Non recursive proof of the KAM theorem *

Giovanni Gallavotti, Guido Gentile [1]

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:

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

where J is the (diagonal) matrix of the inertia moments, $\vec{A}=(A_1,\ldots,A_l)\in R^l$ are their angular momenta and $\vec{\alpha}=(\alpha_1,\ldots,\alpha_l)\in T^l$ are the angles describing their positions: the matrix J will be supposed non singular; but we only suppose that $\min_{j=1,\ldots,l}J_j=J_0>0$, 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:

$$f(\vec{\alpha})=\sum_{0<|\vec{\nu}|\le N} f_{\vec{\nu}}\cos\vec{\nu}\cdot\vec{\alpha},\qquad f_{\vec{\nu}}=f_{-\vec{\nu}}\,,\qquad |\vec{\nu}|=\sum_{j=1}^{l}|\vec{\nu}_j| \tag{1.2}$$

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

$$\bar{C}_0|\vec{\omega}_0\cdot\vec{\nu}|\ge|\vec{\nu}|^{-\tau},\qquad \vec{0}\ne\vec{\nu}\in Z^l \tag{1.3}$$

with diophantine constants $\bar{C}_0,\tau$. The *diophantine vectors* have full measure in R^l if τ 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 ε–analytic family of motions starting at $\vec{\alpha}=\vec{0}$ and having the form:*

$$\vec{A}=\vec{A}_0+\vec{H}(\vec{\omega}_0t;\varepsilon),\qquad \vec{\alpha}=\vec{\omega}_0t+\vec{h}(\vec{\omega}_0t;\varepsilon) \tag{1.4}$$

* Archived in mp_arc@math.utexas.edu; available also by e-mail request to the authors.

[1] Dipartimento di Fisica, Università di Roma, "La Sapienza", P.le Moro 2, 00185 Roma, Italia. E-mail: gallavotti%40221.hepnet@lbl.gov, gentileg%39943.hepnet@lbl.gov.

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

$$\varepsilon_0^{-1} = bJ_0^{-1}(2^\tau \bar C_0)^2 f_0 N^{2+l} e^{cN} e^{\xi N} \tag{1.5}$$

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 ε 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^τ 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 ε 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 $k > 1$:

$$\begin{aligned}
&\vec\omega_0 \cdot \vec\partial\, h_j^{(k)} = J_j^{-1} H_j^{(k)}\\
&\vec\omega_0 \cdot \vec\partial\, H_j^{(k)} = -\sum_{\substack{m_1,\dots,m_l\\ |\vec m|>0}} \frac{1}{\prod_{s=1}^l m_s!} \partial_{\alpha_j}\, \partial^{m_1+\dots+m_l}_{\alpha_1^{m_1}\dots\alpha_l^{m_l}} f(\vec\omega_0 t) \cdot \sum\nolimits^{*} \prod_{s=1}^{l}\prod_{j=1}^{m_s} h_s^{(k_j^s)}(\vec\omega_0 t)
\end{aligned} \tag{1.6}$$

where the $\sum^*$ denotes summation over the integers $k_j^s \ge 1$ 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:

$$\vec h^{(1)}(\vec\psi\,) = -\sum\nolimits_{\vec\nu\neq\vec 0} \frac{iJ^{-1}\vec\nu}{(i\vec\omega_0\cdot\vec\nu)^2} f_{\vec\nu}\, e^{i\vec\nu\cdot\vec\psi} \tag{1.7}$$

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 $\le kN$, 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 ϑ 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:

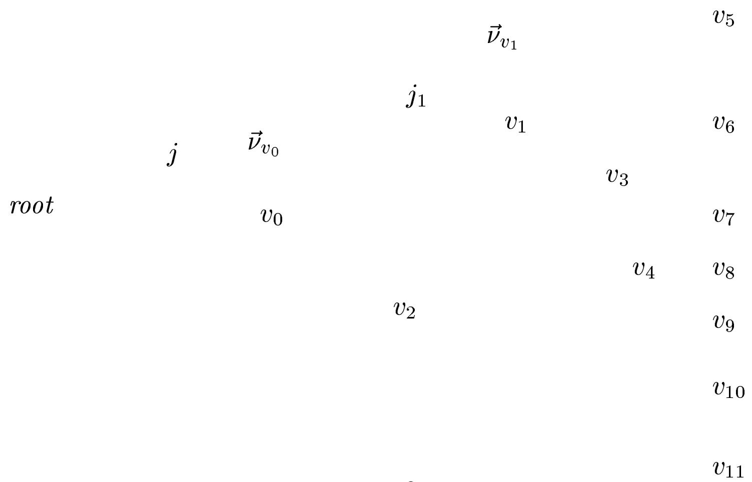

fig. 1: A tree diagram ϑ with $m_{v_0}=2, m_{v_1}=2, m_{v_2}=3, m_{v_3}=2, m_{v_4}=2$ and $m=12$, $\prod m_v! = 2^4 \cdot 6$, and some decorations. The line numbers, distinguishing the lines, are not shown.

To each vertex v we attach a "mode label" $\vec\nu_v \in Z^l$, $|\vec\nu_v| \le N$ and to each branch leading to v we attach a "branch label" $j_v = 1, \ldots, l$. The order of the diagram will be $k =$ number of vertices $=$ 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 superpositon 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 $\vec\nu(v) = \sum_{w \ge v} \vec\nu_w$: 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 $\vec\nu_v$). If from a vertex v emerge m_1 lines carrying a label $j=1$, m_2 lines carrying $j=2$, $\ldots$, it follows that (1.6) can be rewritten:

$$h^{(k)}_{\vec\nu j} = \frac{1}{k!} \sum\nolimits^* \prod_{v \in \vartheta} \frac{(-iJ^{-1}\vec\nu_v)_{j_v}\, f_{\vec\nu_v} \prod_{s=1}^{l} (i\vec\nu_v)_s^{m_s}}{(i\vec\omega_0 \cdot \vec\nu(v))^2} \tag{2.1}$$

with the sum running over the diagrams ϑ of order k and with $\vec\nu(v_0) = \vec\nu$; 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 ϑ can and will be supposed such that $\vec\nu(v) \ne \vec 0$ 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 $\vec\nu(v) = \vec\nu(v')$. If there were no resonant diagrams, it would be straightforward to obtain a bound like (1.5), as it is shown in [G]. However *there are* resonant diagrams.The key remark is that they cancel *almost exactly*. For a more detailed heuristic discussion about the two above remarks we refer again to [G].

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, that we shall call *trees*, on which the bounds are easy. The $(\vec\omega \cdot \vec\nu)^{-2}$ are the *propagators*, in our analogy.

In [G], a *scaling* parameter γ (which could be taken $\gamma = 2$) was fixed. Then, defining the adimensional frequency $\vec\omega \equiv \bar C_0 \vec\omega_0$, a propagator $(\vec\omega \cdot \vec\nu)^{-2}$ was said to be *on scale* n if $2^{n-1} < |\vec\omega \cdot \vec\nu| \le 2^n$, for $n \le 0$, and it was set $n = 1$ if $1 < |\vec\omega \cdot \vec\nu|$. Nevertheless, if we want to eliminate the strong diophantine property, we need to change the decomposition of the propagator. We define a new vector $\vec\omega = 2^\tau \bar C_0 \vec\omega_0$, and we say that a propagator $(\vec\omega \cdot \vec\nu)^{-2}$ is *on scale* n if $\gamma_{n-1} < |\vec\omega \cdot \vec\nu| \le \gamma_n$, for $n \le 0$, and we set $n = 1$ if $\gamma_0 < |\vec\omega \cdot \vec\nu|$. Here $\{\gamma_n\}$ is a sequence such that $1/2 \le \gamma_n 2^{-n} \le 1$, 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 γ_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 ϑ, 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. 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 λ_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 λ_V ends, inside V. *We say* *that such a V is a resonance if the number of lines contained in V is* $\le E\,2^{-n\varepsilon}$, where $n=n_{\lambda_V}$, and E,ε are defined by: $E\equiv 2^{-3\varepsilon}N^{-1}$, $\varepsilon=\tau^{-1}$. We call n_{λ_V} *the resonance scale*, and λ_V a *resonant line.*

3. Multiscale decomposition of the propagator

Given $\vec\omega_0$ verifying (1.3) with some constant $\bar C_0$ and some $\tau>0$ we define $C_0\equiv 2^\tau\bar C_0$: this leaves (1.3) still valid. We define the set B_n of the values of $|\vec\omega\cdot\vec\nu|$ as $\vec\nu$ varies in the set $0\le|\vec\nu|<(2^{n+3})^{-1/\tau}$, with $n=0,-1,-2,\ldots$. 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 $2^\tau(2(2^{n+3})^{-1/\tau})^{-\tau}\ge 2^{n+3}$, by the diophantine property (1.3); also $x\in B_n$ is such that $x>2^{n+3}$, if $x\ne0$.

More abstractly let B_n, $n=0,-1,\ldots$, 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 $\gamma_0,\gamma_{-1},\ldots$ with $\gamma_p\in[2^{p-1},2^p]$ such that:*

$$|x-\gamma_p|\ge 2^{n+1},\qquad\text{if}\quad n\le p\le 0\quad\text{and}\quad x\in B_n \tag{3.1}$$

for all $n\le 0$.

Remark: hence if $x\in\cup_n B_n$ with $|x|\le\gamma_p$ 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 $p\ge n+3$.

Fix $p\le0$ and let $G=[a,b]$ be an interval verifying what we shall call below the *property* $\mathcal P_n$:

$$\mathcal P_n:\qquad |x-\gamma|\ge 2^{m+1}\quad\text{for all}\quad n\le m\le -3,\qquad \gamma\in G,\qquad |G|\ge 2^{n+1} \tag{3.2}$$

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 $x\in B_{p-3}$, $x\ne0$ implies $x\ge 2^p$ by the spacing property, and it is $b_{p-3}\ge 2^{p-1}+2^{p-2}$).

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 rigthmost possible choice.

Clearly we shall be able to define $\gamma_p=\lim_{n\to-\infty}b_n$, 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 $|G_n|\ge 2^{n+2}$ 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 $b_n-x\ge 2^{n+1}$ 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. [2]

Remark: note that the above proof is constructive.

[2] If two consecutive points $x<y$ of B_{n-1} fall inside G_n we must have $y-x\ge 2^{n+2}$ 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 $\mathcal P_{n-1}$.

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 $C_0 = 2^\tau \bar C_0$ and $\vec\omega = C_0\vec\omega_0$; it is possible to find a sequence $\gamma_p \in [2^{p-1}, 2^p]$ such that $\gamma_{p-1} \le \gamma_p$ and:*

$$\big||\vec\omega\cdot\vec\nu| - \gamma_p\big| \ge 2^{n+1}, \qquad \text{if} \quad 0 < |\vec\nu| \le (2^{n+3})^{-\tau^{-1}} \tag{3.3}$$

for all $n \le 0$ and for all $p \ge n$. Furthermore $|\vec\omega\cdot\vec\nu| \ne \gamma_n$, for all $n \le 0$.

Remark:
1) The sequence γ_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 γ_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 one side 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 γ_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 ϑ and its clusters. We wish to estimate the number N_n of lines with scale $n \le 0$ in it, assuming $N_n > 0$ (we remind that a line is on scale n, if the line propagator $(\vec\omega\cdot\vec\nu)^{-2}$ is such that $\gamma_{n-1} < |\vec\omega\cdot\vec\nu| < \gamma_n$, see also the remark after (3.1)).

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

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

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.

Consider a diagram ϑ^1; we define the family $\mathcal{F}(\vartheta^1)$ generated by ϑ^1 as follows. Given a resonance V of ϑ^1 we detach the part of ϑ^1 above λ_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 λ_V contained in V) outside the resonances contained in V. 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 λ in $\tilde V$ have a scale $n_\lambda \ge n_V$.

For each resonance V of ϑ^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 simoultaneously the signs of the vertex modes $\vec\nu_w$, for $w \in \tilde V$[3]: 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 $\exp \sum 2M_V \le e^{2k}$.

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

[3] This can be done without breaking the relationship which has to exist between the lines, as it can be easily checked by observing that $\sum_{w\in\tilde V} \vec\nu_w = \vec 0$, since, for any resonance V, $\sum_{v\in V} \vec\nu_v = \vec 0$.

Since the number of lines inside V is smaller than $\overline{N}_n \equiv E2^{-n\tau^{-1}}$, ($E = 2^{-3\tau^{-1}}N^{-1}$), the quantity $\vec{\omega}\cdot\vec{\nu}_\lambda$ of λ has the form $\vec{\omega}\cdot\vec{\nu}^0_\lambda + \sigma_\lambda\vec{\omega}\cdot\vec{\nu}_{\lambda_V}$ if $\vec{\nu}^0_\lambda$ is the momentum of the line λ "inside the resonance V", *i.e.* it is the sum of all the vertex modes of the vertices preceding λ in the sense of the line arrows, but contained in V; and $\sigma_\lambda = 0, \pm 1$.

Therefore not only $|\vec{\omega}\cdot\vec{\nu}^0_\lambda| > 2^{n+3}$ (because $\vec{\nu}^0_\lambda$ is a sum of $\le \overline{N}_n$ vertex modes, so that $|\vec{\nu}^0_\lambda| \le N\overline{N}_n$) but $\vec{\omega}\cdot\vec{\nu}^0_\lambda$ 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 $\sigma_\lambda\vec{\omega}\cdot\vec{\nu}_{\lambda_V}$). *Hence no line changes scale as ϑ varies in $\mathcal{F}(\vartheta^1)$.*

Let ϑ^2 be a diagram not in $\mathcal{F}(\vartheta^1)$ and construct $\mathcal{F}(\vartheta^2)$, *etc.* We define a collection $\{\mathcal{F}(\vartheta^i)\}_{i=1,2,\ldots}$ 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 ε_V the quantity $\vec{\omega}\cdot\vec{\nu}_{\lambda_V}$ associated with the resonance V with scale n. If λ is a line in $\tilde{V}$, (see paragraphs following (4.1)), we can imagine to write the quantity $\vec{\omega}\cdot\vec{\nu}_\lambda$ as $\vec{\omega}\cdot\vec{\nu}^0_\lambda + \sigma_\lambda\varepsilon_V$, with $\sigma_\lambda = 0, \pm 1$.

We want to show that the product of the propagators is holomorphic in ε_V for $|\varepsilon_V| < \gamma_{n_V-3}$. Let us reason in the following way. If λ is a line on scale n_V, $\gamma_{n_V} > |\vec{\omega}\cdot\vec{\nu}_\lambda| > \gamma_{n_V-1}$; remarking that it is $|\vec{\omega}\cdot\vec{\nu}^0_\lambda| > 2^{n+3}$, we obtain immediately $|\vec{\omega}\cdot\vec{\nu}_\lambda| > 2^{n+3} - 2^n > 2^{n+2}$, so that $n_V \ge n+3$. On the other hand, if $n_V > n+3$, *i.e.* $n_V = n+m$, for some $m > 3$, we note that $|\vec{\omega}\cdot\vec{\nu}^0_\lambda| > \gamma_{n_V-1} - \gamma_n$, 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 $|\varepsilon_V| < \gamma_{n_V-3}$, $|\vec{\omega}\cdot\vec{\nu}^0_\lambda + \sigma_\lambda\varepsilon_V| \ge \gamma_{n_V-1} - \gamma_n - \gamma_{n_V-3} \ge (2^{n_V-2} - 2^{n_V-m}) - \gamma_{n_V-3} \ge (2^{n_V-3} + 2^{n_V-4} + \ldots + 2^{n_V-m+1}) - 2^{n_V-3}$ $\ge 2^{n_V-4}$; otherwise, if $n_V = n+3$, we note that $|\vec{\omega}\cdot\vec{\nu}^0_\lambda| > 2^{n+3}$, so that $|\vec{\omega}\cdot\vec{\nu}^0_\lambda + \sigma_\lambda\varepsilon_V| > 2^{n+3} - \gamma_{n_V-3}$ $\ge 2^{n_V-1}$, for $|\varepsilon_V| < \gamma_{n_V-3}$. Therefore we can conclude that, while ε_V varies in a complex disk of radius γ_{n_V-3} and center in 0, the quantity $|\vec{\omega}\cdot\vec{\nu}^0_\lambda + \sigma_\lambda\varepsilon_V|$ does not become smaller than 2^{n_V-4}. Note the main point here: the quantity γ_{n_V-3} will usually be $\gg \gamma_{n_{\lambda_V}}$ which is the value ε_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_λ the scale of the line λ in ϑ^1, each of the $\prod 2M_V \le e^{2k}$ 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 $\varepsilon_V = \vec{\omega}\cdot\vec{\nu}_{\lambda_V}$, for $|\varepsilon_V| \le \gamma_{n_V-3}$, associated with the resonant clusters V. This even holds if the ε_V are regarded as independent complex parameters.

By construction it is clear that the sum of the $\prod 2M_V \le e^{2k}$ terms, giving the contribution to $\vec{h}^{(k)}$ from the trees in $\mathcal{F}(\vartheta^1)$, vanishes to second order in the ε_V parameters (by the approximate cancellation discussed above). Hence by the maximum principle, and recalling that each of the scalar products in (8) can be bounded by N^2, we can bound the contribution from the family $\mathcal{F}(\vartheta^1)$ by:

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

where:

1) N_n is the number of propagators of scale n in ϑ^1 ($n=1$ does not appear as $|\vec{\omega}\cdot\vec{\nu}| \ge \gamma_0 \ge 2^2$, in such cases, and $2^8 \ge 2^4$);

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 scale $n_{\lambda_{V_i}} = n$ is holomorphic for $|\varepsilon_i| < \gamma_{n_i-3}$ if $n_i \equiv n_{V_i}$, provided $n_i \ge n+3$ (see above);

ii) the resummation says that the dependence on the ε_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 $(|\varepsilon_i|\,\gamma^{-1}_{n_i-3})^2 \le 2^{2(n-n_i+4)}$, if $n_i \ge n+3$ (of course the gain factor can be important only when $\ll 1$).

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:

$$\frac{1}{k!}\,(2^{2\tau}C_0^2J_0^{-1}f_0N^2)^k e^{2k}2^{8k}2^{8k}\prod_n 2^{-8nkE^{-1}2^{\varepsilon n}} \le \frac{1}{k!}B_0^k \tag{4.3}$$

with $B_0 = 2^{18}e^2(2^{2\tau}C_0^2 f_0 J_0^{-1})N^2 \exp[2^{3\tau^{-1}}(8N\log 2)\sum_{n=1}^{\infty} n2^{-n\tau^{-1}}]$. To sum over the trees we note that, fixed ϑ the collection of clusters is fixed. Therefore we only have to multiply (4.3) by the number of diagram shapes for ϑ, ($\le 2^{2k}k!$), by the number of ways of attaching mode labels, ($\le (3N)^{lk}$), so that we can bound $|h^{(k)}_{\vec\nu j}|$ by (1.5).

Appendix : Resonant Siegel-Brjuno bound.

Calling N^*_n the number of non resonant lines carrying a scale label $\le n$. We shall prove first that $N^*_n \le 2k(E2^{-\varepsilon n})^{-1} - 1$ if $N_n > 0$. We fix n and denote N^*_n as $N^*(\vartheta)$.

If ϑ has the root line with scale $> n$ then calling $\vartheta_1, \vartheta_2, \ldots, \vartheta_m$ the subdiagrams of ϑ emerging from the first vertex of ϑ and with $k_j > E\,2^{-\varepsilon n}$ lines, it is $N^*(\vartheta) = N^*(\vartheta_1) + \ldots + N^*(\vartheta_m)$ 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 is certainly the case if E is chosen as in (4.11)[4].

In the other case it is $N^*_n \le 1 + \sum_{i=1}^m N^*(\vartheta_i)$, and if $m = 0$ the statement is trivial, or if $m \ge 2$ 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 ϑ_1 is $k_1 > k - \frac{1}{2}E\,2^{-n\varepsilon}$. 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 ϑ_1 is either a resonant line or it has scale $> n$.

Accepting the last statement it will be: $N^*(\vartheta) = 1 + N^*(\vartheta_1) = 1 + N^*(\vartheta'_1) + \ldots + N^*(\vartheta'_{m'})$, with ϑ'_j being the m' subdiagrams emerging from the first node of ϑ'_1 with orders $k'_j > E\,2^{-\varepsilon n}$: this is so because the root line of ϑ_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 $N^*(\vartheta'_1) = N^*(\vartheta''_1) + \ldots + N^*(\vartheta''_{m''})$, *etc.*, until we reach a trivial case or a diagram of order $\le k - \frac{1}{2}E\,2^{-n\varepsilon}$.

It remains to check that if $k_1 > k - \frac{1}{2}E\,2^{-n\varepsilon}$ then the root line of ϑ_1 has scale $> n$, unless it is entering a resonance.

Suppose that the root line of ϑ_1 has scale $\le n$ and is not entering a resonance. Note that $|\vec\omega\cdot\vec\nu(v_0)| \le \gamma_n, |\vec\omega\cdot\vec\nu(v_1)| \le \gamma_n$, if v_0, v_1 are the first vertices of ϑ and ϑ_1 respectively. Hence $\delta \equiv |(\vec\omega\cdot(\vec\nu(v_0) - \vec\nu(v_1))| \le 2\,2^n$ and the diophantine assumption implies that $|\vec\nu(v_0) - \vec\nu(v_1)| > (2\,2^n)^{-\tau^{-1}}$, or $\vec\nu(v_0) = \vec\nu(v_1)$. The latter case being discarded as $k - k_1 < \frac{1}{2}E\,2^{-n\varepsilon}$ (and we are not considering the resonances: note also that in such case the lines in ϑ/ϑ_1 different from the root of ϑ must be inside a cluster), it follows that $k - k_1 < \frac{1}{2}E\,2^{-n\varepsilon}$ is inconsistent: it would in fact imply that $\vec\nu(v_0) - \vec\nu(v_1)$ is a sum of $k - k_1$ vertex modes and therefore $|\vec\nu(v_0) - \vec\nu(v_1)| < \frac{1}{2}NE\,2^{-n\varepsilon}$ 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 $p \le 2k\,(E2^{-\varepsilon n})^{-1} - 1$. In fact this is true for $k \le E2^{-\varepsilon n}$, (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 $p(\vartheta) = p(\vartheta_1) + \ldots + p(\vartheta_m)$, with the above notation, and the statement follows by induction. If v_0 is in a cluster of scale n we call $\vartheta_1, \ldots, \vartheta_m$ the subdiagrams emerging from the cluster containing v_0 and with orders $k_j > E2^{-\varepsilon n}$. It will be $p(\vartheta) = 1 + p(\vartheta_1) + \ldots + p(\vartheta_m)$. 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 $\le n - 1$. Therefore the cluster V must contain at least $E2^{-\varepsilon n}$ nodes. This means that $k_1 \le k - E2^{-\varepsilon n}$: thus (4.1) is proved.

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`, #93-171.

[4] Note that if $k \le E\,2^{-n\varepsilon}$ it is, for all momenta $\vec\nu$ of the lines, $|\vec\nu| \le NE\,2^{-n\varepsilon}$, *i.e.* $|\vec\omega\cdot\vec\nu| \ge (NE\,2^{-n\varepsilon})^{-\tau} = 2^3\,2^n$ so that there are *no* clusters T with $n_T = n$ and $N^* = 0$.

[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`, #93-172. 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`, #93-164.

[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 transormations 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.