KAM theorem revisited

G. Gentile[†], **V. Mastropietro**[††]

† Dipartimento di Fisica, Università di Roma "La Sapienza", 00185 Roma, Italia.
†† Dipartimento di Matematica, Università di Roma II "Tor Vergata", 00133 Roma, Italia.

Abstract. *A combinatorial proof of the KAM theorem for the Thirring model is presented, by using renormalization group techniques usual in the formalism of quantum field theory.*

Keywords: classical mechanics, KAM theorem, perturbation theory, euclidean quantum field theory, renormalization group

1. Introduction

In this paper we present a simple proof of the KAM theorem[1] by studying directly the convergence of the formal series (*Lindstedt series*, [L]) which expresses the quasiperiodic solution of the equations of motion. Our aim is to show that the methods of quantum field theory represent a very profitable approach to the problem, since they permit to single out in a very natural way the source of problems, and provide simultaneously the necessary tools how to solve them.

Albeit the methods we are going to describe can be applied in order to prove the (analytic) KAM theorem in the most general case, [GM], here we want to insist on the methods rather than on the result, so that it will be convenient to restrict ourselves in a particular case, in which no strictly unavoidable technical intricacy appears: in this way we will be able to present a rather elementary proof. The extension to the general case does not requires conceptually anything new, but of course everything is notationally more involved.

Let us consider the model obtained by perturbing through an action–independent function a system of rotators (*Thirring model*, [T]). The hamiltonian function is then

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

where J is the (non singular) matrix of the moments of inertia, $\vec{\alpha} = (\alpha_1, \ldots, \alpha_l) \in \mathbf{T}^l$ are the angles describing the positions of the rotators, $\vec{A} = (A_1, \ldots, A_l) \in \mathbf{R}^l$ are the conjugated actions, and we suppose that $\min_{j=1,\ldots,l} J_j \equiv J_0 > 0$, if J_j's are the eigenvalues of J. The interaction potential f is an even trigonometric polynomial of degree N

$$f(\vec{\alpha}) = \sum_{|\vec{\nu}| \le N} f_{\vec{\nu}} \cos \vec{\nu} \cdot \vec{\alpha} , \qquad f_{\vec{\nu}} = f_{-\vec{\nu}} , \tag{1.2}$$

and without loss of generality we can suppose $f_{\vec{0}} = 0$.

Our proof takes full advantage from the similarity between the Lindstedt series and the perturbative series in euclidean quantum field theory. Such similarity can be pursued to exhibith explicitly a quantum field theory model in which the one–point Schwinger function coincides with

[1] From Kolmogorov, [K], Arnol′d, [A], and Moser, [M].

the Lindstedt series for the model (1.1), see [G2]. In §2, we shall show that a proof of the KAM theorem can be performed within the usual framework of the renormalization group approach to quantum field theory, (we refer to [BG] for a rewiew of the results of this approach).

We conclude this introductory section by giving a formal statement of the result.

1.1. Theorem. *Given the hamiltonian model (1.1), with interaction potential (1.2), and fixed a "rotation vector" $\vec\omega_0$ satyisfying the* diophantine property $C_0|\vec\omega_0\cdot\vec\nu| \ge |\vec\nu|^{-\tau}$, $\vec 0 \ne \vec\nu \in \mathbf{Z}^l$ *with diophantine constants $C_0, \tau > 0$, there exist two functions $\vec H(\vec\omega_0 t;\varepsilon)$, $\vec h(\vec\omega_0 t;\varepsilon)$ analytic in $\vec\psi \equiv \vec\omega_0 t$ with* $\mathrm{Re}\vec\psi \in \mathbf{T}^l$, *and $|Im\vec\psi_j| < \xi$, and analytic in ε for $|\varepsilon| < \varepsilon_0$, where*

$$\varepsilon_0 = bJ_0^{-1}C_0^2\Big[\max_{0<|\vec\nu|\le N} f_{\vec\nu}\Big]N^{2+l}e^{cN}e^{\xi N} \; , \tag{1.3}$$

for some positive constant b depending only on N and l, such that the functions

$$\begin{aligned}\vec A(t) &= \vec A_0 + \vec H(\vec\omega_0 t;\varepsilon) \; , \\ \vec\alpha(t) &= \vec\omega_0 t + \vec h(\vec\omega_0 t;\varepsilon) \; ,\end{aligned} \tag{1.4}$$

where $\vec A_0 = J\vec\omega_0$, describe an invariant torus which is run quasi–periodically with angular velocity $\vec\omega_0$.

The *twistless property*, *i.e.* the uniformity of the convergence radius of the perturbative series on the twist rate $T = \max_{i=1,\dots,l} J_j$, is due to the form of the interaction, and can be obtained also via the classical proof, as a careful analysis of the latter could show.

2. Proof of the theorem

Calling $\vec H^{(k)}$ and $\vec h^{(k)}$ the k-th order coefficients of the Taylor expansion in powers of ε of the functions $\vec H$ and $\vec h$ in (1.4), and writing the equations of motion as

$$\frac{d\vec A}{dt} = -\varepsilon\partial_{\vec\alpha} f(\vec\alpha,\vec A) \; , \qquad \frac{d\vec\alpha}{dt} = J\vec A \; ,$$

we get immediately recursion relations for $\vec H^{(k)}, \vec h^{(k)}$; if $\vec\psi = \vec\omega_0 t$, one has for $k=1$

$$(\vec\omega_0\cdot\partial_{\vec\psi})\,\vec H^{(1)} = -\partial_{\vec\psi} f \; , \qquad \vec\omega_0\cdot\partial_{\vec\psi}\,\vec h^{(1)} = J\vec H^{(1)} \; ,$$

and for $k>1$

$$(\vec\omega_0\cdot\partial_{\vec\psi})^2\,\vec h^{(k)} = \sum_{(k-1)}{}^{*}(-\partial_{\vec\alpha})\sum_{p\ge0}\frac{1}{p!}\prod_{s=1}^{p}\Big(\vec h^{(k_s)}\cdot\partial_{\vec\alpha}\Big) f(\vec\omega_0 t) \; , \tag{2.1}$$

and $\vec H^{(k)}$ can be trivially expressed in terms of $\vec h^{(k)}$ by $\vec\omega_0\cdot\partial_{\vec\psi}\,\vec h^{(k)} = J\vec H^{(k)}$. In (2.1) the sum $\sum^*_{(k-1)}$ denotes summation over the integers $k_s \ge 1$ with $\sum_{s=1}^p k_s = k-1$, and the derivatives are supposed to apply to the functions $f(\vec\alpha)$, and then evaluated in $\vec\alpha = \vec\omega_0 t$. If $p=0$ the corresponding product is meant as 1. The formal solubility of the above equations follows easily by induction: in fact, if (1) $\vec h^{(k)}$ is odd in $\vec\psi$, and (2) $\vec H^{(k)}$ is even in $\vec\psi$ and with vanishing average, for $1 \le k < k_0$, then from the parity of the interaction the same holds for $\vec h^{(k_0)}$ and $\vec H^{(k_0)}$, so that (2.1) can be solved.

It can be convenient to write the above recursive formulae in the Fourier space. We easily find, for $k=1$ and $\vec\nu \ne \vec 0$,

$$\vec h^{(1)}_{\vec\nu} = \frac{-i\vec\nu\, J f_{\vec\nu}(\vec A_0)}{(i\vec\omega_0\cdot\vec\nu)^2} \; ,$$

and, for $k>1$ and $\vec\nu \ne \vec 0$,

$$(i\vec\omega_0\cdot\vec\nu)^2\,\vec h^{(k)}_{\vec\nu} = \sum_{(k-1)}{}^{*}(-i\vec\nu_0)\sum_{p\ge0}\frac{1}{p!}\prod_{s=1}^{p}\Big(i\vec\nu_0\cdot\vec h^{(k_s)}_{\vec\nu_s}\Big) f_{\vec\nu_0} \; ,$$

where $\sum_{(k-1)}^{*}$ denotes summation over the integers $k_s \geq 1$ with: $\sum_{s=1}^{p} k_s = k-1$, and over the integers $\vec{\nu}_0$, $\vec{\nu}_1$, …, $\vec{\nu}_p$, with: $\vec{\nu}_0 + \sum_{s=1}^{p} \vec{\nu}_s = \vec{\nu}$. Moreover $\vec{H}^{(k)}_{\vec{\nu}} = J^{-1}\vec{h}^{(k)}_{\vec{\nu}}(i\vec{\omega}_0 \cdot \vec{\nu})$ so that from now on we study simply $\vec{h}^{(k)}_{\vec{\nu}}$. The $\vec{0}$-th Fourier components $\vec{h}^{(k)}_{\vec{0}}$ and $\vec{H}^{(k)}_{\vec{0}}$ are identically vanishing for any $k \geq 1$, by the previously remarked parity properties.

It is possible to write $\vec{h}^{(k)}_{\vec{\nu}}$ as sum of contributions each of which can be graphically represented as a Feynman graph with no loops, *i.e.* a *tree graph* (or simply *tree*), which we call ϑ: it will consist of a family of k lines (*branches* or *lines*) arranged to connect a partially ordered set of points (*vertices* or *nodes*), with the higher vertices to the right. The branches are naturally ordered as well; all of them have two vertices at their extremes (possibly one of them is a top vertex), except the lowest or *first branch* which has only one vertex, the *first vertex* v_0 of the graph. The other extreme r of the first branch will be called the *root* of the graph and will not be regarded as a vertex; we shall call the first branch also *root branch*.

Each branch carries an arrow pointing from the vertex to the right v to the vertex to the left v' (*i.e.* directed toward the root): we say that the branch exits from v and enters v'. Note that there is a correspondence 1–to–1 between vertices and branches, if we associate to each vertex v the branch λ_v exiting from it. Note also that the direction of the arrows is opposite to the ordering of the vertices. A possible graph is represented in Fig. 2.1.

Let us define the *order* of a graph as the number of vertices of the graph. We introduce also the notion of *numbered trees*: they are obtained by imagining to have a deposit of k branches numbered from 1 to k and depositing them on the branches of a topological graph with order k.

A group $\mathcal{G}$ of transformations acts on the graphs, generated by the following operations: fix a node $v \in \vartheta$ and permute the subgraphs emerging from it. The graphs will be regarded as identical if superposable by the action of a transformation of the group $\mathcal{G}$, in such a way that all the numbers associated to the lines match. The number of numbered graphs of order k is bounded by $k!2^{2k}$.

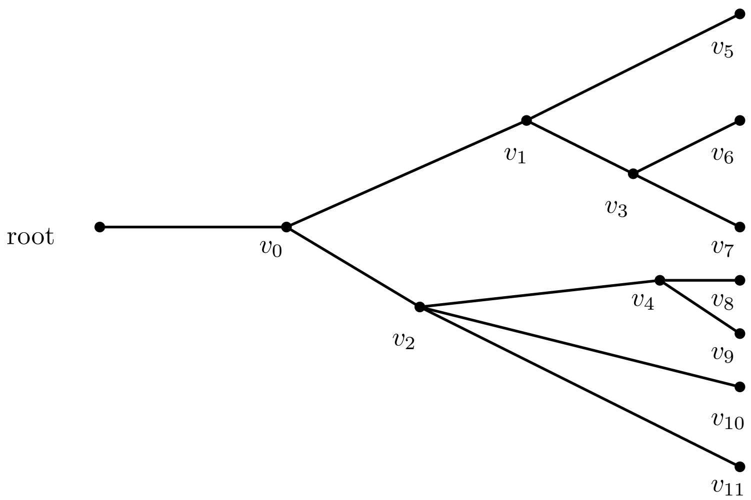

Fig.2.1. A graph ϑ with order $k = 12$.

One associates to each line λ a *momentum* $\vec{\nu}_\lambda$, a label $j_\lambda \in \{1, \ldots, l\}$, and a *propagator* $g_\lambda = [i\vec{\omega} \cdot \vec{\nu}_\lambda]^{-2}$and to each vertex v there corresponds a label $\vec{\nu}_v$ (*mode label*), such that $0 < |\vec{\nu}_v| \leq N$, and a *vertex factor*

$$\mathcal{E}_v = f_{\vec{\nu}_v}(-J^{-1}\vec{\nu}_v)_{j_{\lambda_v}} \prod_{q=1}^{m_v} (i\vec{\nu}_v)_{j_{\lambda_q}} \delta(\vec{\nu}_{\lambda_v} - \vec{\nu}_v - \sum_{q=1}^{m_v} \vec{\nu}_{\lambda_q}) , \tag{2.2}$$

where the product and the sum are over all the m_v lines entering v, and the δ is the Kronecker's delta assuring the momentum conservation, (*i.e.* it imposes that the momentum coming out from a vertex v is given by the sum of the momenta flowing in it plus the "momentum emitted" by the vertex itself, namely the mode $\vec{\nu}_v$).

Let us call $\mathcal{T}_k$ the collection of *labeled graphs*, *i.e.* of decorated numbered graphs, of order k, where the decoration is given by a set of labels associated to the lines λ's, $(\vec{\nu}_\lambda,\ j_\lambda)$, and to the vertices v's, $(\vec{\nu}_v,\ \mathcal{E}_v)$. The *value* of a graph ϑ is given by

$$X(\vartheta) = \prod_{v\in\vartheta} g_{\lambda_v}\mathcal{E}_v \ , \tag{2.3}$$

and the k-th order coefficient of $h_{j\vec{\nu}}$ can be written as $h^{(k)}_{j\vec{\nu}} = \sum^*_{\vartheta\in\mathcal{T}_k} \frac{1}{k!}X(\vartheta)$, where the sum is over all the labeled graphs in $\mathcal{T}_k$, and the $*$ recalls that the graph labels have to be consistent, (*i.e.* the momentum conservation has to be satisfied), and the first branch λ has to have $\vec{\nu}_\lambda = \vec{\nu}$ and $j_\lambda = j$.

The representation of the Lindstedt series in terms of graphs was introduced in [E1] and the analogy with the formal Feynman graphs expansion in field theory was pointed out in [FT].

One of the main difficulties of proving the convergence of $\sum_{k=1}^\infty h^{(k)}_{j\vec{\nu}}$ is that one can easily exhibit graph values of order $O((k!)^\beta)$ for some suitable positive constant β, if k is the perturbative order, so that in order to prove that the Lindstedt series converges one has to show that there is a sharp compensation between these "huge" graph values. The first proof of the convergence of the Lindstedt series was provided by Eliasson, [E1], in the general analytic case; his work, quite difficult to read, has not enjoyed a wide circulation and understanding, (the only printed works being [E2] and [E3], and the fundamental paper [E1] being only an internal report of the University of Stockholm). Eliasson did not perform explicit cancellations on individual terms of the series, but he showed that they must occur because of the symplectic nature of the problem. A very clean proof of the convergence of the model (1.1) was given in [G1] (with the additional condition of the *strong diophantine property*, see §3), showing explicitly some approximate cancellations between terms of the Lindstedt series; as the present work has the same purposes of [G1], and it should be considered an improvement of it, we discuss the main differences at the end, when the reader can appreciate them.[2]

Like in the renormalization group approach to the analysis of the perturbation expansions of renormalizable quantum field theory, [BG], it can be convenient to introduce a multiscale decomposition of the propagators in the following way. Let us define

$$\chi_n(x) = \theta(|x| - 2^{n-1}) - \theta(|x| - 2^n) \ , \quad n \le 0 \ , \qquad \chi_1(x) = \theta(|x| - 1) \ , \tag{2.4}$$

where $\theta(x)$ is the Heaviside function. Such functions realize a sharp partition of unity so that we can decompose the propagator

$$g_\lambda = \frac{1}{[\vec{\omega}\cdot\vec{\nu}_\lambda]^2} = \sum_{n=-\infty}^{1} \frac{\chi_n(\vec{\omega}\cdot\vec{\nu}_\lambda)}{[\vec{\omega}\cdot\vec{\nu}_\lambda]^2} \equiv \sum_{n=-\infty}^{1} g^{(n)}_\lambda \ , \tag{2.5}$$

where $g^{(n)}_\lambda$ is the "propagator at scale n". If $n \le 0$, $g^{(n)}_\lambda$ is a compact support function different from 0 for $2^{n-1} \le |\vec{\omega}\cdot\vec{\nu}_\lambda| \le 2^n$, while $g^{(1)}_\lambda$ has support for $1 \le |\vec{\omega}\cdot\vec{\nu}_\lambda|$. Then one associates to each line λ an extra label n_λ (*scale label*). We associate to the lines of the labeled graphs $\mathcal{T}_k$ also the scale label n_λ and the value of a graph ϑ is now given by $X(\vartheta) = \prod_{v\in\vartheta} g^{(n_{\lambda_v})}_{\lambda_v}\mathcal{E}_v$.

Looking at the scale labels we identify the connected clusters T of vertices which are linked by a continuous path of lines with the same scale label n_T or a higher one and which are maximal: we shall say that "the cluster T has scale n_T". Therefore an inclusion relation is established between the clusters, in such a way that the innermost clusters are the clusters with highest scale, and so on. Let us denote by $K(T)$ the *order* of the cluster T, *i.e.* number of vertices contained in T. Then, given a graph ϑ, let T be the maximal cluster contained in no other clusters, (*i.e.* the cluster containing all the graph), and T_0 the collection of lines and vertices contained inside T but not in any subcluster $T' \subset T$. We can write the graph value $X(\vartheta)$ recursively as

$$X(\vartheta) \equiv X(T) = \prod_{\lambda\in T_0} g^{(n_T)}_\lambda \prod_{T'\subset T} X(T') \prod_{v\in T_0} \mathcal{E}_v \ , \tag{2.6}$$

[2] In [CF] a sligth generalization of the model (1.1) is studied, with techniques remarkably close to [G1].

where $T' \subset T$ means that the cluster T' is a maximal cluster inside T.

2.1. Definition. *Among the clusters we consider the ones with the property that there is only one incoming line, carrying the same momentum of the outgoing line, and we define them* resonances. *If V is one such cluster we denote by λ_V the incoming line: we call n_{λ_V} the* resonance-scale *(which is different from the scale n_V of the resonance V as a cluster) and λ_V a* resonant line.

Given a resonance V with incoming momentum $\vec{\nu}$, we write $X(V) = \mathcal{V}(\vec{\omega} \cdot \vec{\nu}_{\lambda_V})$ in order to show explicitly the dependence on the (sole) momentum of the external lines of the resonance. Note that, if V is a resonance, $\sum_{w \in V} \vec{\nu}_w = 0$ for the momentum conservation (see (2.2))

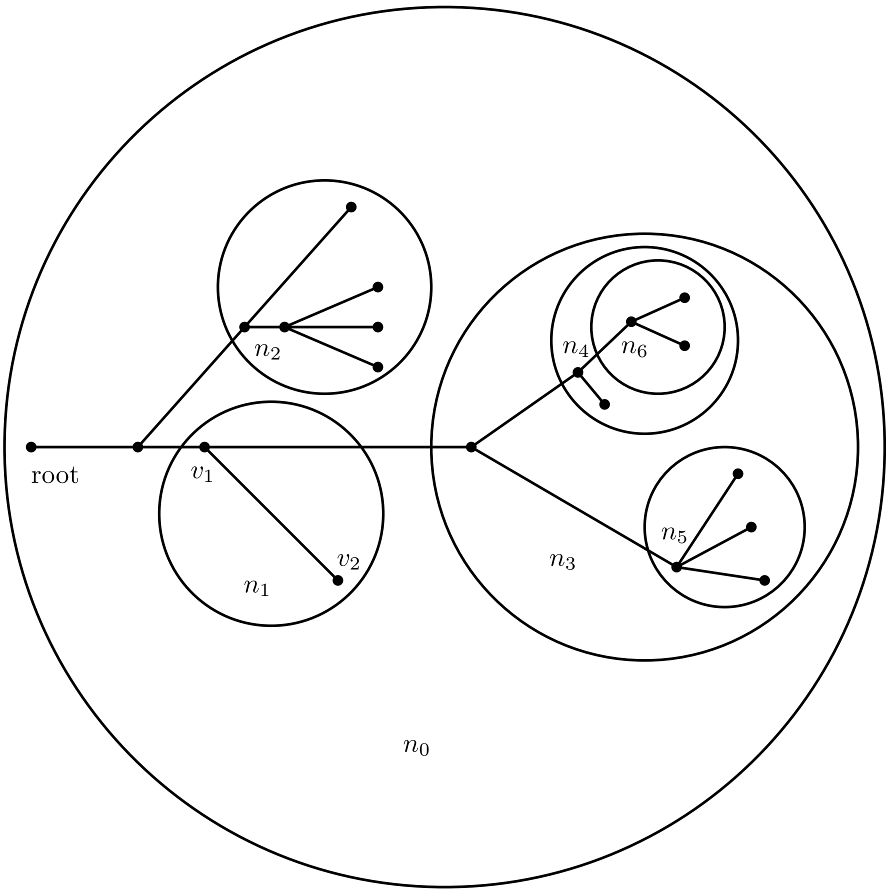

Fig.2.2 A graph with scale labels associated to the lines and a cluster structure.

Given a graph ϑ, let us define $N_n(\vartheta)$ the number of lines in ϑ with scale $n \leq 0$. Then, as we can deduce from (2.3), the graph value $X(\vartheta)$ admits the bound

$$\mathcal{C}^k \prod_{n \leq 0} 2^{-2nN_n(\vartheta)} , \tag{2.7}$$

where $\mathcal{C} = 2^2 J_m^{-1} C_0^2 N^2 [\max_{\vec{\nu}} f_{\vec{\nu}}]$, and $N_n(\vartheta)$ verifies the bound (*Siegel-Bryuno's lemma*, [S,B])

$$N_n(\vartheta) \leq \frac{4k}{E 2^{-n\tau^{-1}}} + \sum_{T, n_T = n} [-1 + m_T(\vartheta)] , \tag{2.8}$$

where $m_T(\vartheta)$ is the number of resonances V inside the cluster T of ϑ, with resonance-scale $n_{\lambda_V} = n_T$, and E can be chosen $E = 2^{-3\tau^{-1}} N^{-1}$. The proof of the above bound is in the Appendix.

Then, if there were no resonances, we would obtain a convergent bound. The resonances are then the really dangerous contributions and we have to deal with them. Note that our definition of resonances is different from the one in [E1] or in [CF] in which the resonances are simply subgraphs among two lines carring the same momentum: in this way the well known problem in quantum field theory of the *overlapping divergences* appear and one has to eliminate it by distinguishing between resonances really "dangerous" or not. The introduction of clusters "automatically" leads to isolate the dangerous contributions in the perturbative expansion and the problem of the overlapping divergences never appear (the multiscale decomposition was in fact created in field theory just to solve this problem). We stress moreover that the Siegel-Bryuno's bound is the only point in the proof in which the diophantine condition has a role; from now on, *i.e.* in the study of the resonances, the diophantine condition will never be used.

Proceeding as in quantum field theory we introduce a *localization operator*. If $X(T)$ is the value of a cluster T, the action of the localization operator $\mathcal{L}$ on T is defined as follows. If T is not a resonance, $\mathcal{L}X(T) = 0$, while, if T is a resonance, $T = V$, we set

$$\mathcal{L}X(V) \equiv \mathcal{L}\mathcal{V}(\vec{\omega} \cdot \vec{\nu}_{\lambda_V}) = \mathcal{V}(0) + (\vec{\omega} \cdot \vec{\nu}_{\lambda_V})\dot{\mathcal{V}}(0) \ , \tag{2.9}$$

where $\dot{\mathcal{V}}(0)$ denotes the first derivative of $\mathcal{V}(\vec{\omega} \cdot \vec{\nu}_{\lambda_V})$ with respect to its argument, computed in $\vec{\omega} \cdot \vec{\nu}_{\lambda_V} = 0$. Then we split each cluster value as $X(T) = \mathcal{L}X(T) + \mathcal{R}X(T)$, with $\mathcal{R} = 1\!\!1 - \mathcal{L}$.

2.2. LEMMA. *The sum over graphs containing resonances on which $\mathcal{L}$ applies is vanishing.*

2.3 *Proof of Lemma 2.2.* We consider all the graphs we obtain by detaching from each resonance the subgraph with first vertex w_2, if w_2 is the vertex from which the incoming line of the resonance (resonant line) exits, then reattaching it to all the vertices $w \in V$. We call this set of contributions *resonance family*. If $\vec{\omega} \cdot \vec{\nu}_{\lambda_V} = 0$, no propagator changes, and the only effect of our operation is that one of the factors appearing in the product $\prod_{w \in V} \mathcal{E}_w$ changes by taking successively the values $(i\vec{\nu}_w)_{j_{\lambda_v}}$, $w \in V$. In this way we build a quantity proportional to $\sum_{w \in V} (i\vec{\nu}_w)_{j_{\lambda_w}} = \vec{0}$. If we sum also on an overall change of sign of the $\vec{\nu}_w$'s, and we take into account the parity of the propagator and (2.2), we obtain a second order zero. ∎

It is convenient to write the effect of $\mathcal{R}$ on a resonance V as

$$\mathcal{R}\mathcal{V}(\vec{\omega} \cdot \vec{\nu}) = (\vec{\omega} \cdot \vec{\nu})^2 \int_0^1 dt \; t\ddot{\mathcal{V}}(t\vec{\omega} \cdot \vec{\nu}) \ , \tag{2.10}$$

where $\ddot{\mathcal{V}}$ denotes the second derivative. As there are resonances enclosed in other resonances the above formula can suggest that there are propagators derived up to $\approx k$ times, if k is the order of the graph.

However this is not the case as the propagators are derived at most two times. Let be n the resonance-scale of the maximal resonance V, and let us write $\mathcal{R}X(V)$ as

$$\mathcal{R}\Big(\prod_{\lambda \in V_0} g_\lambda^{(n_V)} \prod_{T' \in V} \mathcal{R}X(T') \prod_{v \in V_0} \mathcal{E}_v \Big) \ , \tag{2.11}$$

where, for any resonance $\tilde{V} \subseteq V$, $\mathcal{R}X(\tilde{V})$ can be written either as in (2.10), or as a difference $\mathcal{R}X(\tilde{V}) = X(\tilde{V}) - \mathcal{L}X(\tilde{V})$, in according to which expression turns out to be more convenient to deal with.

Then the first step is to write the action of $\mathcal{R}$ on the maximal cluster as in (2.10), leaving the other terms $\mathcal{R}X(T')$ written as differences when T' are resonances: so (2.11) can be expressed by the Leibniz's rule as a sum of terms, and the derivatives of $\mathcal{R}$ apply either on some propagator $g_\lambda^{(n_V)}$ or on some $\mathcal{R}X(T')$. In the end there can be either no derivative, or one derivative, or two derivatives applied on each $\mathcal{R}\mathcal{V}(T')$. When T' is not a resonance, $\mathcal{R} = 1$, and trivially, $\partial^p \mathcal{R}X(T') \equiv \partial^p X(T')$, for $0 \leq p \leq 2$.

When T' is a resonance, $T' = V'$, if only one derivative acts on $\mathcal{R}X(V') = \mathcal{R}\mathcal{V}(\vec{\omega} \cdot \vec{\nu})$, then we write $\partial \mathcal{R}\mathcal{V}(\vec{\omega} \cdot \vec{\nu}) = \partial \mathcal{V}(\vec{\omega} \cdot \vec{\nu}) - \dot{\mathcal{V}}(0) = (\vec{\omega} \cdot \vec{\nu}) \int_0^1 dt \ddot{\mathcal{V}}(t\vec{\omega} \cdot \vec{\nu})$, while, if two derivatives act on

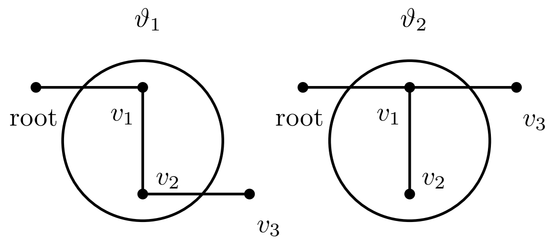

Fig.2.3. The simplest resonance family.

$\mathcal{R}X(V')$, then we write $\partial^2\mathcal{R}\mathcal{V}(\vec\omega\cdot\vec\nu)=\ddot{\mathcal{V}}(\vec\omega\cdot\vec\nu)$. Then two derivatives act on each resonance V' in any case (*i.e.* also if it is enclosed in another resonance), and the procedure can be iterated, since the resonances V' can be dealt with as the resonance V. Let us consider the effect of a derivative on a propagator $\dot g_\lambda^{(n)}=-2[\vec\omega\cdot\vec\nu_\lambda]^{-3}\chi_n(\vec\omega\cdot\vec\nu_\lambda)+[\vec\omega\cdot\vec\nu_\lambda]^{-2}\dot\chi_n(\vec\omega\cdot\vec\nu_\lambda)$, (where, as usual, the dot denotes derivative with respect to the argument). We write then the resonances as sum of two kinds of terms: the terms of the first kind contain no factors $\dot\chi_n$, while in the other ones derivatives of $\dot\chi_n$ appear. In the latter contributions the action of $\mathcal{R}$ produces a gain factor $2^{n_{\lambda_V}-n'}2^{n_{\lambda_V}-n''}$, where n' and n'' are the scales of two lines λ' and λ'', possibly coinciding, contained in some clusters T' and T'' inside V.

The terms in which there is at least a factor $\dot\chi_n$ could be sources of troubles, as

$$\dot\chi_n(x)=\delta(|x|-2^{n-1})-\delta(|x|-2^n)\ .$$

Hovewer this is not the case. In fact, let us consider the sum of all the graphs obtained replacing the resonance V we are considering with any other resonance with the same scale n_{λ_V} and with any choice of the scale labels n_λ of the lines λ inside the resonance with the only condition that $n_\lambda\le n_V$ and with any choice of the labels $\mathcal{L}$ or $\mathcal{R}$. If we sum all the contributions coming from all these resonances in which the χ_n function of a particular line is derived, the deltas cancel each other except the one coming from the case in which the line we are considering has frequency $n_V=n_{\lambda_V}+1$, but in this case $2^{n_{\lambda_V}-n_V}=2$.

In conclusion: (1) no more than two derivatives can ever act on any propagators; (2) a gain $2^{2(n_{\lambda_V}-n_V)}$ is obtained for any resonance V; (3) the total number of terms generated by the derivation is bounded by k^2.

Therefore (2.7) can be replaced with

$$\mathcal{C}_0^k\prod_{n\le0}2^{-2nN_n(\vartheta)}\prod_{V\in\vartheta}2^{2(n_{\lambda_V}-n_V)}\le\mathcal{C}_0^k\exp\Big[-k(8E^{-1}\log2)\sum_{n=-\infty}^{0}n2^{n/\tau}\Big]\ ,$$

where $\mathcal{C}_0=e^2\mathcal{C}$, as $k^2\le e^{2k}$, the last product in the left hand side is over the resonances in ϑ, $\prod_{V\in\vartheta}2^{2n_{\lambda_V}}\prod_n\prod_{T,n_T=n}2^{-2nm_T}=1$ and $\prod_{V\in\vartheta}2^{-2n_V}\prod_n\prod_{T,n_T=n}2^{2n}<1$ as it immediately follows from the definitions. Since the sum over the graphs is bounded by $2^{2k}N^{2k}2^kl^k(2N+1)^k$, this completes the proof of Theorem 1.1.

3. Conclusive comments

Finally we discuss the differences of our work with respect to [G1], where the gain factor $2^{2(n_{\lambda_V}-n_V)}$ was obtained by summing over all the resonances and using the maximum principle (in the form of the Schwarz's lemma) for analytic functions. Then one has to prove that the resonances are analytic as function of the scalar product of $\vec\omega$ times the momentum flowing in the resonance, and this property, of course, is not true if the resonances are defined like in our paper. This is the reason why in [G1] the resonances were defined with the further condition that number of lines inside each resonance V was bounded by $N^{-1}2^{-(n_{\lambda_V}+3)/\tau}$ and one assumed moreover that $\vec\omega_0$ obeyed to the *strong diophantine property* $\min_{0\geq p\geq n} |C_0|\vec{\omega}\cdot\vec{\nu}| - 2^{-p}| > 2^{-(n+1)}$, for $0 < |\vec{\nu}| \leq (2^{n+3})^{-\tau^{-1}}$; this assumption was eliminated in [GG] performing a particular partition of the unity depending on $\vec{\omega}_0$.

With our method we have to prove simply that the resonances are twice differentiable, and this allows us to give a simpler definition by requiring that the resonances are the clusters having only one incoming line (see Definition 2.1). The deep reason of this simplification is that there is an overcompensation, *i.e.* there are more cancellations than it would be necessary in order to make the series convergent.

We think that our method, very close to quantum field theory, should be useful to clarify some open problems in the KAM theory.

Aknowledgements. We thank G. Gallavotti for having introduced us to KAM theory and quantum field theory, and for many enlightening discussions.

Appendix. Resonant Siegel-Bryuno's bound.

We prove by induction on the graph order that, if $N_n^*(\vartheta)$ is defined as the number of non resonant lines in ϑ carring a scale label $\le n$, then $N_n^*(\vartheta) \le 2M(\vartheta)2^{(n+2)/\tau} - 1$, if $N_n(\vartheta) \ne 0$.

Let ϑ be a graph of order k. If ϑ has the root line with scale $> n$ then calling $\vartheta_1, \vartheta_2, \ldots, \vartheta_m$ the subgraphs of ϑ emerging from the first vertex of ϑ and with $M(\vartheta_j) > 2^{-(n+2)/\tau}$ lines, it is $N_n^*(\vartheta) = N_n^*(\vartheta_1) + \ldots + N_n^*(\vartheta_m)$ and the statement is inductively implied from its validity for $k' < k$ provided it is true that $N_n^*(\vartheta) = 0$ if $M(\vartheta) < 2^{-(n+2)/\tau}$, which is is certainly the case.

In the other case, (*i.e.* if the root branch has scale label $\le n$), it is $N_n^*(\vartheta) \le 1 + \sum_{i=1}^m N_n^*(\vartheta_i)$, 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 it is $M(\vartheta_1) > M(\vartheta) - 2^{-1}2^{-(n+2)/\tau}$. But in the latter case, it turns out that the root line of ϑ_1 is either a resonant line or it has scale $> n$.

Accepting the last statement we have: $N_n^*(\vartheta) = 1 + N_n^*(\vartheta_1) = 1 + N_n^*(\vartheta'_1) + \ldots + N_n^*(\vartheta'_{m'})$, with ϑ'_j's being the m' subgraphs emerging from the first vertex of ϑ'_1 with $M(\vartheta'_j) > 2^{-(n+2)/\tau}$: this is so because the root line of ϑ_1 will not contribute its unit to $N_n^*(\vartheta_1)$. Going once more through the analysis the only non trivial case is if $m' = 1$ and in that case $N_n^*(\vartheta'_1) = N_n^*({\vartheta''}_1) + \ldots + N_n^*({\vartheta''}_{m''})$, *etc.*, until we reach either a trivial case or a graph $\tilde{\vartheta}$ such that $M(\tilde{\vartheta}) < M(\vartheta) - 2^{-1}2^{-n/\tau}$.

It remains to check that, if $M(\vartheta_1) > M(\vartheta) - 2^{-1}2^{-(n+2)/\tau}$, 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 $|\vec{\omega} \cdot \vec{\nu}_{\lambda_{v_0}}| \le 2^{n+1}$, $|\vec{\omega} \cdot \vec{\nu}_{\lambda_{v_1}}| \le 2^{n+1}$, if v_0 and v_1 are the first vertices of ϑ and ϑ_1 respectively. Hence $\delta \equiv |(\vec{\omega} \cdot (\vec{\nu}_{\lambda_{v_0}} - \vec{\nu}_{\lambda_{v_1}})| \le 2\,2^{n+1}$ and the diophantine assumption implies that $|\vec{\nu}_{\lambda_{v_0}} - \vec{\nu}_{\lambda_{v_1}}| > (2^{n+2})^{-1/\tau}$, or $\vec{\nu}_{\lambda_{v_0}} = \vec{\nu}_{\lambda_{v_1}}$. The latter case being discarded as we are not considering the resonances,[3] it follows that $M(\vartheta) - M(\vartheta_1) < 2^{-1}\,2^{-(n+2)/\tau}$ is inconsistent: it would in fact imply that $\vec{\nu}_{\lambda_{v_0}} - \vec{\nu}_{\lambda_{v_1}}$ is a sum of $k - k_1$ vertex modes such that $|\vec{\nu}_{\lambda_{v_0}} - \vec{\nu}_{\lambda_{v_1}}| < 2^{-1}2^{-(n+2)/\tau}$, hence $\delta > 2^{n+3}$ which is contradictory with the above opposite inequality.

Analogously, we can prove that, if $N_n(\vartheta) > 0$, then the number $p_n(\vartheta)$ of clusters of scale n verifies the bound $p_n(\vartheta) \le 2M(\vartheta)2^{(n+2)/\tau}) - 1$. In fact this is true for a graph ϑ such that $M(\vartheta) \le 2^{(n+2)/\tau}$. Otherwise, if the first graph vertex v_0 is not in a cluster of scale n, it is $p_n(\vartheta) = p(\vartheta_1) + \ldots + p_n(\vartheta_m)$, with the above notation, and the statement follows by induction. If v_0 is in a cluster on scale n we call $\tilde{\vartheta}_1, \ldots, \tilde{\vartheta}_m$ the subdiagrams emerging from the cluster containing v_0 and such that $M(\vartheta_j) > 2^{-(n+2)/\tau}$, $j = 1, \ldots, m$: it will be $p_n(\vartheta) = 1 + p(\tilde{\vartheta}_1) + \ldots + p_n(\tilde{\vartheta}_m)$. Again we can assume $m = 1$, the other cases being trivial. But in such a case there will be only one branch entering the cluster T on scale n containing v_0 and it will have a momentum of scale $n' \le n - 1$. Therefore the cluster T must contain vertices such that at least $\sum_{v \in T} |\vec{\nu}_v| > 2^{-(n+2)/\tau}$ vertices, (otherwise, if λ is a line on scale n contained in T, and $\vec{\nu}_\lambda^0$ is the sum of the mode labels corresponding to the vertices following v_0 but inside T, we would have $|\vec{\omega} \cdot \vec{\nu}_\lambda| \le 2^{n+1}$ and, simultaneously, $|\vec{\omega} \cdot \vec{\nu}_\lambda| \ge 2^{n+3} - 2^{n-1} > 2^{n+2}$, which would lead to a contradiction). This means that $M(\vartheta_1) \le M(\vartheta) - 2^{-(n+2)/\tau}$.

From the above proven results, (2.8) follows, if we note that $\sum_{T, n_T = n} 1 = p_n(\vartheta)$.

[3] Note that $M(\vartheta) - M(\vartheta_1) < 2{-}12^{-(n+2)/\tau}$ implies that $|\vec{\omega} \cdot \lambda| \ge 2^{1+(n+2)/\tau}$ for all the lines λ preceding the root line and not contained in ϑ_1, so that the set composed by such lines, if $\vec{\nu}_{\lambda_{v_0}} = \vec{\nu}_{\lambda_{v_1}}$, is a resonance on scale $n' > n$.

References

[A] V.I. Arnol′d: Proof of a A.N. Kolmogorov's theorem on conservation of conditionally periodic motions under small perturbations of the hamiltonian, *Uspekhi Mat. Nauk* **18**, No. 5, 13–40 (1963); english translation in *Russ. Math. Surv.* **18**, No. 5, 9–36 (1963).

[BG] G.Benfatto G.Gallavotti: *Renormalization group*, Princeton University Press, Princeton (1995).

[B] A.D. Bryuno: The analytic form of differential equations I, *Trans. Moscow Math. Soc.* **25**, 131–288 (1971); II, *Trans. Moscow Math. Soc.* **26**, 199–239 (1972).

[CF1] L. Chierchia, C. Falcolini: A direct proof of a theorem by Kolmogorov in hamiltonian systems, *Ann. Scuola Norm. Super. Pisa Cl. Sci. Ser. IV* **21**, No. 4, 541–593 (1995).

[E1] L.H. Eliasson: Absolutely convergent series expansions for quasi-periodic motions, Report 2–88, 1–31, Department of Mathematics, University of Stockholm (1988).

[E2] L.H. Eliasson: Hamiltonian systems with linear normal form near an invariant torus, in *Nonlinear dynamics*, Ed. G. Turchetti, Bologna Conference, 30/5 to 3/6 1988, World Scientific, Singapore (1989).

[E3] L.H. Eliasson: Generalization of an estimate of small divisors by Siegel, in *Analysis, et cetera*, Ed. E. Zehnder & P. Rabinowitz, book in honor of J. Moser, Academic press (1990).

[FT] J. Feldman, E. Trubowitz: Renormalization on Classical Mechanics and Many Body Quantum Field Theory, *J. Anal. Math.* **58**, 213–247 (1992).

[G1] G. Gallavotti: Twistless KAM tori, *Comm. Math. Phys.* **164**, 145–156 (1994).

[G2] G. Gallavotti: Perturbation Theory, in *Mathematical physics towards the XXI century*, 275–294, Ed. R. Sen & A. Gersten, Ben Gurion University Press, Ber Sheva (1994).

[GG] G. Gallavotti, G. Gentile: Majorant series convergence for twistless KAM tori, *Ergodic Theory Dynam. Systems* , to appear (1995).

[GM] G. Gentile, V. Mastropietro: Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications, *Rev. Math. Phys.*, to appear (1995).

[K] A.N. Kolmogorov: On the preservation of conditionally periodic motions, *Dokl. Akad. Nauk* **96**, 527–530 (1954); english translation in G. Casati, J. Ford: *Stochastic behavior in classical and quantum hamiltonian*, Lectures Notes in Physics **93**, Springer, Berlin (1979).

[L] A. Lindstedt: *Mémoires de l'Académique de Saint-Pétersbourg* (1882), quoted in H. Poincaré: *Les méthodes nouvelles de la mécanique céleste*, Vol.II, Gauthier-Villars, Paris (1893).

[M] J. Moser: On invariant curves of an area preserving mapping of the annulus, *Nachr. Akad. Wiss. Göttingen Math.–Phys. Kl. II*, 1–20 (1962).

[S] K. Siegel: Iterations of analytic functions, *Ann. of Math.*, **43**, No. 4, 607–612 (1942).

[T] W. Thirring: *Course in Mathematical Physics*, Vol. 1, Springer, New York (1978).