Lindstedt series and Kolmogorov theorem.

*Giovanni Gallavotti**

Abstract: the KAM theorem from a combinatorial viewpoint.

Lindstedt, Newcomb and Poincaré introduced a remarkable trigonometric series motivated by the analysis of the three body problem, [P]. In modern language, [G2], it is the generating function, that I call *Lindstedt series* here, $\vec{h}(\vec{\psi}) = \sum_{k=1}^{\infty} \varepsilon^k \vec{h}^{(k)}(\vec{\psi})$ of the sequence $\vec{h}^{(k)}(\vec{\psi})$ of trigonometric polynomials associated with well known combinatorial objects, namely *rooted trees*.

It is necessary to recall, first, the notion of rooted tree as used here. We lay down one after the other on a plane k pairwise distinct unit segments oriented from one endpoint to the other (respectively the *initial point* and the *endpoint* of the oriented segment also called *arrow* or *branch*).

The rule is that after laying down the first segment, the *root branch*, with the endpoint at the origin and otherwise arbitrarily, the others are laid down one after the other by attaching an endpoint of a new branch to an initial point of an old one and leaving free the branch initial point. The set of initial points of the object thus constructed will be called the set of the tree *nodes*. A tree is therefore a partially ordered set with top point the endpoint of the root branch, *also* called the root (which is not a node).

The angles at which the segments are attached will be irrelevant: *i.e.* the operation of changing the angles between arrows emerging from the same node (each arrow carrying along, unchanged, the subtree of arrows possibly attached to its initial point) generates a group of transformations and two trees that can be overlapped by acting on them with a group element are regarded as identical. The number of trees with k branches is thus bounded by $4^k k!$.

With each tree node v we associate an *incoming momentum*, or "decoration", which is simply an integer component vector $\vec{\nu}_v$; with the root of the tree (which is not regarded as a node) we associate a label $j = 1, \ldots, l$.

With each branch $\lambda = v'v$, with final point v' and initial point v, we associate another integer component vector, the *branch momentum* "flowing through the branch", defined by $\vec{\nu}(v) = \sum_{w \le v} \vec{\nu}_w$ (we shall also denote $\vec{\nu}(v)$ by the symbol $\vec{\nu}(\lambda)$). Then, given a positive matrix J and a trigonometric polynomial $f(\vec{\psi}) = \sum_{0<|\vec{\nu}|<N} f_{\vec{\nu}} \cos \vec{\nu} \cdot \vec{\psi}$, $f_{\vec{\nu}} = f_{-\vec{\nu}}$, we consider from now on only decorated trees ϑ with k branches, *such that* $\vec{\nu}(v) \ne \vec{0}$ *for all* v, and associate with each decorated tree the *value*:

$$V_f(\vartheta)_j = -i \prod_{v<r} f_{\vec{\nu}_v} \frac{\vec{\nu}_{v'} \cdot J^{-1} \vec{\nu}_v}{(i\vec{\omega} \cdot \vec{\nu}(v))^2} \tag{1}$$

where v' is the node immediately following v in ϑ; $\vec{\omega} \cdot \vec{\nu}(v)$ will be called the *divisor* of $v'v$; here $\vec{\nu}_r$ denotes the unit vector in the j-th direction $\vec{e}_j$, $j = 1, \ldots, l$. The momentum flowing through the root will be denoted also $\vec{\nu}(\vartheta)$. The *Lindstedt–Newcomb–Poincarè* ("LNP") polynomial $\vec{h}^{(k)}(\vec{\psi})$ is *defined* by $\sum_{\vec{\nu}} \vec{h}^{(k)}_{\vec{\nu}} e^{i\vec{\nu}\cdot\vec{\psi}}$ with:

* Dipartimento di Fisica, Università di Roma, P. Moro 2, 00183 Roma, Italia. Expanded version of a talk at the "Congress on dynamical systems" Montevideo, Uruguay, march 27- april 1, 1995; archived in *mp_arc@ math. utexas. edu*, #95– and in *chao-dyn@ xyz. lanl. gov*, # 9503??, last version available by mosaic: *http://chimera. roma1. infn. it*.

$$\vec{h}^{(k)}_{\vec{\nu},j} = \frac{1}{k!} \sum_{\vartheta,\vec{\nu}(\vartheta)=\vec{\nu}} V_f(\vartheta)_j \tag{2}$$

The main result is, if $|\vec{\omega}\cdot\vec{\nu}|^{-1} < C|\vec{\nu}|^{\tau}$ for some $C,\tau > 0$:

Theorem: *The Lindstedt series is convergent for ε small enough.*

Remark: The first proof (of a more general result) is due to Kolmogorov, [K]; a conceptually new proof came much later and is due to Eliasson, [E], quite difficult to read but definitely *very nice and correct*. Further proofs were given by [G2] (extracted from [G1]) and by [CF1]: somewhat different from the proof in [E] (and remarkably close to each other (except for the size), although independent). The main contribution of the latter papers was to realize that the proof could be greatly simplified by restricting it to a special case. The above case is the simplest and it is discussed in [G1],[G2]: even mild generalizations of the simplest case, [CF1], tend to become quite involved particularly if inappropriate notations are used: a comparative analysis of the different proofs will appear in [GM4].

proof: Let us call a branch $v'v$ of a tree ϑ a *scale n branch*, if $2^{n-1} < C\,|\vec{\nu}(v)\cdot\vec{\omega}| \le 2^n$ where $n = 0,-1,-2,\ldots$ or a scale 1 branch if $C\,|\vec{\omega}\cdot\vec{\nu}(v)| > 1$.

We call a *cluster* of scale n in the tree ϑ a maximal connected set of branches of scale $m \ge n$.

Notable clusters consist of the clusters V with only one entering branch λ_V (and, of course, only one exiting branch λ'_V). If n_V is the scale of the cluster then the scale n_{λ_V} of the branch λ_V is necessarily $n_{\lambda_V} < n_V$. If $v \in V$ call $\vec{\nu}^0(v) = \sum_{w\in V, w\le v} \vec{\nu}_w$.

A cluster V of the above type is called a *resonance* or a *resonant cluster* if $\vec{\nu}(\lambda_V) \equiv \vec{\nu}(\lambda'_V)$ and *also* the number $\mathcal{N}$ of branches inside the cluster is not too large: $\mathcal{N} < \frac{1}{N} 2^{-(n_{\lambda_V}+3)/\tau}$; hence it is so small that $|C\,\vec{\omega}\cdot\vec{\nu}^0(v)| > 2^{n_{\lambda_V}+3}$. A resonant cluster may contain resonant subclusters: we shall denote $\tilde{V}$ the set of nodes in V which are outside the possible resonant subclusters of V, and we call them the *free nodes* of V.

A resonance has two scales attached to it: the scale n_V of the resonant cluster and the scale n_{λ_V} of the branch entering the resonance. We shall distinguish them by calling them the *inner scale* and, respectively, the *outer scale* of the resonance: it is of course $n_V > n_{\lambda_V}$. But the latter inequality can be made stronger: in fact the previous paragraph shows that the scale n_λ of a branch λ in V is determined by the size of the "divisor" $C\,(\vec{\omega}\cdot\vec{\nu}^0(\lambda)+\varepsilon)$ with $C|\varepsilon| < 2^{n_{\lambda_V}}$, so that the size of the divisor is $\ge 2^{n_{\lambda_V}+3} - 2^{n_{\lambda_V}} > 2^{n_{\lambda_V}+2}$, hence:

$$n_\lambda \ge n_V \ge n_{\lambda_V} + 3 \tag{3}$$

for all λ in a resonance V with outer scale n_{λ_V} and inner scale n_V.

Given a branch with endpoint on a node v of a tree ϑ we define the operation of *attaching* or *shifting* the branch to another node w: it consists in detaching *the whole subtree* preceding v (which has v as root) and of attaching it to another node w: the result is a new tree ϑ'.

One then considers all the trees ϑ obtained from a given tree ϑ_0 by one of the following operations:

(i) detaching the endpoint of a branch λ_V entering a resonance V and attaching it (in the above sense) to all possible free nodes $v \in \tilde{V}$: an act that will be called the shift operation on V.
(ii) changing the sign of all (simoultaneously) the incoming momenta $\vec{\nu}_v$ of the free nodes $v \in \tilde{V}$: the parity operation on V.

This produces a family $\mathcal{F}(\vartheta_0)$ of trees containing not too many trees: at most $\prod_V 2m_V$ if m_V is the number of free nodes inside a resonance V. Since $\sum_V m_V \le k$ the number of elements of $\mathcal{F}(\vartheta_0)$ does not exceed $\prod_V 2m_V \le 2^{2k}$, (using $m \le 2^m$).

Furthermore a condition will be imposed on $\vec{\omega}$ (see below) such that given two trees ϑ_0 and ϑ_1 either $\mathcal{F}(\vartheta_0) = \mathcal{F}(\vartheta_1)$ or $\mathcal{F}(\vartheta_0) \cap \mathcal{F}(\vartheta_1) = \emptyset$. This can be done for the following reason.

Consider a branch λ_V with scale n_{λ_V} entering a resonant cluster V with inner scale $n_V > n_{\lambda_V}$ and attached to a node v_1 in V; then if λ_V is shifted to be attached to a node $v_2 \in V$ the momenta flowing through the branches inside V may change by an amount $\vec{\nu}(\lambda_V)$: therefore their scale may change by at most $2^{n_{\lambda_V}}$.

If the resonance V is inside many resonances $V_2, V_3, \ldots$ with outer scales $n_2 > n_3 > ..$ (note that $n_{\lambda_V} > n_2$) then by shifting the branches λ_{V_i} entering V_i in all possible ways to be attached at different free nodes inside V_i the scale of the branches in V cannot change by more than $2^{n_{\lambda_V}} + 2^{n_2} + 2^{n_3} + \ldots < 2^{n_{\lambda_V}+1}$.

Since the number $\mathcal{N} < \frac{1}{N} 2^{-(n_{\lambda_V}+3)/\tau}$ of branches inside the resonance V is not too large the vector $\vec{\nu}^0(v) = \sum_{w \in V, w \le v} \vec{\nu}_w$ is not longer than $2^{-(n_{\lambda_V}+3)/\tau}$ for any $v \in V$. Hence assuming that the following property holds for $\vec{\omega}$:

$$\min_{n \le p \le 0} |C|\vec{\omega} \cdot \vec{\nu}| - 2^p| > 2^{n+1} \qquad \text{if} \quad n \le 0, \ |\vec{\nu}| < 2^{-(n+3)/\tau} \tag{4}$$

we see that the number $x_v = C\vec{\omega} \cdot \vec{\nu}^0(v)$ is not only $> 2^{n_{\lambda_V}+3}$, but also it is further away from the extremes of the interval $I_p = (2^{p-1}, 2^p]$ or, for $p = 1$, $I_1 = (1, +\infty)$ containing x_v, by an amount $2^{n_{\lambda_V}+1}$ at least.

This means that by shifting the branches entering the resonances external to a resonant cluster V of outer scale n_{λ_V} one cannot change the scales of the branches inside V (because for $v \in V$ it is determined by the size of $C\vec{\omega} \cdot \vec{\nu}^0(v)$ plus a quantity which is at most $2^{n_{\lambda_V}+1}$, while the distance of $C\vec{\omega} \cdot \vec{\nu}^0(v)$ from the extremes of the intervals I_p is greater).

The scales of corresponding branches in the elements of a family $\mathcal{F}(\vartheta_0)$ do not change by shifting operations nor they do change by the parity operations, provided Eq. Eq. (4) holds: in particular $2^{n_\lambda - 1} < |C\vec{\omega} \cdot \vec{\nu}^0(\lambda)| \le 2^{n_\lambda}$, for any λ in V. Here two branches of two trees in $\mathcal{F}(\vartheta_0)$ are "corresponding" if they can be superposed by performing suitable sequences of shift or parity transformations in (i), (ii) above. The part of the statement concerning the parity is easily checked, too.

Therefore, assuming Eq. Eq. (4) for the time being (note, however, that almost all $\vec{\omega}$'s verify Eq. Eq. (4)) we shall use the following representation of the function $\vec{h}^{(k)}(\vec{\psi})$:

Lindstedt series representation:

(a) consider all the trees with k branches and collect them into disjoint families: each family is obtained from a tree ϑ in it by the operation of shifting the branches entering a resonance V to the various free nodes in V, or of reversing simoultaneously the signs of the momenta $\vec{\nu}_v$ incoming into the free nodes of V, for all resonances V, independently.

(b) define the value of a family to be the sum of the values of the trees in it.

(c) sum all the values of the families.

This is a natural representation for $\vec{h}^{(k)}$: it was introduced in [G1],[G2] and with a minor variation, based on the slightly different notion of resonance used by [E], in [CF1] (independently).

The above representation seems also close to the one proposed in [GL]. In fact it is probably very natural if one recalls that there is a symplectic structure in the problem that generates the Lindstedt series, see comments (7,8,9) below. Evidence in this direction is provided by the alternative scheme of proof followed by [E], where the bound on the convergence radius of the series is deduced from an argument in some respect simpler (but abstract) than the one below, where the cancellations happening when one sums all the tree values of trees in a given family are checked "indirectly" as a consequence of the symplectic symmetry of the mechanical system whose theory leads to the series, through an idea that goes back to [P], see [GM4].

The proof will now proceed to the easy check that the value of each family is bounded by B^k for some $B > 0$: since there are at most as many families as trees (*i.e.* $\le k!4^k$) this implies the theorem.

Fix a family $\mathcal{F}(\vartheta)$ of trees. And consider the set of the innermost, or *first order*, resonances $\mathcal{V}_1$. Then we shall call $\varepsilon_V = C\vec{\omega} \cdot \vec{\nu}(\lambda_V)$ where $V \in \mathcal{V}_1$ is a resonant cluster with incoming momentum $\vec{\nu}(\lambda_V)$ and we shall sum over the trees obtained from ϑ by the shift and parity operations on the resonances in $\mathcal{V}_1$: if $2m_1 \stackrel{def}{=} \sum_{V\in\mathcal{V}_1} 2m_V$ the number of trees involved in the sum is $\prod_{V\in\mathcal{V}_1} 2m_V \le 2^{2m_1}$, (by $2m < 2^{2m}$).

The "value $F_1(\vartheta)$ of the subfamily" of $\mathcal{F}(\vartheta)$ so obtained *can be regarded* as a function of the complex parameters $\{\varepsilon_V\}$, with V varying in $\mathcal{V}_1$ defined as follows. Note that the momentum flowing in a branch $\lambda = v'v \subset V$ is $\vec\omega\cdot\vec\nu^0(v) + \sigma_v\,\vec\omega\cdot\vec\nu(\lambda_V)$ with $\sigma_v = 0,1$; then the function is constructed by replacing the quantities $(\vec\omega\cdot\vec\nu^0(v) + \sigma_v\,\vec\omega\cdot\vec\nu(\lambda_V))^{-2}$ by $(\vec\omega\cdot\vec\nu^0(v)\lambda + \sigma_v\varepsilon_V)^{-2}$ for all the branches $v'v \subset V$.

Note that this implies that in regarding the value of $\mathcal{F}_1(\vartheta)$ as a function of the ε_V's the divisors associated with the branches entering or exiting the resonance V *are not* represented as ε_V^{-2}.

It is now easy to see that the value of a family is holomorphic in the variables ε_V in the disks $|\varepsilon_V| < 2^{n_V-3}$. This is because in any tree of the family it is $|C\,\vec\omega\cdot\vec\nu^0(v)| > 2^{n_V-1}$ (by the remarks following Eq. Eq. (4)) so that for $|\varepsilon_V| < 2^{n_V-3}$ it is $|C\,\vec\omega\cdot\vec\nu^0(v) + \sigma\varepsilon_V| > 2^{n_V-3}$. In fact, for later use, one sees that the last inequality holds even in the larger disk $|\varepsilon_V| < 2^{n_V-3}(1+\frac{1}{7})$ (the $\frac{1}{7}$ will arise later as the sum of the geometric series with ratio $\frac{1}{8}$).

This means that the value $F_1(\{\varepsilon_V\}_{V\in\mathcal{V}_1})$ of the subfamily $\mathcal{F}_1(\vartheta)$ can be bounded above, for $|\varepsilon_V| < 2^{n_V-3}$ (and in fact even for $|\varepsilon_V| < 2^{n_V-3}\frac{8}{7}$), by trivially bounding the divisors in the contributions of the various trees in terms of their scales, or by using $|C\,\vec\omega\cdot\vec\nu^0(v) + \sigma\varepsilon_V| \ge 2^{n_V-3}$. One finds therefore, from Eq. Eq. (1) , the bound:

$$2^{2m_1} X \stackrel{def}{=} 2^{2m_1}\frac{1}{k!}\Big(\frac{f_0 C^2 N^2}{J_0}\Big)^k \prod_\lambda 2^{-2(n_\lambda-3)} = 2^{2m_1}\Big(\frac{f_0 2^6 C^2}{J_0}\Big)^k \prod_n 2^{-2nN_n} \tag{5}$$

if N_n is the number of branches with scale n, $f_0 = \max|f_{\vec\nu}|$ and J_0 is a lower bound to J.

It is easy to check that the function $F_1(\{\varepsilon_V\})$ vanishes in *each* ε_V to *second order* for $\varepsilon_V = 0$. This is simply because the sum of the momenta $\vec\nu_{v_j}$ incoming into the nodes $v_j \in \tilde V$, $\sum_j \vec\nu_{v_j}$, in each resonance V, vanishes, (by definition of resonance). Furthermore the only variation of the value of a tree when the branch entering a resonance is shifted to the node v_j from the node $v_{j'}$ is the replacement of a factor $\vec\nu_{v_{j'}}\cdot J^{-1}\vec\nu'$, for some $\vec\nu'$, with the factor $\vec\nu_{v_j}\cdot J^{-1}\vec\nu'$, when $\varepsilon_V = 0$. Hence if, $\varepsilon_V = 0$ for some V, $F_1(\{\varepsilon_V\})$ is proportional to $\sum_{v\in\tilde V}\vec\nu_v = \vec 0$ and vanishes to first order in each of the ε_V. The summation over the sign reversals turns the function $F_1(\{\varepsilon_V\})$ into an even function of the ε_V, hence vanishing to second order.

Therefore the above bound Eq. Eq. (5) in the domain $|\varepsilon_V| < 2^{n_V-3}$ implies, by the maximum principle, that the bound can be improved by a factor $\prod_{V\in\mathcal{V}_1}\big(\frac{|\varepsilon_V|}{2^{n_V-3}}\big)^2$; which becomes, if one replaces ε_V by its actual value:

$$Y_1 \stackrel{def}{=} \prod_{V\in\mathcal{V}_1} 2^{2(n_{\lambda_V}-n_V+3)} \qquad \text{if} \quad n_{\lambda_V} \le n_V - 3 \tag{6}$$

Note that only the case $n_{\lambda_V} < n_V - 3$ needs the maximum priciple.

We now consider the *second order resonances, i.e.* the set of resonances $\mathcal{V}_2$ which only contain first order resonances. Then we denote again $\varepsilon_V = C\vec\omega\cdot\vec\nu(\lambda_V)$ where $V \in \mathcal{V}_2$ is a resonant cluster with incoming momentum $\vec\nu(\lambda_V)$ and we shall set $2m_2 = \sum_{V\in\mathcal{V}_2} 2m_V$, if m_V is the number of free nodes in V.

Summing over the $< 2^{2m_1+2m_2}$ trees obtained from ϑ by the shift and parity operations on the resonances in $\mathcal{V}_1, \mathcal{V}_2$ we define the "value $F_2(\vartheta)$ of the subfamily" of $\mathcal{F}_2(\vartheta)$ so obtained. It can be regarded as a function of the complex parameters $\{\varepsilon_V\}$, with V varying in $\mathcal{V}_2$. The momentum flowing in a branch $\lambda = v'v \subset V$ is $\vec\omega\cdot\vec\nu^0(v) + \sigma_v\vec\omega\cdot\vec\nu(\lambda_V)$ with $\sigma_v = 0,1$ and the function is constructed, as in the previous case, by replacing the quantities $(\vec\omega\cdot\vec\nu^0(v) + \sigma_v\vec\omega\cdot\vec\nu(\lambda_V))^{-2}$ by $(\vec\omega\cdot\vec\nu^0(v)\lambda + \sigma_v\varepsilon_V)^{-2}$ for all $v'v \subset V$.

Furthermore when the variable ε_V varies in the disk $|\varepsilon_V| < 2^{n_V-3}$ the divisors of the branches entering the first order resonances $W \subset V$ have the form $\vec\omega\cdot\vec\nu^0(\lambda) + \sigma_v\vec\omega\cdot\vec\nu(\lambda_V)$ with $2^{n_\lambda-1} < C|\vec\omega\cdot\vec\nu^0(v)| \le 2^{n_\lambda}$ and therefore if $n_V < n_W - 3$ they are bounded above by $2^{n_V} + 2^{n_V-3} \le 2^{n_W-4} + 2^{n_W-6} < 2^{n_W-3}$.

Hence if $n_V < n_W - 3$ the value of the function is bounded by $2^{2m_2+2m_1} Y_1 X$, for $|\varepsilon_V| < 2^{n_V-3}$: this remains true even if $n_V = n_W - 3$ for one or more $W \subset V$. In the latter cases, in fact, the momentum entering the resonance W needs not be $< 2^{n_W-3}$ but it is $\le 2^{n_V} + 2^{n_V-3} = 2^{n_W-3} + 2^{n_W-6} < 2^{n_V-3}\frac{8}{7}$ so that the branches λ inside W are still such that $C|\vec{\omega}\cdot\vec{\nu}^0(\lambda) + \sigma_v \varepsilon_W| > 2^{n_W-1} - 2^{n_W-3} - 2^{n_W-6} \ge 2^{n_W-3}$ and taking $Y_1 = 1$ provides a correct bound.

Again the value $F_2(\vartheta)$ vanishes to second order in the variables ε_V so that it can be bounded, if one replaces ε_V by its actual value, by $2^{m_2+m_1} Y_1 Y_2 X$ where Y_2 is defined as in Eq. Eq. (6) with $\mathcal{V}_2$ replacing $\mathcal{V}_1$, *i.e.* $\prod_{V\in\mathcal{V}_2} 2^{2(n_{\lambda_V}-n_V+3)}$. One now considers the third order resonances and so on until one finds the bound on the value of the complete family $\mathcal{F}(\vartheta)$ given by $2^{2m_1+2m_2+\cdots} Y_1 Y_2 \ldots X$. Since $2m_1 + \ldots \le 2k$ one finds, recalling the definition of X in Eq. Eq. (5) :

$$\frac{1}{k!}\Big(\frac{f_0 2^6 C^2}{J_0}\Big)^k \prod_n 2^{-2nN_n} \prod_{n\le 0} \prod_{T, n_T=n} \prod_{i=1}^{m_T} 2^{2(n-n_i+3)} \tag{7}$$

if T are the clusters of the family of trees under consideration, m_T is the number of maximal resonances contained inside the cluster T, and n_i denotes the scale of the branch entering the i-th of such resonances.

The number N_n of branches of scale n can be bounded by imitating the key bound due to Siegel and Brjuno in the case of trees without resonances (the bound seems to have been well known since Siegel's theorem for the similar case of the linearization of the complex maps, which essentially corresponds to the analysis of the generating function of polynomials like the ones considered here and generated by resonanceless trees).

If m_T is the number of maximal resonances inside the cluster T then N_n is bounded by:

$$N_n \le 4N 2^{(n+3)/\tau} k + \sum_{T, n_T=n} (m_T - 1) \tag{8}$$

the proof in [G2] is reproduced, for completeness, in the appendix.

Hence the product in Eq. Eq. (7) is bounded by combining Eq. Eq. (7) and Eq. Eq. (8) , after some simple power counting, by:

$$\frac{1}{k!}\Big(\frac{f_0 2^{12} C^2 N^2}{J_0}\Big)^k \prod_{n=-\infty}^{1} 2^{-8nN2^{(n+3)/\tau}k} = \frac{1}{k!} B_0^k \tag{9}$$

with $B_0 = C^2 f_0 J_0^{-1} N^2 2^{12} 2^{-\sum_n 8nN2^{(n+3)/\tau}}$.

The number of families cannot exceed the number of trees, *i.e.* $k!2^{2k}$, so that the series Eq. Eq. (1) converges for $|\varepsilon| < (4B_0)^{-1}$, which is even a rather good estimate, at least as a first estimate, for N small.

This completes the proof in the case in which the inequalities Eq. Eq. (4) hold. Such inequalities were called *strong diophantine property* in [G1], [G2]. If one examines carefully their role (*i.e.* keeping the scales unchanged while performing the operations (i), (ii)) one sees that it is very reasonable that they can be replaced by a suitable similar property which holds in the general case in which only the diophantine inequality (following Eq. Eq. (2) above) holds: this is discussed at the end of [G2] and proved in [GG].

Comments:

(1) The restriction that f is a trigonometric polynomial can be lifted easily to treat analytic f's, as shown in [GM2], (in [CF1] too the f is not supposed a polynomial). The restriction that f is even can be lifted simply by replacing the parity operation by the *new* operation of shifting the node to which the *outgoing branches* are attached in $\tilde{V}$, as shown in [CF1]. Note that the even case is extremely interesting for the applications as the perturbation functions f, in many cases, are even because this property is related to the time reversal symmetry (see [BCG]).

(2) The above method can be extended to study the generating functions of quantites originated in a way analogous to the Lindstedt polynomials in the perturbation theory of whiskered tori, see [G1]: the first such results are in [Ge1], [Ge2]. In the later review paper [CF2], aware of such references, a part of such results are also discussed and apparently claimed as new, if so improperly because the analysis in [Ge1],[Ge2] is correct and complete (and, furthermore, it goes quite far beyond what is discussed in [CF2] in the cases common to both papers).

(3) The method of [E] is not quite the same as the above, see [GM4].

(4) The above method reminds very much of field theory methods: the similarity was noted first in [FT], at a formal level, without taking advantage of it to study a proof of the theorem. The results of [GM1], [GM2] take full advantage of the similarity and build a proof (see comment (9) below).

(5) In [G3] it is shown that there is a euclidean field theory whose one point Schwinger function is the function Eq. Eq. (2) . This analogy is pushed further in [GGM], where a connection with the theory of the singularity in ε of the function Eq. Eq. (1) as ε grows is heuristically attempted, along an early suggestion in [PV].

(6) the convergence of the Lindstedt series generated by the general KAM theorem can be proved along the above lines, and as one should expect, it does not involve any new ideas, in fact see [CF2] for a somewhat more detailed sketch and [GM3] for a proof.

(7) there are other ways besides the one of [E] to reach a representation similar to the above and giving immediately an absolutely convergent series for the generating function of the LNP polynomial. This has been shown in [GL] and it seems closely related to the fact that the problem is associated with the theory of hamiltonian evolutions. Also the method of [E] does not perform explicit cancellations on individual terms of the series, but it shows that they must happen because of the symplectic nature of the equations that the convergence of the LNP series would solve; hence it would be interesting to establish a clear connection between the method in [GL] and that of [E]: this may help establishing a more clear connection also between the above method and [G1],[G2] or [CF1].

(8) one can ask whether the above proof can be used for numerical purposes. At first sight one may think that the maximum principle used in the proof is an obstacle. However one should note that the sum over the tree values in a given family is an algebraic operation: hence the analytic estimate can be replaced by an algebraic one (this is exploited in [GM1], [GM2]).

When one performs the sum corresponding to the shift of the incoming branch, or to the parity transformation, for a resonance V, one takes a linear combination of $m_V - 1$ products of divisors. Therefore the result is a polynomial in ε_V. The above remarks show that it starts at second order: hence one can factor ε_V^2 (*by a finite number of algebraic operations*) which cancels one of the two divisors associated with the resonance incoming and outgoing branches.

This means that each resonance only contributes a divisor not exceeding $\varepsilon_V^{-2}2^{-2n_V} \le 2^{-2n_{\lambda_V}-2n_V}$, in the estimates, *instead of* $\varepsilon_V^{-4} \le 2^{-4n_{\lambda_V}}$: in other words the "extra" divisor on scale n_{λ_V} is replaced by a divisor on the generally much higher scale n_V (hence much larger). It is not difficult to see that the phenomenon of accumulation of divisors expressed by the sum in Eq. Eq. (8) is precisely due to the fact that without taking into account the above remark the resonances contribute *two* divisors on the same scale (*i.e.* they generate a divisor ε_V^{-4}). Hence once the above algebraic operation is performed one has no more divisors in excess, in comparison with the Siegel's situation: this idea could also be used to provide a quick proof of Eq. Eq. (8) if one assumes it to be valid for resonanceless trees, a case in which it was well known being implicit in Siegel's and Brjuno's work.

The remarkable and important result in [GL], that justifies the apparent absence of a cancellation analysis, is that the above collection of terms is automatically performed if one proceeds by keeping always explicitly into account the symplectic structure of the problem that is solved by the Lindstedt series. In this respect the work [GL] is much closer to [E] than [G1],[G2],[CF1].

(9) of course one could define the notion of resonant cluster by forgetting the condition on the number of branches: this would lead, if suitably complemented with a redefinition of the families of trees, to an *overcompensation* and to worse final estimates (and to a trivially more involved proof). This is done in [E], and in some way also in [GL] (but not in [G1],[G2],[CF1]). On the other hand it can provide a conceptual simplification as it has been shown in [GM2],[GM3] where the possibility

of overcompensating is even exploited to use deeply the analogies of the above ideas with those of renormalization theory in quantum field theory (see (4) above). The result is what I think is the most transparent and elementary proof of the KAM theorem, even simpler than the one decribed in the present paper (although the latter is likely to be better for numerical applications).

(10) finally it is important to note that the results depend only on the smallest eigenvalue J_0 of the matrix J and *not* on the maximum. Given the hamiltonian $H = \vec\omega\cdot\vec A + \frac{1}{2}\vec A J^{-1}\vec A + \varepsilon\varphi(\vec\alpha)$ where $\vec A\in R^l$ and $\vec\alpha$ is in the torus T^l (*Thirring model*, see [G2]), the Lindstedt series solves, see [G2], the problem of the existence of a KAM torus close to $\vec A = \vec 0$ and with rotation vector $\vec\omega$: hence we see that the condition on the size of ε is *independent* on the *twist rate*, which is the *maximum* eigenvalue of J. In the latter model the twist condition, that plays such an important role in the general KAM theorem, is *not necessary*: thus tori arising from the above analysis were called, in [G1],[G2], *twistless KAM tori*.

Appendix A1: Resonant Siegel-Brjuno bound.

Calling N^*_n the number of branches of scale $\le n$ which are *not* entering a resonance ("non resonant branches"). We shall prove first that $N^*_n \le 2k(N2^{-(n+3)/\tau})^{-1}-1$ if $N_n>0$. We fix n and denote N^*_n as $N^*(\vartheta)$. Let $\varepsilon=\frac{1}{\tau}$ and $E=N^{-1}2^{-3\varepsilon}$.

If the root branck scale of ϑ is $>n$ and if $\vartheta_1,\vartheta_2,\ldots,\vartheta_m$ are the subtrees of ϑ attached to the root branch of ϑ and with $k_j > E\,2^{-\varepsilon n}$ branches, it is $N^*(\vartheta)=N^*(\vartheta_1)+\ldots+N^*(\vartheta_m)$ and the statement is inductively implied from its validity for $k'<k$ provided it is true that $N^*(\vartheta)=0$ if $k<E2^{-\varepsilon n}$, which is is certainly the case if E is chosen as above.[1]

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

If $m=1$ we once more have a trivial case unless the order k_1 of ϑ_1 is $k_1>k-\frac{1}{2}E\,2^{-n\varepsilon}$. Finally, and this is the real problem as the analysis of a few examples shows, we claim that in the latter case the root line of ϑ_1 is either a resonant line or it has scale $>n$.

Accepting the last statement it will be: $N^*(\vartheta)=1+N^*(\vartheta_1)=1+N^*(\vartheta'_1)+\ldots+N^*(\vartheta'_{m'})$, with ϑ'_j being the m' subdiagrams emerging from the first node of ϑ'_1 with orders $k'_j > E\,2^{-\varepsilon n}$: this is so because the root line of ϑ_1 will not contribute its unit to $N^*(\vartheta_1)$. Going once more through the analysis the only non trivial case is if $m'=1$ and in that case $N^*(\vartheta'_1)=N^*(\vartheta"_1)+\ldots+N^*(\vartheta"_{m"})$, *etc*, until we reach a trivial case or a diagram of order $\le k-\frac{1}{2}E\,2^{-n\varepsilon}$.

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

Suppose that the root line of ϑ_1 has scale $\le n$ and is not entering a resonance. Note that $|\vec\omega\cdot\vec\nu(v_0)|\le 2^n, |\vec\omega\cdot\vec\nu(v_1)|\le 2^n$, if v_0,v_1 are the first vertices of ϑ and ϑ_1 respectively. Hence $\delta\equiv|(\vec\omega\cdot(\vec\nu(v_0)-\vec\nu(v_1))|\le 2\,2^n$ and the diophantine assumption implies that $|\vec\nu(v_0)-\vec\nu(v_1)|>(2\,2^n)^{-\tau^{-1}}$, or $\vec\nu(v_0)=\vec\nu(v_1)$. The latter case being discarded as $k-k_1<\frac{1}{2}E\,2^{-n\varepsilon}$ (and we are not considering the resonances: note also that in such case the lines in ϑ/ϑ_1 different from the root of ϑ must be inside a cluster, see footnote 3)), it follows that $k-k_1<\frac{1}{2}E\,2^{-n\varepsilon}$ is inconsistent: it would in fact imply that $\vec\nu(v_0)-\vec\nu(v_1)$ is a sum of $k-k_1$ vertex modes and therefore $|\vec\nu(v_0)-\vec\nu(v_1)|<\frac{1}{2}NE\,2^{-n\varepsilon}$ hence $\delta>2^3\,2^n$ which is contradictory with the above opposite inequality.

The total number of branches with scale n is therefore $N^*_n+\sum_{T,n_T=n} m_T$ if m_T is the number of resonances contained in T. A similar, far easier, induction can be used to prove that if $N^*_n>0$ then the number p of clusters of scale n verifies the bound $p\le 2k\,(E2^{-\varepsilon n})^{-1}-1$; it will be left out (see [GG]). Thus Eq. (8) is proved.

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

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

Acknowledgements: I am indebted to G. Gentile and V. Mastropietro for many discussions and for explaining important details of the work of Eliasson, that I had missed, and for clarifying comments on his work as well as on their own work. I am grateful to G. Gentile for his careful comments on the details of the paper. This work is part of the research program of the European Network on: "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460.

References

[BCG] Benettin, G., Carati, A., Gallavotti, G.: *A rigorous implemenation of the Jeans–Landau–Teller approximation for adiabatic invariants*, archived in *mp_arc@ math. utexas. edu*, # 95–???.
[CF1] Chierchia, L., Falcolini, C.: *direct proof of a theorem by Kolmogorov in hamiltonian systems*, Annali della Scuola Normale Superiore di Pisa, **21**, 541–593, 1994.
[CF2] Chierchia, L., Falcolini, C.: *Compensations in small divisor problems*, Archived in *mp_arc@ math. utexas. edu*, #94-270.
[E] Eliasson L. H.: *Hamiltonian systems with linear normal form near an invariant torus*, ed. G. Turchetti, Bologna Conference, 30/5 to 3/6 1988, World Scientific, 1989. And *Generalization of an estimate of small divisors by Siegel*, ed. E. Zehnder, P. Rabinowitz, book in honor of J. Moser, Academic press, 1990. But mainly: *Absolutely convergent series expansions for quasi–periodic motions*, report 2–88, Dept. of Math., University of Stockholm, 1988.
[FT] Feldman, J., Trubowitz, E.: *Renormalization in classical mechanics and many body quantum field theory*, Journal d'Analyse Mathématique, **58**, 213-247, 1992.
[G1] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable hamiltonian systems. A review.*, Reviews on Mathematical Physics, **6**, 343– 411, 1994.
[G2] Gallavotti, G.: *Twistless KAM tori*, Communications in Mathematical Physics, **164**, 145–156, 1994.
[G3] Gallavotti, G.: *Perturbation Theory*, in "Mathematical physics towards the XXI century", p. 275–294, ed. R. Sen, A. Gersten, Ben Gurion University Press, Ber Sheva, 1994.
[GG] 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, in print on Ergodic Theory and Dynamical Systems with the title *Majorant series convergence for twistless KAM tori.*
[Ge1] Gentile, G.: *A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable systems*, preprint Roma "La Sapienza" n. 996, february 1994, to appear in Forum Mathematicum. archived in *mp_arc@ math. utexas. edu*, #94-33, in print on Forum Mathematicum.
[Ge2] Gentile, G.: *Whiskered tori with prefixed frequencies and Lyapunov spectrum*, preprint Roma "La Sapienza" n. 1015, april 1994. archived in *mp_arc@ math. utexas. edu*, #94-92.
[GL] Giorgilli, A., Locatelli, U.: *Kolmogorov theorem and classical perturbation theory*, preprint Dipartimento Matematica, Milano, january 1995.
[GM1] Gentile, G., Mastropietro, V.: *KAM theorem revisited*, archived in *mp_arc@ math. utexas. edu*, #94-403.
[GM2] Gentile, G., Mastropietro, V.: *Tree expansion and multiscale analysis for KAM tori*, archived in *mp_arc@ math. utexas. edu*, #94-405.
[GM3] Gentile, G., Mastropietro, V.: in preparation, preprint, 1995.
[GM4] Gentile, G., Mastropietro, V.: in preparation, preprint, 1995.
[GGM] Gallavotti, G., Gentile, G., Mastropietro, V.: *Field theory and KAM theorem*, Archived in *mp_arc@ math. utexas. edu*, #95-???.
[K] Kolmogorov, N.: *On the preservation of conditionally periodic motions*, Doklady Akademia Nauk SSSR, **96**, 527– 530, 1954. See also: Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.M.: *A proof of Kolmogorov theorem on invariant tori using canonical transormations defined by the Lie method*, Nuovo Cimento, **79 B**, 201– 223, 1984.

[Pö] Pöschel, J.: *Invariant manifolds of complex analytic mappings*, Les Houches, XLIII (1984), vol. II, p. 949– 964, Ed. K. Osterwalder, R. Stora, North Holland, 1986.
[PV] Percival, I., Vivaldi, F.: *Critical dynamics and diagrams*, Physica D, **33**, 304– 313, 1988.
[P] Poincaré, H. *Les méthodes nouvelles de la mécanique céleste*, vol. II, ch. 9, 1893.