ( $E = 2^{-3\tau^{-1}} N^{-1}$ ), the quantity  $\vec{\omega} \cdot \vec{v}_\lambda$  of  $\lambda$  has the form  $\vec{\omega} \cdot \vec{v}_\lambda^0 + \sigma_\lambda \vec{\omega} \cdot \vec{v}_{\lambda_V}$  if  $\vec{v}_\lambda^0$  is the momentum of the line  $\lambda$  "inside the resonance  $V$ ", *i.e.* it is the sum of all the vertex modes of the vertices preceding  $\lambda$  in the sense of the line arrows, but contained in  $V$ ; and  $\sigma_\lambda = 0, \pm 1$ .

Therefore not only  $|\vec{\omega} \cdot \vec{v}_\lambda^0| > 2^{n+3}$  (because  $\vec{v}_\lambda^0$  is a sum of  $\leq \bar{N}_n$  vertex modes, so that  $|\vec{v}_\lambda^0| \leq N \bar{N}_n$ ) but  $\vec{\omega} \cdot \vec{v}_\lambda^0$  is "in the middle" of the interval of scales containing it and, by the proposition in section 3 (in [G] this was a consequence of the strong diophantine property) does not get out of it if we add a quantity bounded by  $2^{n+1}$  (like  $\sigma_\lambda \vec{\omega} \cdot \vec{v}_{\lambda_V}$ ). Hence no line changes scale as  $\vartheta$  varies in  $\mathcal{F}(\vartheta^1)$ .

Let  $\vartheta^2$  be a diagram not in  $\mathcal{F}(\vartheta^1)$  and construct  $\mathcal{F}(\vartheta^2)$ , etc. We define a collection  $\{\mathcal{F}(\vartheta^i)\}_{i=1,2,\dots}$  of pairwise disjoint families of diagrams. We shall sum all the contributions to  $\vec{h}^{(k)}$  coming from the individual members of each family. This is similar to the *Eliasson's resummation*.

We call  $\varepsilon_V$  the quantity  $\vec{\omega} \cdot \vec{v}_{\lambda_V}$  associated with the resonance  $V$  with resonance-scale  $n$ . If  $\lambda$  is a line in  $\tilde{V}$ , (see paragraphs following (4.1)), we can imagine to write the quantity  $\vec{\omega} \cdot \vec{v}_\lambda$  as  $\vec{\omega} \cdot \vec{v}_\lambda^0 + \sigma_\lambda \varepsilon_V$ , with  $\sigma_\lambda = 0, \pm 1$ .

We want to show that the product of the propagators is holomorphic in  $\varepsilon_V$  for  $|\varepsilon_V| < \gamma_{n_V-3}$ . Let us reason in the following way. If  $\lambda$  is a line on scale  $n_V$ ,  $\gamma_{n_V} > |\vec{\omega} \cdot \vec{v}_\lambda| > \gamma_{n_V-1}$ ; remarking that it is  $|\vec{\omega} \cdot \vec{v}_\lambda^0| > 2^{n+3}$ , we obtain immediately  $|\vec{\omega} \cdot \vec{v}_\lambda| > 2^{n+3} - 2^n > 2^{n+2}$ , so that  $n_V \geq n + 3$ . On the other hand, if  $n_V > n + 3$ , i.e.  $n_V = n + m$ , for some  $m > 3$ , we note that  $|\vec{\omega} \cdot \vec{v}_\lambda^0| > \gamma_{n_V-1} - \gamma_n$ , because the resonance-scales and the scales of the resonant clusters (and of all the lines) do not change, so that it follows that, for  $|\varepsilon_V| < \gamma_{n_V-3}$ ,  $|\vec{\omega} \cdot \vec{v}_\lambda^0 + \sigma_\lambda \varepsilon_V| \geq \gamma_{n_V-1} - \gamma_n - \gamma_{n_V-3} \geq (2^{n_V-2} - 2^{n_V-m}) - \gamma_{n_V-3} \geq (2^{n_V-3} + 2^{n_V-4} + \dots + 2^{n_V-m+1}) - 2^{n_V-3} \geq 2^{n_V-4}$ ; otherwise, if  $n_V = n + 3$ , we note that  $|\vec{\omega} \cdot \vec{v}_\lambda^0| > 2^{n+3}$ , so that  $|\vec{\omega} \cdot \vec{v}_\lambda^0 + \sigma_\lambda \varepsilon_V| > 2^{n+3} - \gamma_{n_V-3} \geq 2^{n_V-1}$ , for  $|\varepsilon_V| < \gamma_{n_V-3}$ . Therefore we can conclude that, while  $\varepsilon_V$  varies in a complex disk of radius  $\gamma_{n_V-3}$  and center in 0, the quantity  $|\vec{\omega} \cdot \vec{v}_\lambda^0 + \sigma_\lambda \varepsilon_V|$  does not become smaller than  $2^{n_V-4}$ . Note the main point here: the quantity  $\gamma_{n_V-3}$  will usually be  $\gg \gamma_{n_{\lambda_V}}$  which is the value  $\varepsilon_V$  actually can reach in every diagram in  $\mathcal{F}(\vartheta^1)$ ; this can be exploited in applying the maximum principle, as done below.

It follows that, calling  $n_\lambda$  the scale of the line  $\lambda$  in  $\vartheta^1$ , each of the  $\prod 2M_V \leq e^{2k}$  products of propagators of the members of the family  $\mathcal{F}(\vartheta^1)$  can be bounded above by  $\prod_\lambda 2^{-2(n_\lambda-4)} = 2^{8k} \prod_\lambda 2^{-2n_\lambda}$ , if regarded as a function of the quantities  $\varepsilon_V = \vec{\omega} \cdot \vec{v}_{\lambda_V}$ , for  $|\varepsilon_V| \leq \gamma_{n_V-3}$ , associated with the resonant clusters  $V$ . This even holds if the  $\varepsilon_V$  are regarded as independent complex parameters.

By construction it is clear that the sum of the  $\prod 2M_V \leq e^{2k}$  terms, giving the contribution to  $\vec{h}^{(k)}$  from the trees in  $\mathcal{F}(\vartheta^1)$ , vanishes to second order in the  $\varepsilon_V$  parameters (by the approximate cancellation discussed above in §2). Hence by the maximum principle, and recalling that each of the scalar products in (2.1) can be bounded by  $N^2$ , we can bound the contribution from the family  $\mathcal{F}(\vartheta^1)$  by:

$$\left[ \frac{1}{k!} \left( \frac{f_0 2^{2\tau} C_0^2 N^2}{J_0} \right)^k 2^{8k} e^{2k} \prod_{n \leq 0} 2^{-2nN_n} \right] \left[ \prod_{n \leq 0} \prod_{T, n_T=n} \prod_{i=1}^{m_T} 2^{2(n-n_i+4)} \right] \quad (4.2)$$

where:

- 1)  $N_n$  is the number of propagators of scale  $n$  in  $\vartheta^1$  ( $n = 1$  does not appear as  $|\vec{\omega} \cdot \vec{v}| \geq \gamma_0 \geq 2^2$ , in such cases, and  $2^8 \geq 2^4$ );
- 2) the first square bracket is the bound on the product of individual elements in the family  $\mathcal{F}(\vartheta^1)$  times the bound  $e^{2k}$  on their number;
- 3) the second term is the part coming from the maximum principle (in the form of Schwarz's lemma), applied to bound the resummations, and is explained as follows:
  - i) the dependence on the variables  $\varepsilon_{V_i} \equiv \varepsilon_i$  relative to resonances  $V_i \subset T$  with resonance-scale  $n_{\lambda_{V_i}} = n$  is holomorphic for  $|\varepsilon_i| < \gamma_{n_i-3}$  if  $n_i \equiv n_{V_i}$ , provided  $n_i \geq n + 3$ , which holds in our case (see above);
  - ii) the resummation says that the dependence on the  $\varepsilon_i$ 's has a second order zero in each. Hence the maximum principle tells us that we can improve the bound given by the first factor in (4.2) by the product of factors  $(|\varepsilon_i| \gamma_{n_i-3}^{-1})^2 \leq 2^{2(n-n_i+4)}$ , if  $n_i \geq n + 3$  (of course the gain factor can be important only when  $\ll 1$ ). Appendix B illustrates the details.

Hence substituting (4.1) into (4.2) we see that the  $m_T$  is taken away by the first factor in  $2^{2n} 2^{-2n_i}$ , while the remaining  $2^{-2n_i}$  are compensated by the  $-1$  before the  $+m_T$  in (4.1), taken from the factors with  $T = V_i$  (note that there are always enough  $-1$ 's).

Hence the product (4.2) is bounded by:

$$\frac{1}{k!} (2^{2\tau} C_0^2 J_0^{-1} f_0 N^2)^k e^{2k} 2^{8k} 2^{8k} \prod_n 2^{-8nkE^{-1} 2^{\varepsilon n}} \leq \frac{1}{k!} B_0^k \quad (4.3)$$

with  $B_0 = 2^{18} e^2 (2^{2\tau} C_0^2 f_0 J_0^{-1}) N^2 \exp[2^{3\tau-1} (8N \log 2) \sum_{n=1}^{\infty} n 2^{-n\tau-1}]$ . To sum over the trees we note that, fixed  $\vartheta$  the collection of clusters is fixed. Therefore we only have to multiply (4.3) by the number of diagram shapes for  $\vartheta$ , ( $\leq 2^{2k} k!$ ), by the number of ways of attaching mode labels, ( $\leq (3N)^{lk}$ ), so that we can bound  $|h_{\vec{v}_j}^{(k)}|$  by (1.5).

## Appendix A: Resonant Siegel-Brjuno bound.

Calling  $N_n^*$  the number of non resonant lines carrying a scale label  $\leq n$ , (the notion of resonant line was defined at the end of §2), we shall prove first that  $N_n^* \leq 2k(E2^{-\varepsilon n})^{-1} - 1$  if  $N_n > 0$ . We fix  $n$  and denote  $N_n^*$  as  $N^*(\vartheta)$ .

Note that if  $k \leq E2^{-n\varepsilon}$  it is, for all momenta  $\vec{v}$  of the lines,  $|\vec{v}| \leq NE2^{-n\varepsilon}$ , *i.e.*  $|\vec{\omega} \cdot \vec{v}| \geq (NE2^{-n\varepsilon})^{-\tau} = 2^3 2^n$  so that there are *no* clusters  $T$  with  $n_T = n$  and  $N^*(\vartheta) = 0$ . If  $\vartheta$  has the root line with scale  $> n$  then calling  $\vartheta_1, \vartheta_2, \dots, \vartheta_m$  the subdiagrams of  $\vartheta$  emerging from the first vertex of  $\vartheta$  and with  $k_j > E2^{-\varepsilon n}$  lines, it is  $N^*(\vartheta) = N^*(\vartheta_1) + \dots + N^*(\vartheta_m)$  and the statement is inductively implied from its validity for  $k' < k$  provided it is true that  $N^*(\vartheta) = 0$  if  $k < E2^{-\varepsilon n}$ , which is certainly the case if  $E$  is chosen as in (4.1).

In the other case it is  $N_n^* \leq 1 + \sum_{i=1}^m N^*(\vartheta_i)$ , and if  $m = 0$  the statement is trivial, or if  $m \geq 2$  the statement is again inductively implied by its validity for  $k' < k$ .

If  $m = 1$  we once more have a trivial case unless the order  $k_1$  of  $\vartheta_1$  is  $k_1 > k - \frac{1}{2}E2^{-n\varepsilon}$ . Finally, and this is the real problem as the analysis of a few examples shows, we claim that in the latter case the root line of  $\vartheta_1$  is either a resonant line or it has scale  $> n$ .

Accepting the last statement it will be:  $N^*(\vartheta) = 1 + N^*(\vartheta_1) = 1 + N^*(\vartheta'_1) + \dots + N^*(\vartheta'_{m'})$ , with  $\vartheta'_j$  being the  $m'$  subdiagrams emerging from the first node of  $\vartheta'_1$  with orders  $k'_j > E2^{-\varepsilon n}$ : this is so because the root line of  $\vartheta_1$  will not contribute its unit to  $N^*(\vartheta_1)$ . Going once more through the analysis the only non trivial case is if  $m' = 1$  and in that case  $N^*(\vartheta'_1) = N^*(\vartheta''_1) + \dots + N^*(\vartheta''_{m''})$ , *etc.*, until we reach a trivial case or a diagram of order  $\leq k - \frac{1}{2}E2^{-n\varepsilon}$ .

It remains to check that if  $k_1 > k - \frac{1}{2}E2^{-n\varepsilon}$  then the root line of  $\vartheta_1$  has scale  $> n$ , unless it is entering a resonance.

Suppose that the root line of  $\vartheta_1$  has scale  $\leq n$  and is not entering a resonance; then  $|\vec{\omega} \cdot \vec{v}(v_0)| \leq \gamma_n, |\vec{\omega} \cdot \vec{v}(v_1)| \leq \gamma_n$ , if  $v_0, v_1$  are the first vertices of  $\vartheta$  and  $\vartheta_1$  respectively. Hence  $\delta \equiv |(\vec{\omega} \cdot (\vec{v}(v_0) - \vec{v}(v_1)))| \leq 22^n$  and the diophantine assumption implies that  $|\vec{v}(v_0) - \vec{v}(v_1)| > (22^n)^{-\tau^{-1}}$ , or  $\vec{v}(v_0) = \vec{v}(v_1)$ . The latter case being discarded as  $k - k_1 < \frac{1}{2}E2^{-n\varepsilon}$  (and we are not considering the resonances: note also that in such case the lines in  $\vartheta/\vartheta_1$  different from the root of  $\vartheta$  must be inside a cluster), it follows that  $k - k_1 < \frac{1}{2}E2^{-n\varepsilon}$  is inconsistent: it would in fact imply that  $\vec{v}(v_0) - \vec{v}(v_1)$  is a sum of  $k - k_1$  vertex modes and therefore  $|\vec{v}(v_0) - \vec{v}(v_1)| < \frac{1}{2}NE2^{-n\varepsilon}$  hence  $\delta > 2^3 2^n$  which is contradictory with the above opposite inequality.

A similar, far easier, induction can be used to prove that if  $N_n^* > 0$  then the number  $p$  of clusters of scale  $n$  verifies the bound  $p \leq 2k(E2^{-\varepsilon n})^{-1} - 1$ . In fact this is true for  $k \leq E2^{-\varepsilon n}$ , (see footnote 4). Let, therefore,  $p(\vartheta)$  be the number of clusters of scale  $n$ : if the first tree node  $v_0$  is not in a cluster of scale  $n$  it is  $p(\vartheta) = p(\vartheta_1) + \dots + p(\vartheta_m)$ , with the above notation, and the statement follows by induction. If  $v_0$  is in a cluster of scale  $n$  we call  $\vartheta_1, \dots, \vartheta_m$  the subdiagrams emerging from the cluster containing  $v_0$  and with orders  $k_j > E2^{-\varepsilon n}$ . It will be  $p(\vartheta) = 1 + p(\vartheta_1) + \dots + p(\vartheta_m)$ . Again we can assume that  $m = 1$ , the other cases being trivial. But in such case there will be only one branch entering the cluster  $V$  of scale  $n$  containing  $v_0$  and it will have a momentum of scale  $\leq n - 1$ . Therefore the cluster  $V$  must contain at least  $E2^{-\varepsilon n}$  nodes. This means that  $k_1 \leq k - E2^{-\varepsilon n}$ : thus (4.1) is proved.

## Appendix B: Dimensional estimate of the order of zero in $\varepsilon_i$ .

Consider a family  $\mathcal{F}(\vartheta^1) \equiv \mathcal{F}$ . Let  $B$  be the first factor in (4.2) without the  $e^{2k}$ , *i.e.* a “naive” bound on the sum of the values of each of the trees in the family.

Between the resonances there exists an inclusion relation; let us define “first generation resonances” the innermost resonances, *i.e.* the resonances  $V_{j_1}^1$ ,  $j_1 \geq 1$ , containing no other resonances, “second generation resonances” the next to innermost resonances  $V_{j_2}^2$ ,  $j_2 \geq 1$ , *i.e.* the resonances which become innermost if all the original innermost ones are regarded as single nodes, and so on. Let  $\varepsilon_{j_i}^i = \vec{\omega} \cdot \vec{\lambda}_{V_{j_i}^i}$ : each  $\varepsilon_{j_i}^i$  is a function of the values  $\varepsilon_{j_k}^k$ ,  $k > i$ , corresponding to the resonances with resonance-scale  $> n_{\lambda_{V_{j_i}^i}}$ .

Consider a first generation resonance  $V_1^1$  of scale  $n_{V_1^1}$ : it is  $|\varepsilon_1^1| < \gamma_{n_{\lambda_{V_1^1}}}$ , and the values of the trees in  $\mathcal{F}$  are analytic in  $\varepsilon_1^1$  for  $|\varepsilon_1^1| < \gamma_{n_{V_1^1}-3}$ ; note that if the  $\varepsilon_{j_i}^i$ 's,  $i > 1$ , vary in their analyticity domains,  $\varepsilon_1^1$ , considered as a function of them, can assume a value outside its own analyticity domain when  $n_{V_1^1} < n_{\lambda_{V_1^1}} + 5$ , although the contribution of the factors corresponding to the tree branches to the first square bracket in (4.2) remains as in (4.2).

Then if we sum the values of the considered trees collecting them into families (of  $2M_{V_1^1}$  terms) corresponding



to the Eliasson's resummation related to the resonance  $V_1^1$ , *only*, we obtain a sum of functions each of which has a zero of second order in  $\varepsilon_1^1$ , independently on the other values  $\varepsilon_{j_i}^i$ 's.

Therefore the considered sums are, each, bounded by:

$$B \left( 2M_{V_1^1} 2^{2(n_{\lambda_{V_1^1}} - n_{V_1^1} + 4)} \right),$$

when the “gain factor” can be left also when it is not obtained, (*i.e.* if  $n_{V_1^1} < n_{\lambda_{V_1^1}} + 5$ ), since, in such a case, the dimensional bound, which cannot be improved, is given by  $B$ , and  $n_{\lambda_{V_1^1}} - n_{V_1^1} + 4 = 0, 1$ .

We then consider another innermost resonance  $V_2^1$  (if existent). We perform the same resummation, obtaining a bound:

$$B \prod_{i=1}^2 \left( 2M_{V_i^1} 2^{2(n_{\lambda_{V_i^1}} - n_{V_i^1} + 4)} \right)$$

on each of the subfamilies of  $\mathcal{F}$  that we consider in this way (each consisting of  $2M_{V_1^1} 2M_{V_2^1}$  trees).

And we continue until all the  $N_1$  innermost resonances have been considered.

Then we consider a second generation resonance  $V_1^2$ . As we perform the resummation related to the  $2M_{V_1^2}$  terms associated with the new resonance, we regard the sum of the values of each of the groups of trees as a function of  $\varepsilon_1^2 = \vec{\omega} \cdot \vec{\nu}_{\lambda_{V_1^2}}$  (also the values  $\varepsilon_{j_1}^1$ 's corresponding to the innermost resonances contained in  $V_1^2$  are regarded as dependent on  $\varepsilon_1^2$ ): for all the values of  $\varepsilon_1^2$ , with  $|\varepsilon_1^2| < \gamma_{n_{V_1^2}-3}$ , such a sum is analytic if  $n_{V_1^1} \geq n_{\lambda_{V_1^1}} + 5$  for each first generation resonance, and is bounded, *in every case* by:

$$B_1 \equiv B \prod_{i=1}^{N_1} \left( 2M_{V_i^1} 2^{2(n_{\lambda_{V_i^1}} - n_{V_i^1} + 4)} \right)$$

as it can be argued analogously to the previous discussion. The further sum over the values of the  $2M_{V_1^2}$  elements involved in the new resummation creates a function of  $\varepsilon_1^2$  with a second order zero so that we can improve the bound of such a larger collection of trees by:

$$B_1 2M_{V_1^2} 2^{2(n_{\lambda_{V_1^2}} - n_{V_1^2} + 4)}$$

and we can continue in this way until the second generation of resonances is exhausted, and so on until no resonances are left, and there is only a big group of terms collected in the successive resummations (containing all the values of the trees in  $\mathcal{F}$ ) and the bound (4.2) is consequently obtained.

## References

- [A] Arnold, V.: *Proof of a A.N. Kolmogorov theorem on conservation of conditionally periodic motions under small perturbations of the hamiltonian function*, Uspeki Matematicheskii Nauk, 18, 13– 40, 1963.
- [B] Brjuno, A.: *The analytic form of differential equations*, I: Transactions of the Moscow Mathematical Society, 25, 131– 288, 1971; and II: 26, 199–239, 1972.
- [CF] Chierchia, L., Falcolini, C.: *A direct proof of the Theorem by Kolmogorov in Hamiltonian systems*, preprint, archived in [mp\\_arc@math.utexas.edu](mailto:mp_arc@math.utexas.edu), #93-171.
- [E] Eliasson, L. H.: *Hamiltonian systems with linear normal form near an invariant torus*, ed. G. Turchetti, Bologna Conference, 30/5 to 3/6 1988, World Scientific, 1989. And: *Generalization of an estimate of small divisors by Siegel*, ed. E. Zehnder, P. Rabinowitz, book in honor of J. Moser, Academic press, 1990. But mainly: *Absolutely convergent series expansions for quasi-periodic motions*, report 2–88, Dept. of Math., University of Stockholm, 1988.
- [G] Gallavotti, G.: *Twistless KAM tori.*, archived in [mp\\_arc@math.utexas.edu](mailto:mp_arc@math.utexas.edu), #93-172, to appear in Communications in Mathematical Physics. See also: Gallavotti, G.: *Invariant tori: a field theoretic point of view on Eliasson's work*, talk at the meeting in honor of the 60-th birthday of G. Dell'Antonio (Capri, may 1993), Marseille, CNRS-CPT, preprint
- [G2] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable hamiltonian systems. A review.*, deposited in the archive [mp\\_arc@math.utexas.edu](mailto:mp_arc@math.utexas.edu), #93-164, to appear in Reviews in Mathematical Physics.

- [G3] Gallavotti, G.: *The elements of Mechanics*, Springer, 1983.
- [G4] Gallavotti, G.: *Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods*, Reviews in Modern Physics, 57, 471- 572, 1985. See also: Gallavotti, G.: *Quasi integrable mechanical systems*, Les Houches, XLIII (1984), vol. II, p. 539- 624, Ed. K. Osterwalder, R. Stora, North Holland, 1986.
- [K] Kolmogorov, N.: *On the preservation of conditionally periodic motions*, Doklady Akademia Nauk SSSR, 96, 527- 530, 1954. See also: Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.M.: *A proof of Kolmogorov theorem on invariant tori using canonical transformations defined by the Lie method*, Nuovo Cimento, 79B, 201- 223, 1984.
- [M] Moser, J.: *On invariant curves of an area preserving mapping of the annulus*, Nachrichten Akademie Wiss. Göttingen, 11, 1- 20, 1962.
- [P] Pöschel, J.: *Invariant manifolds of complex analytic mappings*, Les Houches, XLIII (1984), vol. II, p. 949- 964, Ed. K. Osterwalder, R. Stora, North Holland, 1986.
- [PV] Percival, I., Vivaldi, F.: *Critical dynamics and diagrams*, Physica, D33, 304- 313, 1988.
- [S] Siegel, K.: *Iterations of analytic functions*, Annals of Mathematics, 43, 607- 612, 1943.
- [T] Thirring, W.: *Course in Mathematical Physics*, vol. 1, p. 133, Springer, Wien, 1983.