KAM theorem revisited

G.Gentile[1], **V.Mastropietro**[2]

Abstract: *A terse direct proof of the KAM theorem in the Thirring model is presented, by using a multiscale decomposition technique usual in quantum field theory formalism.*

Keywords: KAM theorem, perturbation theory, quantum field theory, renormalization group

The KAM theorem proves in a indirect way that the formal perturbation series for the invariant tori of quasi integrable systems are convergent, if some conditions are fulfilled by the hamiltonian. Of course one should be able to prove the convergence studing directly the perturbative series and this has been performed in recent times by Eliasson, [1]. In [2] the ideas of Eliasson are exposed via the analysis of a simple particular hamiltonian model, which is a system of rotators interacting via a potential, [3], which is chosen in [2] depending only on the angle variables. The hamiltonian function is

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

where J is the (non singular) matrix of the inertia moments, $\vec{A} = (A_1, \ldots, A_l) \in \mathbf{R}^l$ are their angular momenta and $\vec{\alpha} = (\alpha_1, \ldots, \alpha_l) \in \mathbf{T}^l$ are the angles describing their positions 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:

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

and without loss of generality we can suppose $f_{\vec{0}} = 0$. In [2] the convergence of the KAM tori for the above model is proven under the extra assumption that the rotation vectors satisfy a *strong diophantine condition*, *i.e.* : $\min_{0\ge p\ge n} |C_0|\vec{\omega}\cdot\vec{\nu}| - \gamma^p| > \gamma^{n+1}$, for $0 < |\vec{\nu}| \le (\gamma^{n+3})^{-\tau^{-1}}$, with $\gamma > 1$, (*e.g.* $\gamma = 2$). The proof combines Eliasson's ideas with the multiscale decomposition of the propagators of the Feynman graphs used in field theory, (see, *e.g.* , [4] for a review): the decomposition is performed by dividing $\mathbf{R}^+$ into diadic intervals, but such a partition has some technical inconvenients which can be bypassed if the assumption that the vector $\vec{\omega}$ is a strong diophantine one is accepted. Such assumption is removed in [5], by using a different suitable partion of $\mathbf{R}^+$, depending on the fixed rotation vector $\vec{\omega}$, and explicitly computable.

In this paper a proof semplified with respect to [2],[5] for the convergence of the KAM tori is given, following more closely the field theory techniques (see for instance [4]), adopting a smooth decomposiotion of the propagator and defining a renormalized perturbative expansion via the introduction of a *localization* operator. In this way the strong diophantine condition is not necessary to make simpler the proof, in distinction to [2], and the analogy with the quantum field theory techniques is more explicitly focused.

More precisely we prove the following result:

Theorem. *Given the hamiltonian model (1), with interaction potential (2), and fixed a "rotation vector" $\vec{\omega}$ satyisfying the* diophantine property $C_0|\vec{\omega}\cdot\vec{\nu}| \ge |\vec{\nu}|^{-\tau}$, $\vec{0} \ne \vec{\nu} \in \mathbf{Z}^l$ *with diophantine constants $C_0, \tau > 0$, there exist two functions $\vec{H}(\vec{\omega}t;\varepsilon)$, $\vec{h}(\vec{\omega}t;\varepsilon)$ analytic in in $\vec{\psi} \equiv \vec{\omega}t$ with $Re\vec{\psi} \in$*

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

[2] Address: Dipartimento di Matematica, Università di Roma II "Tor Vergata", Via della Ricerca Scientifica, 00133 Roma, Italia.

$\mathbf{T}^l$, *and* $|Im\vec{\psi}_j| < \xi$, *and analytic for* $|\varepsilon| < \varepsilon_0$, *where*

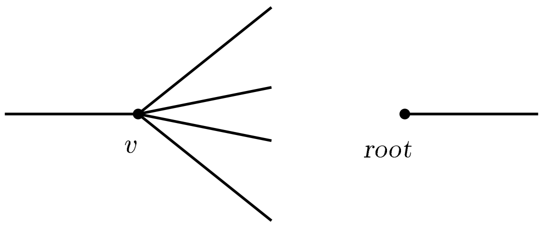
$$\varepsilon_0 = bJ_0^{-1}C_0^2[\max_{0<|\vec{\nu}|\le N} f_{\vec{\nu}}]N^{2+l}e^{cN}e^{\xi N} , \tag{3}$$

(for some positive constant b depending only on N and l), such that $\vec{A} = \vec{A}_0 + \vec{H}(\vec{\omega}t;\varepsilon)$, $\vec{A}_0 = J\vec{\omega}$, *and* $\vec{\alpha} = \vec{\omega}t + \vec{h}(\vec{\omega}t;\varepsilon)$ *describe a family of invariant tori which is run quasi-periodically with angular velocity* $\vec{\omega}$.

One recognizes a version of the KAM theorem, [6]. 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.

Calling $\vec{H}^{(k)}_{\vec{\nu}}$, $\vec{H}^{(k)}_{\vec{\nu}}$, $\vec{h}^{(k)}_{\vec{\nu}}$ the $\vec{\nu}$-th Fourier components of the k-th order coefficients of the Taylor expansion of $\vec{H}(\vec{\psi};\varepsilon)$, $\vec{h}(\vec{\psi};\varepsilon)$ in powers of ϵ, from the analysis of the equations of motion we deduce that: 1) the components with $\vec{\nu} = \vec{0}$ are identically vanishing; 2) the $\vec{H}^{(k)}_{\vec{\nu}}$'s can be trivially expressed as functions of the $\vec{h}^{(k)}_{\vec{\nu}}$'s, *i.e.* $J^{-1}\vec{H}^{(k)}_{\vec{\nu}} = (i\vec{\omega}\cdot\vec{\nu})\vec{h}^{(k)}_{\vec{\nu}}$; 3) $\vec{h}^{(k)}_{\vec{\nu}}$'s admit a graphical representation, in terms of *Feynman's graphs*.

A Feynman's graph θ of order k is obtained in according to the following rules. One lays down k *graphs elements* formed by a point v (*vertex*), from which only one (oriented) half-line comes out and which $m_v \ge 0$ (oriented) half-lines enters, plus a unique graph element formed by a point (*root*), which only one line enters and from which no lines come out: the root is not counted among the vertices. Then one considers all the possibles ways of joining togheter lines in pairs so that no line is left over unpaired and only lines with consistent, (*i.e.* opposite), orientations are allowed to form a pair; if a line λ enters a vertex v, we write $\lambda = \lambda_v$. By construction no loop can be formed: the graphs we obtain are trees. To order k, not considering the labels which will be introduced below, the number of graphs is bounded by 2^{2k}.



One associates to each line λ a *momentum* $\vec{\nu}_\lambda$, a label $j_\lambda \in \{1,\dots,l\}$, and a *propagator* $g_\lambda = [i\vec{\omega}\cdot\vec{\nu}_\lambda]^{-2}$.

To each vertex v there corresponds a label $\vec{\nu}_v$ (*mode label*), such that $0 < |\vec{\nu}_v| \le N$, and a *vertex factor*

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

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 $\vec{\nu}_v$).

Like in the renormalization group approach to the analysis of the perturbation expansions of field theory, it can be convenient to introduce a multiscale decomposition of the propagators, [4], in the following way. Let $\chi(x)$ be a C^∞ function such that $\chi(x) = 0$, if $|x| \ge 2$ and $\chi(x) = 1$ if $|x| \le 1$, and let $\chi_n(x) = \chi(2^{-n}x) - \chi(2^{-(n-1)}x)$, $n \le 0$, and $\chi_1(x) = 1 - \chi(x)$: such functions realize a C^∞ partition of unity so that we can decompose the propagator

$$g_\lambda = \frac{1}{[\vec{\omega}\cdot\vec{\nu}_\lambda]^2} = \sum_{n=-\infty}^{1} \frac{\chi_n(\vec{\omega}\cdot\vec{\nu}_\lambda)}{[\vec{\omega}\cdot\vec{\nu}_\lambda]^2} \equiv \sum_{n=-\infty}^{1} g^{(n)}_\lambda , \tag{4}$$

where $g_\lambda^{(n)}$ is the "propagator at scale n". If $n \le 0$, $g_\lambda^{(n)}$ is a C^∞ compact support function different from 0 for $2^{n-1} \le |\vec{\omega} \cdot \vec{\nu}_\lambda| \le 2^{n+1}$, while $g_\lambda^{(1)}$ has support for $1 \le |\vec{\omega} \cdot \vec{\nu}_\lambda|$. In the domain where it is different from zero, the propagator verifies the bound $\left|\partial^p g_\lambda^{(n)}(x)\right| \ge a(p) 2^{-n(2+p)}$ for any $p \in \mathbf{N}$, $2^{n-1} \le |x| \le 2^{n+1}$, being the derivative with respect to the argument $\vec{\omega} \cdot \vec{\nu}_\lambda$, and $a(p)$ a suitable constant, such that $a(0) = 2^2$. Then one associates to each line λ an extra label n_λ, (*scale label*).

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

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 $T' \subset T$. One can associate to θ a *graph value* $X(\theta)$, recursively defined as

$$X(\theta) \equiv X(T) = \prod_{\lambda \in T_0} g_\lambda^{(n_T)} \prod_{T' \subset T} X(T') \prod_{v \in T_0} \mathcal{E}_v \;, \tag{5}$$

where $T' \subset T$ means that the cluster T' is a maximal cluster inside T. We can iterate (5), until we obtain

$$X(\theta) = \prod_{v \in \theta} g_{\lambda_v}^{(n_{\lambda_v})} \mathcal{E}_v \;, \tag{6}$$

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

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*. Finally, 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.

Given a graph θ, let us define $N_n(\theta)$ the number of lines in θ with scale $n \le 0$. Then, as we can deduce from (6), the graph value $X(\theta)$ admits the bound

$$\mathcal{C}^k \prod_{n \le 0} 2^{-2n N_n(\theta)} \;, \tag{7}$$

where $\mathcal{C} = 2^2 J_m^{-1} C_0^2 N^2 [\max_{\vec{\nu}} f_{\vec{\nu}}]$, and $N_n(\theta)$ verifies the bound (*Brjiuno's lemma*, [7])

$$N_n(\theta) \le \frac{4k}{E 2^{-n\tau^{-1}}} + \sum_{T, n_T = n} [-1 + m_T(\theta)] \;, \tag{8}$$

where $m_T(\theta)$ 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 $E = 2^{-3\tau^{-1}} N^{-1}$.

We stress that the only point in the proof of the theorem where the diophantine property is used is exactly in Brjiuno's lemma: the proof is standard, and can be found in [8], and it is in turn an adaptation from [2]. If there were not the resonances, from (7) and (8), we would obtain a convergent bound. However *there are resonances*, so we have to deal with them.

This situation is strongly reminescent of renormalizable field theory. One has a perturbative expansion in term of Feynman's graphs which is convergent only if some classes of subgraphs, (the resonances in our case), are avoided: therefore the renormalized expansion is implemented via the introduction of 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

$$\mathcal{L}X(V)\equiv\mathcal{L}\mathcal{V}(\vec{\omega}\cdot\vec{\nu}_{\lambda_V})=\mathcal{V}(0)+(\vec{\omega}\cdot\vec{\nu}_{\lambda_V})\dot{\mathcal{V}}(0)\ , \tag{9}$$

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 $\mathcal{R}=1-\mathcal{L}$. The sum over graphs containing one or more resonances on which $\mathcal{L}$ applies is vanishing, because of cancellation mechanisms which work exactly to second order: imagine to detach from a graph the subgraph with first line λ_V entering a resonance V, hence attach it to all the remaining vertices $v\in V$, external to the inner resonances, *i.e.* $v\in V_0$. The expression we obtain is vanishing to first order, as the various contributions differ because a term $(i\vec{\nu}_v)_{j_{\lambda_V}}$ is successively associated to a different vertex $v\in V_0$, and $\sum_{v\in V_0}\vec{\nu}_v=\vec{0}$, and to second order by parity properties, so that we can rule out all such contributions and consider simply the graphs in which $\mathcal{R}$ applies.

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

$$\mathcal{R}\mathcal{V}(\vec{\omega}\cdot\vec{\nu})=(\vec{\omega}\cdot\vec{\nu})^2\int_0^1 dt\ t\ddot{\mathcal{V}}(t\vec{\omega}\cdot\vec{\nu})\ , \tag{10}$$

where $\ddot{V}$ 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. This would be of course a source of problems, as $a(p)>p!$, where $a(p)$ is defined after (4). However it is not so: in fact 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

$$\mathcal{R}\Big(\prod_{\lambda\in V_0}g_\lambda^{(n_V)}\prod_{T'\in V}\mathcal{R}X(T')\prod_{v\in V_0}\mathcal{E}_v\Big)\ , \tag{11}$$

where, for any resonance $\tilde{V}\subseteq V$, $\mathcal{R}X(\tilde{V})$ can be written either as in (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 (10), leaving the other terms $\mathcal{R}X(T')$ written as differences when T' are resonances: so (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, $\partial^p\mathcal{R}X(T')\equiv\partial^pX(T')$, for $0\le p\le 2$.

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 $\partial\mathcal{R}\mathcal{V}(\vec{\omega}\cdot\vec{\nu})=\partial\mathcal{V}(\vec{\omega}\cdot\vec{\nu})-\dot{\mathcal{V}}(0)=(\vec{\omega}\cdot\vec{\nu})\int_0^1dt\ddot{\mathcal{V}}(t\vec{\omega}\cdot\vec{\nu})$, while, if two derivatives act on $\mathcal{R}X(V')$, 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, and the procedure can be iterated, since the resonances V' can be dealt with as the resonance V.

The effect of the $\mathcal{R}$ operator is to obtain the gain factor $2^{n-n'}2^{n-n''}$, where n' and n'' are the scales of two lines λ' and λ'', possibly coinciding, contained in some clusters T' and T'' inside V. So we can rewrite, *e.g.* ,the first factor as $2^{n-n'}=2^{n-n_1}\dots 2^{n_q-n'}$, where n_i is the scale of the cluster $T_i\supset T_{i+1}$, with $T_1=V$ and $T_q=T'$. Analogous considerations hold for n'', so that we can conclude that: 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 (7) can be replaced with

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

where the last product in the left hand side is over the resonances in θ, and $\mathcal{C}_0=a(2)e^2\mathcal{C}$, as $k^2\le e^{2k}$, and $a(p)\le a(2)$, for $0\le p\le 2$, and the right hand side follows from the above discussion about the gain factors and the Brjiuno bound (8). Since the sum over the graphs is bounded by $2^{2k}N^{2k}2^k l^k(2N+1)^k$, this completes the proof of the Theorem.

Acknowledgements. We want to thank G.Gallavotti for proposing us the subject of this work, and for introducing us to a field theory approach to the KAM theorem.

References

[1] L.H.Eliasson: *Absolutely convergent series expansions for quasi-periodic motions*, report 2-88, Dept. of Math., University of Stockholm (1988).

[2] G.Gallavotti: *Twistless KAM tori*, Communications in Mathematical Physics, **164**, 145-156 (1994)

[3] W.Thirring: *Course in Mathematical Physics*, vol. 1, p. 133, Springer, Wien (1983).

[4] G.Gallavotti: *Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods*, Reviews in Modern Physics, **57**, 471- 572 (1985)

[5] G.Gallavotti, G.Gentile: *Majorant series convergence for twistless KAM tori*, preprint, 1993, to appear in Ergodic theory and dynamical systems, archived in `mp_arc@math. utexas. edu`, #93-229.

[6] A.N.Kolmogorov: *On the preservation of conditionally periodic motions*, Doklady Akademia Nauk SSSR, **96**, 527-530 (1954). V.Arnold: *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). J.Moser: *On invariant curves of an area preserving mapping of the annulus*, Nachrichten Akademie Wiss. Göttingen, **11**, 1-20 (1962).

[7] A.Brjuno: *The analytic form of differential equations*, I: Transactions of the Moscow Mathematical Society, **25**, 131-288, 1971; II: **26**, 199-239 (1972).

[8] G.Gentile, V.Mastropietro: *Tree expansion and multiscale analysis for KAM tori*, preprint (1994).