Commun. Math. Phys. 164, 145–156 (1994)

Communications in
**Mathematical
Physics**

Twistless KAM Tori

Giovanni Gallavotti [1]

CNRS-CPT Luminy, case 907, F-13288 Marseille, France

Received: 21 June 1993/in revised form: 1 November 1993

Abstract: A selfcontained proof of the KAM theorem in the Thirring model is discussed.

I shall particularize the Eliasson method, [E], for KAM tori to a special model, of great interest, whose relevance for the KAM problem was pointed out by Thirring, [T] (see [G] for a short discussion of the model). The idea of exposing Eliasson's method through simple particular cases appears in [V], where results of the type of the ones discussed here, and more general ones, are announced.

The connection between the methods of [E] and the tree expansions in the renormalization group approaches to quantum field theory and many body theory can be found also in [G]. The connection between the tree expansions and the breakdown of invariant tori is discussed in [PV].

The Thirring model is a system of rotators interacting via a potential. It is described by the hamiltonian (see [G] for a motivation of the name):

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

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 nonsingular; but we only suppose that $\min\limits_{j=1,\ldots,l} J_j = J_0 > 0$, 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 the name "twistless": this is not a contradiction with the necessity of a twist rate in the general problems, see problems 1, 16, 17 in Sect. 5.11 of [G2], and [G]). We suppose f to be an even

Archived in mp_arc@math.utexas.edu#93-172; to get a TeX version, send an empty E-mail message

[1] E-mail: gallavotti%40221.hepnet@lbl.gov; permanent address: Dipartimento di Fisica, Università di Roma, "La Sapienza," P. Moro 2, I-00185 Roma, Italy

trigonometric polynomial of degree N:

$$f(\vec{\alpha}) = \sum_{0<|\vec{\nu}|\le N} f_{\vec{\nu}} \cos \vec{\nu}\cdot\vec{\alpha}\,, \qquad f_{\vec{\nu}} = f_{-\vec{\nu}}\,. \tag{2}$$

We shall consider a "rotation vector" $\vec{\omega}_0 = (\omega_1, \ldots, \omega_l) \in R^l$ verifying a *strong diophantine property* (see, however, the final comments) with dophantine constants C_0, τ, γ, c; this means that:

$$\begin{aligned} &1) \quad C_0|\vec{\omega}_0\cdot\vec{\nu}| \ge |\vec{\nu}|^{-\tau}\,, \qquad \vec{0} \ne \vec{\nu} \in Z^l\,,\\ &2) \quad \min_{0\ge p\ge n} |C_0|\vec{\omega}_0\cdot\vec{\nu}| - \gamma^p| > \gamma^{n+1} \quad \text{if} \quad n \le 0,\, 0<|\vec{\nu}| \le (\gamma^{n+c})^{-\tau^{-1}}\,, \end{aligned} \tag{3}$$

and it is easy to see that the *strongly diophantine vectors* have full measure in R^l if $\gamma > 1$ and c are fixed and if τ is fixed $\tau > l-1$: we take $\gamma = 2$, $c = 3$ for simplicity; note that 2) is empty if $n > -3$ or $p < n+3$. We shall set $\vec{A}_0 = J\vec{\omega}_0$. A special example can be the model $f_0(\vec{\alpha}) = J_0\vec{\omega}_0^2(\cos\alpha_1 + \cos(\alpha_1+\alpha_2))$.

We look for an ε-analytic family of motions starting at $\vec{\alpha} = \vec{0}$ and having the form:

$$\vec{A} = \vec{A}_0 + \vec{H}(\vec{\omega}_0 t;\varepsilon)\,, \qquad \vec{\alpha} = \vec{\omega}_0 t + \vec{h}(\vec{\omega}_0 t;\varepsilon) \tag{4}$$

with $\vec{H}(\vec{\psi};\varepsilon)$, $\vec{h}(\vec{\psi};\varepsilon)$ analytic in $\vec{\psi} \in T^l$ and in ε close to 0. *We shall prove that such functions exist and are analytic for $|\operatorname{Im}\psi_j| < \xi$ for $|\varepsilon| < \varepsilon_0$ with:*

$$\varepsilon_0^{-1} = bJ_0^{-1}C_0^2 f_0 N^{2+l}e^{cN}e^{\xi N}\,, \tag{5}$$

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 Eq. (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 $\vec{A} = \vec{A}_0$, $\vec{\alpha} = \vec{\psi} \in T^l$. One recognizes a version of the* KAM *theorem. The proof that follows simplifies the one reported in* [*G*].

Supposing $J_0 \equiv J_1 < J_2$ the uniformity in J_2 (i.e. what we call the *twistless* property) implies that the same ε_0 can be used as an estimate of the radius of convergence in ε of the power series describing the KAM tori with rotation vector $\vec{\omega}_0 = (\omega_1, \omega_2)$ in the system $(2J_1^{-1})A_1^2 + \omega_2 A_2 + \varepsilon f_0(\alpha_1, \alpha_2)$, which is one of the most studied hamiltonian systems. The estimate can be improved. Note that a careful analysis of the proof of the KAM theorem also shows the uniformity in the twist rate in the case Eq. (1).

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 $\vec{\omega}_0 \cdot \vec{\partial} h_j^{(k)} = J_j^{-1} H_j^{(k)}$ and, for $k > 1$:

$$\vec{\omega}_0 \cdot \vec{\partial} H_j^{(k)} = -\sum_{\substack{m_1,\ldots,m_l\\ |\vec{m}|>0}} \frac{1}{\prod_{s=1}^{l} m_s!}\,\partial_{\alpha_j}\partial^{m_1+\ldots+m_l}_{\alpha_1^{m_1}\ldots\alpha_l^{m_l}} f(\vec{\omega}_0 t)\cdot \sum{}^{*} \prod_{s=1}^{l}\prod_{j=1}^{m_s} h_s^{(k_j^s)}(\vec{\omega}_0 t)\,, \tag{6}$$

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 $k = 1$ one easily finds:

$$\vec{h}^{(1)}(\vec{\psi}) = -\sum_{\vec{\nu}\neq\vec{0}} \frac{iJ^{-1}\vec{\nu}}{(i\vec{\omega}_0 \cdot \vec{\nu})^2} f_{\vec{\nu}} e^{i\vec{\nu}\cdot\vec{\psi}} . \tag{7}$$

Suppose that $\vec{h}^{(k)}(\vec{\psi})$ is a trigonometric polynomial of degree $\leq kN$, odd in t, for $1 \leq k < k_0$. Then we see immediately that the r.h.s. of Eq. (6) is odd in t. This means that the r.h.s. of Eq. (6) has zero average in t, hence in $\vec{\psi}$, and the second of Eq. (6) can be solved for $k = k_0$. It yields an even function $\vec{H}^{(k_0)}(\vec{\psi})$ which is defined up to a constant which, however, must be taken such that $\vec{H}^{(k_0)}(\vec{\psi})$ has zero average, to make $\vec{\omega}\cdot\vec{\partial} h_j^{(k)} = J_j^{-1} H_j^{(k)}$ soluble. Hence the equation for $\vec{h}^{(k)}$ can be solved (because the r.h.s. has zero average) and its solution is a trigonometric polynomial in $\vec{\psi}$, odd if $\vec{h}^{(k)}$ is determined by imposing that its average over $\vec{\psi}$ vanishes.

Hence Eq. (7) provides an algorithm to evaluate a formal power series solution to our problem. It has been remarked, [E, V], see also [G], that Eq. (6) yields a *diagrammatic expansion* of $\vec{h}^{(k)}$. We simply "iterate" it until only $h^{(1)}$, given by Eq. (7), appears.

Let ϑ be a tree diagram: it will consist of a family of "lines" (i.e. segments) numbered for 1 to k arranged to form a (rooted) tree diagram as in Fig. 1:

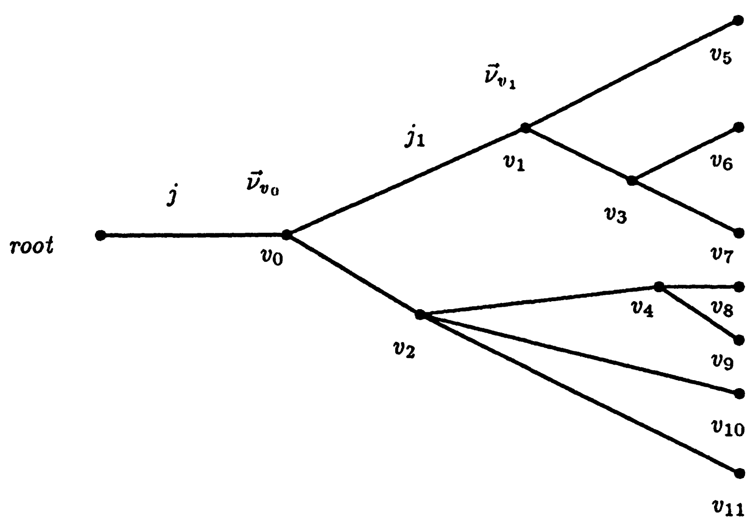


Fig. 1. A tree diagram ϑ with $m_{v_0} = 2$, $m_{v_1} = 2$, $m_{v_2} = 3$, $m_{v_3} = 2$, $m_{v_4} = 2$ and $m = 12$, $\prod m_v! = 2^4 \cdot 6$, and some decorations. The line numbers, distinguishing the lines, are not shown

To each vertex v we attach a "mode label" $\vec{\nu}_v \in Z^l$, $|\vec{\nu}_v| \leq N$ 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 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).

The mode labels, the branch labels and the number labels "decorating" the tree will be called also decorations (i.e. the labels attached to the tree and the decorations are synonymous below). Since the tree is partially ordered not every pair of vertices will be related by the order relation: we say that two vertices are comparable if they are related by the order relation (which we denote $\leq$).

A reader not used to diagrammatic expansions, like the ones in quantum field theory or in statistical mechanics (see [G3, Br]) may find the above concepts somewhat strange and new (they are not), and get easily confused. Therefore a more formal description of the expansion is in Appendix R.

We define the "momentum" entering v as $\vec{\nu}(v) = \sum_{w>v} \vec{\nu}_w$. If from a vertex v emerge m_1 lines carrying a label $j = 1$, m_2 lines carrying $j = 2, \ldots$, it follows that Eq. (6) can be rewritten:

$$\vec{h}^{(k)}_{\vec{\nu}j} = \frac{1}{k!} \sum^{*} \prod_{v \in \vartheta} \frac{(-iJ^{-1}\vec{\nu}_v)_{j_v} f_{\vec{\nu}_v} \prod_{s=1}^{l} (i\vec{\nu}_v)^{m_s}}{(i\vec{\omega}_0 \cdot \vec{\nu}(v))^2} \tag{8}$$

with the sum running over the diagrams ϑ of order k and with $\vec{\nu}(v_0) = \vec{\nu}$; and the combinatories can be checked from Eq. (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 $\vec{\nu}(v) \neq \vec{0}$ for all $v \in \vartheta$ (by the above remarked parity properties). Note that Eq. (8) is implied by the corresponding (6.23) of [G]: one can check that the two formulae coincide (by summing over what in [G] are called the "fruit values"). The theory in [G] is in fact a little more general, although it is really applied to the same Thirring model.

There are other diagrams, however, which we would like to eliminate. They are the diagrams with nodes v', v, with $v' < v$, not necessarily nearest neighbours, such that $\vec{\nu}(v) = \vec{\nu}(v')$. Such diagrams are, according to Eliasson's terminology, *in resonance.*

To see this, *for the purpose of illustration*, we shall first restrict the sum in Eq. (8) to a sum over diagrams such that:

I) $\vec{\nu}(v) \neq \vec{0}$ if v is any vertex.

II) $\vec{\nu}(v) \neq \vec{\nu}(v')$ 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 $\vec{\nu}_v$ can be chosen in a number of ways bounded by $(2N + 1)^{lk} < (3N)^{lk}$. Therefore, if

$f_0 = \max_{\vec{\nu}} |f_{\vec{\nu}}|$:

$$|h^{(k)}_{\vec{\nu}j}| \le (3N)^{lk} 2^{2k} l^k \frac{f_0^k C_0^{2k}}{J_0^k} N^{2k-1} \max_{\vartheta} \prod_{v\in\vartheta} (C_0 \vec{\omega}_0 \cdot \vec{\nu}(v))^{-2}$$
$$\le (f_0 C_0^2 J_0^{-1})^k N^{(l+2)k-1} (4l3^l)^k M \,, \tag{9}$$

where the maximum is over the diagrams ϑ verifying 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 τ is the diophantine constant in Eq. (3): then $C_0|\vec{\omega}_0 \cdot \vec{\nu}| \ge 3\zeta(q)$ if $0 < |\vec{\nu}| \le q$: we can say that $\vec{\nu} \in Z^l$ is "q-singular" if $C_0|\vec{\omega}_0 \cdot \vec{\nu}| < \zeta(q)$. Then the following (extension) of a lemma by Brjuno, see [P], holds for diagrams of degree k verifying I), II) above. Fixed $q \ge 1$, let $N(k)$ be the number of q-singular harmonics, among the vertex momenta $\vec{\nu}(v)$ of the vertices $v \in \vartheta$. Then $N(k) \le 2\,\dfrac{k}{2}$.

Assuming the above claim true (we shall not use it outside the present heuristic argument), we fix an exponentially decreasing sequence γ^n; the choice $\gamma = 2$ recommends itself. The number of $q = \gamma^{-n}$-singular harmonics which are not $q = \gamma^{-(n-1)}$-singular is bounded by $2k\gamma^n$, (being trivially bounded by the number of γ^{-n}-singular harmonics!). Hence

$$\prod_{v\in\vartheta} \frac{1}{(C_0\vec{\omega}_0 \cdot \vec{\nu}(v))^2} \le \prod_{n\le -1}^{-\infty} \zeta(\gamma^{-(n-1)})^{-4k\gamma^n}$$
$$\le \exp 4k \sum_{n\ge 1} 2^{-n} \log(3 \cdot 2^{\tau(n-1)} N^\tau) = M \,, \tag{10}$$

therefore the series for the approximation to $\vec{h}^{(k)}_{\vec{\nu}}$, that we are considering because of the extra restriction in the sum Eq. (8), has radius of convergence in ε bounded below.

However *there are* resonant diagrams. The key remark is that they cancel *almost exactly*. The reason is very simple: imagine to detach from a diagram ϑ the subdiagram ϑ_2 with first node v. Then attach it to all the *remaining* vertices $w \ge v'$, $w \in \vartheta/\vartheta_2$ (this means the graph ϑ, with the subgraph 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 $\vec{\nu}(v)$: this means that some of the denominators $(\vec{\omega} \cdot \vec{\nu}(w))^{-2}$ have become $(\vec{\omega} \cdot \vec{\nu}(w) \pm \varepsilon)^{-2}$ if $\varepsilon \equiv \vec{\omega}_0 \cdot \vec{\nu}(v)$; and:

2) because there is one of the vertex factors which changes by taking successively the values ν_{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 $\vec{\omega} \cdot \vec{\nu} = \varepsilon = 0$ we would build in this *resummation* a quantity proportional to: $\sum \vec{\nu}_w = \vec{\nu}(v') - \vec{\nu}(v)$ which is zero, because $\vec{\nu}(v) = \vec{\nu}(v')$ means that the sum of the $\vec{\nu}_w$'s vanishes. Since $\vec{\omega} \cdot \vec{\nu} = \varepsilon \neq 0$ we can expect to see a sum of order ε^2, if we sum as well on a overall change of sign of the $\vec{\nu}_w$ values (which sum up to $\vec{0}$).

But this can be true only if $\varepsilon \ll \vec{\omega} \cdot \vec{\nu}'$, for any line momentum $\vec{\nu}'$ of a line in ϑ/ϑ_2. If the latter property is not true this means that $\vec{\omega} \cdot \vec{\nu}'$ is small and that there are many vertices in ϑ/ϑ_2 of order of the amount needed to create a momentum with small divisors of order ε.

Examining carefully the proof of Brjuno's lemma one sees that such an extreme case would be essentially also treatable. Therefore the problem is to show that the two regimes just envisaged (and their "combinations") do exhaust all possibilities.

Such problems are very common in renormalization theory and are called "overlapping divergences." Their systematic analysis is made through the renormalization group methods. We argue here that Elliasson's method can be interpreted in the same way.

The above introduced diagrams will play the role of *Feynman diagrams*; 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.

We fix a *scaling* parameter γ, which we take $\gamma = 2$ for consistency with Eq. (3), and we also define $\vec{\omega} \equiv C_0\vec{\omega}_0$: it is an adimensional frequency. Then we say that a propagator $(\vec{\omega}\cdot\vec{\nu})^{-2}$ is *on scale* n if $2^{n-1} < |\vec{\omega}\cdot\vec{\nu}| \le 2^n$, for $n \le 0$, and we set $n = 1$ if $1 < |\vec{\omega}\cdot\vec{\nu}|$.

Proceeding as in quantum field theory, see [G3], given a diagram ϑ we can attach a *scale label* to each line $v'v$ in Eq. (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*.

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 $\le E2^{-n\varepsilon}$*, 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.*

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 $N_n > 0$.

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:

$$N_n \le \frac{4k}{E2^{-\varepsilon n}} + \sum_{T, n_T = n} (-1 + m_T) \tag{11}$$

with $E = N^{-1}2^{-3\varepsilon}$, $\varepsilon = \tau^{-1}$. This is a version of Brjuno's lemma: a proof is in appendix.

Consider a diagram ϑ^1 we define the family $\mathscr{F}(\vartheta^1)$ 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. Note that all the lines λ in V (i.e. contained in V and with at least one point in $\tilde{V}$) have a scale $n_\lambda \geq n_V$.

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 simultaneously the signs of the vertex modes $\vec{\nu}_w$, for $w \in \tilde{V}$: the change of sign is performed independently for the various resonant clusters. This defines a family of $\prod 2M_V$ diagrams that we call $\mathscr{F}(\vartheta_1)$. The number $\prod 2M_V$ will be bounded by $\exp \sum 2M_V \leq e^{2k}$.

It is important to note that the definition of resonance is such that the above operation (of shift of the vertex to which the line entering V is attached) does not change too much the scales of the diagram lines inside the resonances: the reason is simply that inside a resonance of scale n the number of lines is not very large, being $\leq \bar{N}_n \equiv E2^{-n\varepsilon}$.

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 at most a change in the size of the propagator of λ by at most $2^{n_1} + 2^{n_2} + \ldots < 2^{n+1}$.

Since the number of lines inside V is smaller than $\bar{N}_n$ 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 $|\vec{\omega} \cdot \vec{\nu}^0_\lambda| \geq 2^{n+3}$ (because $\vec{\nu}^0_\lambda$ is a sum of $\leq \bar{N}_n$ vertex modes, so that $|\vec{\nu}^0_\lambda| \leq N\bar{N}_n$) but $\vec{\omega} \cdot \vec{\nu}^0_\lambda$ is "in the middle" of the diadic interval containing it and by Eq. (3) 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 $\mathscr{F}(\vartheta^1)$, if $\vec{\omega}_0$ verifies Eq. (3).

This implies, by the strong diophantine hypothesis on $\vec{\omega}_0$, Eq. (3), that the resonant clusters of the diagrams in $\mathscr{F}(\vartheta^1)$ all contain the same sets of lines, and the same lines go in or out of each resonance (although they are attached to generally distinct vertices inside the resonances: the identity of the lines is here defined by the number label that each of them carries in ϑ^1). Furthermore the resonance scales and the scales of the resonant clusters, and of all the lines, do not change.

Let ϑ^2 be a diagram not in $\mathscr{F}(\vartheta^1)$ and construct $\mathscr{F}(\vartheta^2)$, etc. We define a collection $\{\mathscr{F}(\vartheta^i)\}_{i=1,2,\ldots}$ 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 the *Eliasson's resummation.*

We call ε_V the quantity $\vec{\omega} \cdot \vec{\nu}_{\lambda_V}$ associated with the resonance V. If λ is a line in $\tilde{V}$, see above, 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 $\sigma_\lambda = 0, \pm 1$. Since $|\vec{\omega} \cdot \vec{\nu}_\lambda| > 2^{n_V - 1}$ we see that the product of the propagators is holomorphic in ε_V for $|\varepsilon_V| < 2^{n_V-3}$.[2] While ε_V varies in such complex disk the quantity $|\vec{\omega} \cdot \vec{\nu}^0_\lambda + \sigma_\lambda \varepsilon_V|$ does not become smaller than 2^{n_V-3}. Note the main point here: the quantity 2^{n_V-3} will usually be $\gg 2^{n_{\lambda_V}}$ which is the value ε_V actually can reach in every diagram in $\mathscr{F}(\vartheta^1)$; 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 $\prod 2M_V \leq e^{2k}$ products of propagators of the members of the family $\mathscr{F}(\vartheta^1)$ can be bounded above by

[2] In fact $|\vec{\omega} \cdot \vec{\nu}^0_\lambda| \geq 2^{n+3}$ because V is a resonance; therefore $|\vec{\omega} \cdot \vec{\nu}_\lambda| \geq 2^{n+3} - 2^{n+1} > 2^{n+2}$ so that $n_V \geq n + 3$. On the other hand we note that $|\vec{\omega} \cdot \vec{\nu}^0_\lambda| > 2^{n_V-1} - 2^{n+1}$, so that it follows that $|\vec{\omega} \cdot \vec{\nu}^0_\lambda + \sigma_\lambda \varepsilon_V| \geq 2^{n_V-1} - 2^{n+1} - 2^{n_V-3} \geq 2^{n_V-3}$, for $|\varepsilon_V| < 2^{n_V-3}$

$\prod_\lambda 2^{-2(n_\lambda-3)} = 2^{6k} \prod_\lambda 2^{-2n_\lambda}$, if regarded as a function of the quantities $\varepsilon_V = \vec{\omega}\cdot\vec{\nu}_{\lambda_V}$, for $|\varepsilon_V| \le 2^{n_V-3}$, 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 $\mathscr{F}(\vartheta^1)$, 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 Eq. (8) can be bounded by N^2, we can bound the contribution from the family $\mathscr{F}(\vartheta^1)$ by:

$$\left[\frac{1}{k!}\left(\frac{f_0 C_0^2 N^2}{J_0}\right)^k 2^{6k} e^{2k} \prod_{n\le 0} 2^{-2nN_n}\right]\left[\prod_{n\le 0}\prod_{T,n_T=n}\prod_{i=1}^{m_T} 2^{2(n-n_i+3)}\right], \tag{12}$$

where:

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

2) the first square bracket is the bound on the product of individual elements in the family $\mathscr{F}(\vartheta^1)$ times the bound e^{2k} on their number,

3) The second term is the part coming from the maximum principle, 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 $|\varepsilon_i| < 2^{n_i-3}$ if $n_i \equiv n_{V_i}$, provided $n_i > n+3$ (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 Eq. (12) by the product of factors $(|\varepsilon_i| 2^{-n_i+3})^2$ if $n_i > n+3$. If $n_i \le n+3$ we cannot gain anything: but since the contribution to the bound from such terms in Eq. (12) is > 1 we can leave them in it to simplify the notation, (of course this means that the gain factor can be important only when $\ll 1$).

Hence substituting Eq. (11) into Eq. (12) 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 Eq. (11), taken from the factors with $T = V_i$, (note that there are always enough -1's).

Hence the product Eq. (12) is bounded by:

$$\frac{1}{k!}(C_0^2 J_0^{-1} f_0 N^2)^k e^{2k} 2^{12k} \prod_n 2^{-8nkE^{-1}2^{\varepsilon n}} \le \frac{1}{k!} B_0^k \tag{13}$$

with B_0 suitably chosen.

To sum over the trees we note that, with fixed ϑ the collection of clusters if fixed. Therefore we only have to multiply Eq. (13) 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 Eq. (5).

Comments. The strong diophantine condition is quite unpleasant as it seems to put an extra requirement on $\vec{\omega}_0$: I think that in fact such a condition is not necessary. Given $\vec{\omega}_0$ verifying the first of Eq. (3) with some constant $\bar{C}_0$ and some τ we can imagine to define $C_0 \equiv 2^{r+1}\bar{C}_0$: this leaves the first of Eq. (3) still valid. One should then note that the $0 < |\vec{\nu}| \le (2^{n+3})^{-\tau^{-1}}$ will imply that if $0 < |\vec{\nu}| \le (2^{n+3})^{-\tau^{-1}}$

the numbers $|\vec{\omega} \cdot \vec{\nu}|$ are spaced by at least $(2(2^{\tau}2^{n+3})^{-\tau^{-1}})^{-\tau} = 2^{n+3}$. Hence we can find a sequence γ_n such that $1 \leq \gamma_n 2^{-n} \leq 2$ and $||\vec{\omega} \cdot \vec{\nu}| - \gamma_p| \geq 2^{n+2}$ for $0 \geq p \geq n$. Defining a propagator x^{-2} to have "scale n" if $\gamma_{n-1} < |x| \leq \gamma_n$ it should be possible to perform the proof without the second assumption in Eq. (3). This will be interesting to check if it really works as it would be fully constructive, given $\vec{\omega}_0$ verifying the first of Eq. (3), and as it would cover the most studied cases like the case $l = 2$ and $\vec{\omega} = (\omega_1, \omega_2) = (1, \omega)$ with ω a quadratic irrational, or a number with uniformly bounded continued fraction entries. It has also the conceptual advantage that the sequence of scales γ_n is not "prescribed a priori" but it is determined by the arithmetic properties of the rotation vector $\vec{\omega}_0$. While this paper was being refereed G. Gentile and I succeeded in proving the above conjecture, [G5].

I have not considered at all the cases in which f is analytic in $\vec{\alpha}$ (rather than a trigonometric polynomial). Nor have I considered the cases in which the perturbation is not even. My aim here was to consider the simplest case I could find in which the KAM theorem needed a proof and prove it by just bounding the series (called in the literature since Poincaré the *Lindstedt series*) order by order.

Appendix A1: Resonant Siegel-Brjuno Bound

Calling N_n^* the number of non-resonant lines carrying a scale label $\leq n$, 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 $N^*(\vartheta)$.

If ϑ has the root line with scale $> n$ then calling $\vartheta_1, \vartheta_2, \dots, \vartheta_m$ the subdiagrams of ϑ emerging from the first vertex of ϑ 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 Eq. (11).[3]

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 ϑ_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 ϑ_1 is either a resonant line or it has scale $> n$.

Accepting the last statement we have: $N^*(\vartheta) = 1 + N^*(\vartheta_1) = 1 + N^*(\vartheta'_1) + \dots + N^*(\vartheta'_{m'})$, with ϑ'_j being the m' subdiagrams emerging from the first node of ϑ'_1 with orders $k'_j > E2^{-\varepsilon n}$: 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 $N^*(\vartheta'_1) = N^*(\vartheta''_1) + \dots + N^*(\vartheta''_{m''})$, 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 ϑ_1 has scale $> n$, unless it is entering a resonance.

[3] Note that if $k \leq E2^{-n\varepsilon}$ it is, for all momenta $\vec{\nu}$ of the lines, $|\vec{\nu}| \leq NE2^{-n\varepsilon}$, i.e. $|\vec{\omega} \cdot \vec{\nu}| \geq (NE2^{-n\varepsilon})^{-\tau} = 2^3 2^n$ so that there are *no* clusters T with $n_T = n$ and $N^* = 0$. The choice $E = N^{-1}2^{-3\varepsilon}$ is convenient: but this, as well as the whole lemma, remains true if 3 is replaced by any number larger than 1. The choice of 3 is made only to simplify some of the arguments based on the resonance concept

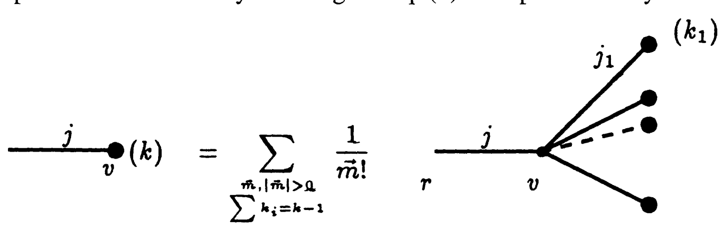
Suppose that the root line of ϑ_1 has scale $\leq n$ and is not entering a resonance. Note that $|\vec{\omega} \cdot \vec{\nu}(v_0)| \leq 2^n$, $|\vec{\omega} \cdot \vec{\nu}(v_1)| \leq 2^n$, if v_0, v_1 are the first vertices of ϑ and ϑ_1 respectively. Hence $\delta \equiv |(\vec{\omega} \cdot (\vec{\nu}(v_0) - \vec{\nu}(v_1))| \leq 2\,2^n$ 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 $k - k_1 < \frac{1}{2}\, E2^{-n\varepsilon}$ (and we are not considering the resonances: note also that in such a case the lines in ϑ/ϑ_1 different from the root of ϑ must be inside a cluster, see footnote 3)), it follows that $k - k_1 < \frac{1}{2}\, E2^{-n\varepsilon}$ 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 $|\vec{\nu}(v_0) - \vec{\nu}(v_1)| < \frac{1}{2}\, NE2^{-n\varepsilon}$, 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 $p \leq 2k(E2^{-\varepsilon n})^{-1} - 1$. Thus Eq. (11) is proved.

Remark. The above argument is a minor adaptation of Brjuno's proof of Siegel's theorem, as remarkably exposed by Pöschel, [P].

Appendix R: The Tree Representation

The tree representation arises by thinking of Eq. (6) as represented by:



where the l.h.s. is a symbol for $h_j^{(k)}$ and in the r.h.s. we have a "simple tree" consisting of a "root' r", a "root branch" rv "leading to the vertex v" and $|\vec{m}| \equiv m_v$ branches "emerging from v."

The vector $\vec{m} = (m_1, \ldots, m_l)$ describes the multiplicity m_i with which the value i occurs in $(j_1, j_2, \ldots, j_{|\vec{m}|})$ and $\vec{m}! = \prod_{i=1}^{l} m_i!$; and $|m!| \equiv \sum_{i=1}^{l} m_i$. We adopt (temporarily, see below) the convention that the first m_1 among the j's are 1, the next m_2 are 2, etc. from top to bottom (in the above figure).

The length of the branches and the angles at which they are drawn (with respect to the sides of the present sheet of paper) are irrelevant.

The vertex v symbolizes $-\partial_{\alpha_j} \partial_{\alpha_{j_1}} \ldots \partial_{\alpha_{j_{|\vec{m}|}}} f(\vec{\omega}_0 t)$ (which means that one takes the indicated derivatives of $f(\vec{\alpha})$ and evaluates them at $\vec{\omega}_0 t$).

The branches ending in the heavy dots with label (k_j) and carrying a label j, see the above figure, represent $h_j^{(k_j)}$, see Eq. (6).

At this point, if one notes that Eq. (6) is multilinear in the $h^{(k_j)}$ (of degree 1 in each), it is clear that we can just replace each of the branches ending in heavy dots with the same graphical expression in the r.h.s. of the above figure. And so on, until the labels (k_j) on all the branches ending in a heavy dot ("top branches") are equal to 1. In this case they will represent $h_j^{(1)}$'s given by Eq. (7).

Thus we have represented our $h^{(k)}$ as a "sum over trees," with k branches and k vertices (as we do not regard the root as a true vertex), of suitable "tree values," which are simply a product of derivatives symbolized by the vertices and of the $h^{(1)}$'s corresponding to the top vertices. The labels attached to the trees determine exactly which derivatives and which components of the $h^{(1)}$ we have to take for a given tree. The top vertices carry a label (k) with, necessarily, $k = 1$ which will therefore be omitted.

However the above tree expansion has rather complicated combinatorial weights $\prod_v m_v!^{-1}$. There is a combinatorially better representation, in which more "trees" are involved, but all with the same weight $k!^{-1}$. It is obtained by simply adding a "number label" n_b to each of the k branches, paying attention that no two branches get the same label and $n_b = 1, 2, \ldots, k$. There are $k!$ ways of doing this on each of the above trees and one realizes that if one evaluates the value of such new "numbered trees" as if the number labels were absent and if one sums the values divided by $k!$ over all the numbered trees, i.e. over all trees whose branches bear the labels j_b, n_b, ("branch label" and "number label") one still gets the correct result (there are in fact $\frac{k!}{\prod_v m_v!}$ distinct numbered trees per each of the previously defined trees.) Of course the way they are drawn on a sheet of paper is irrelevant, as long as by adjusting the length of the drawn branches, and by suitably permuting the branches which emerge from one vertex one can superpose the trees (branch and number labels included). This defines the group of actions on the tree drawings defined in the text.

By construction a tree comes with a natural partial ordering (as *any* rooted tree) we spell it out to avoid misunderstandings. We say that a vertex v follows v' (another vertex), and we write $v' \leq v$ if there is a path that starting from the root proceeds along the tree branches without ever visiting a point twice, and reaching v' first and then v. If $v' \leq v$ or if $v \leq v'$ we say, following the standard notations of set theory, that v and v' are comparable.

Finally since f and thus $h^{(1)}$ are a sum of Fourier components, see Eq. (2), and the tree value is multilinear in the derivatives of f, symbolized by the vertices, we see that we can develop each f, $h^{(1)}$ in its Fourier sum Eq. (2), Eq. (7). This can be simply represented by adding one more decoration $\vec{\nu}_v$, the "mode label," on each vertex v symbolizing that in the Fourier expansion of the f corresponding to the vertex v we have selected the component $f_{\vec{\nu}} e^{i\vec{\nu}\cdot\vec{\alpha}}$ (there is no factor $\frac{1}{2}$ coming from the Euler formula of the cosine, because in Eq. (2) the modes $\pm\vec{\nu}$ are both counted). If $\vec{h}^{(k)}(\vec{\omega}_0 t)$ is also regarded as a sum over $\vec{\nu}$ of $e^{i\vec{\omega}\cdot\vec{\nu}t}$ times coefficients $\vec{h}^{(k)}_{\vec{\nu}}$, then the net result is Eq. (8). All the above analysis is well known in graph theory, and it has been exploited many times in recent years (see for instance [G3]). Since the order of the tree is k, equal to the number of vertices, the above trees can be interpreted as Feynman graphs. This paper shows that the analogy is in fact extremely close. Since this paper was completed I have been able, with the help of A. Berretti and G. Parisi, to find a field theory on the l dimensional torus, whose one point Schwinger function is exactly $h(\vec{\psi})_j$. This is interesting as it gives further weight to my idea of a connection of the KAM theorem with field theory which I have been pursuing for many years (however the field theory that one writes is not easy and in fact its theory coincides with the KAM theorem described above: as it should, I believe). This will be the subject of a further paper, see [G4].

Acknowledgements. I am indebted to J. Bricmont and M. Vittot for many clarifying discussions; and in particular to S. Miracle-Solé, G. Benfatto and G. Gentile for the same reasons and for their interest and suggestions.

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 Mat. Nauk **18**, 13–40 (1963)

[BG] Benfatto, G., Gallavotti, G.: Perturbation theory of the Fermi surface in a quantum liquid. A general quasi particle formalism and one dimensional systems. J. Stat. Phys. **59**, 541–664 (1990)
Benfatto, G., Gallavotti, G.: Renormalization group application to the theory of the Fermi surface. Phys. Rev. B **42**, 9967–9972 (1990)

[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)

[Br] Brydges, D.: Functional integrals and their applications. Notes with the collaboration of R. Fernandez; (Notes for a course for the Troisieme Cycle de la Physique en Suisse Romande given in Lausanne, Switzerland, 1992, archived in mp_arc@math.utexas.edu, #93-24, 1993

[E] Eliasson, L.H.: Hamiltonian systems with linear normal form near an invariant torus. Turchetti, G. (ed.) Bologna Conference, 30/5 to 3/6 1988, Singapore: World Scientific, 1989. And Generalization of an estimate of small divisors by Siegel, ed. E. Zehnder, P. Rabinowitz, book in honor of J. Moser, New York, London: 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, quasi flat homoclinic intersection, 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

[G2] Gallavotti, G.: The elements of mechanics. Berlin, Heidelberg, New York: Springer 1983

[G3] Gallavotti, G.: Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods. Rev. Mod. Phys. **57**, 471–572 (1985). See also, Gallavotti, G.: Quasi integrable mechanical systems. Les Houches, XLIII (1984), vol. II, pp. 539–624, Osterwalder, K., Stora, R. (eds.) Amsterdam: North Holland, 1986

[G4] Gallavotti, G.: Perturbation theory. Notes in margin to the Mathematical Physics towards the XXI century conference, University of Negev, Beer Sheva, 14–19 March, 1993. This text will be archived in the archive mp_arc@math.utexas.edu, and it will appear in the Conference Proceedings

[G5] Gallavotti, G., Gentile, G.: Non-recursive proof of the KAM theorem. Preprint, p. 1–8. Agosto, 1993, Roma, archived in mp_arc@math.utexas.edu, #93-229

[K] Kolmogorov, N.: On the preservation of conditionally periodic motions. Dokl. Akad. 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 **79**B, 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, Osterwalder, K., Stora, R. (eds.) Amsterdam: North-Holland, 1986

[PV] Percival, I., Vivaldi, F.: Critical dynamics and diagrams. Physica D **33**, 304–313 (1988)

[S] Siegel, K.: Iterations of analytic functions. Ann. Math. **43**, 607–612 (1943)

[T] Thirring, W.: Course in mathematical physics Vol. 1, p. 133, Wien: Springer 1983

[V] Vittot, M.: Lindstedt perturbation series in hamiltonian mechanics: Explicit formulation via a multidimensional Burmann-Lagrange formula. Preprint CNRS-Luminy, case 907, F-13288, Marseille, 1992

Communicated by J.-P. Eckmann