JB

METHODS IN THE THEORY OF QUASI PERIODIC MOTIONS

Giovanni GALLAVOTTI

swg728

Institut des Hautes Études Scientifiques
35, route de Chartres
91440 – Bures-sur-Yvette (France)

Mars 1997

IHES/P/97/30

Methods in the theory of quasi periodic motions

Giovanni Gallavotti

Fisica, Università di Roma, P.le Moro 2, 00185 Roma, Italy
I.H.E.S., 91440 Bures s/Yvette, France

Abstract. Recent results on the theory of the Hamilton-Jacobi equation and the regularity of their solutions, in spite of the non regularity of the data, are reviewed, and discussed with attention to the cancellation mechanisms that make regularity possible.

§1 The Hamilton Jacobi equation.

This review deals on the work developed in collaboration with F. Bonetto, G. Gentile, V. Mastropietro.

The Hamilton Jacobi equation for an invariant torus in a ℓ dimensional hamiltonian system close to an integrable system takes one of the forms:

$$\begin{aligned}(\omega_0 \cdot \partial)^2 \mathbf{h}(\psi) &= \varepsilon (\partial f)(\psi + \mathbf{h}(\psi)) \\ (\omega_0 \cdot \partial) \mathbf{h}(\psi) &= \varepsilon \mathbf{f}(\psi + \mathbf{h}(\psi)) + \mathbf{N}\end{aligned}\tag{1.1}$$

where ψ is a point on the ℓ -dimensional torus, $\mathbf{h}(\psi)$ a R^ℓ -valued function, $f, \mathbf{f}$ are functions on T^ℓ , $\mathbf{N}$ is a vector in R^ℓ and ω_0 is a diophantine vector in R^ℓ such that, considering the integer components lattice Z^ℓ , there are constants $C, \tau > 0$ with:

$$C|\omega_0 \cdot \nu| \geq |\nu|^{-\tau} \quad \forall \nu \in Z^\ell, \nu \neq 0\tag{1.2}$$

where $|\nu| = \sqrt{\nu_1^2 + \dots}$. The gradient ∂f is evaluated first and then computed at the point $\psi + \mathbf{h}(\psi)$ (i.e. it is not the gradient of $f(\psi + \mathbf{h}(\psi))$).

The first of (1.1) corresponds to the construction of KAM tori in the case of the Thirring models (a general class of hamiltonian systems, [1]), and the second corresponds to the KAM tori for perturbations of non resonant harmonic oscillators, [2], [3]: in the latter case not only $\mathbf{h}$ is unknown but also $\mathbf{N}$, called a *counterterm* because one can think of it as a quantity that must be fixed in order to make the equations soluble.

We shall not be concerned here with the physical interpretations of (1.1).

It was pointed out explicitly by Moser that the second equation is in some sense the fundamental one. Its theory is essentially equivalent to the KAM theorem theory, [2].

Here we shall consider a wider class of equations:

$$\begin{aligned}
(1) \quad Dh(\psi) &= \varepsilon (\partial f)(\psi + h(\psi)) \\
(2) \quad Dh(\psi) &= \varepsilon f(\psi + h(\psi)) + N
\end{aligned} \tag{1.3}$$

where D is a pseudodifferential operator like, respectively:

$$\begin{aligned}
(1) \quad D &= (\omega_0 \cdot \partial)^2, \quad \text{or} \quad D = (-\Delta)^r \quad r \geq 0 \\
(2) \quad D &= (\omega_0 \cdot \partial), \quad \text{or} \quad D = (-\Delta)^r L \quad r \geq 0
\end{aligned} \tag{1.4}$$

where Δ is the Laplace operator and L is a linear elliptic operator ("Dirac operator": for instance $\gamma \cdot \partial$ where γ are the Dirac's *gamma matrices*). The cases $r = 0$ in the first equation in (1.3) are remarkable as they are *not* differential equations.

It is easy to show by using the contraction method that (1.3), (1.4) with $r \geq 0$ do have a solution h or, respectively, h, N for small ε if f is regular enough. This remains true, but it is far less easy, in the two remaining cases, in which D is a power of $(\omega_0 \cdot \partial)$:

- The classical results that hold when f is assumed analytic on T^ℓ give analyticity in ε, ψ of the solution h and N , for ε small, [2].
- In the differentiable case the first equation with $r = 0$ admits a solution in class $C^{(p-1)}(T^\ell)$ if $f \in C^{(p)}(T^\ell)$ and $p > 1$: this is an elementary result (contraction principle) for ε small. The solution is also $p-1$ times differentiable in ε .
- In the case $D = (\omega_0 \cdot \partial)^2$ and $f \in C^{(p)}(T^\ell)$ it admits a solution in class $C^{(p-4\tau-2)}$ if $p > 4\tau + 2$, a consequence of [4], where the contraction method is "replaced" by Nash's implicit functions theorem. The solution is also smooth in ε .
- In the case $D = \omega_0 \cdot \partial$ and $f \in C^{(p)}(T^\ell)$ it admits a solution in class $C^{(p-2\tau-2)}$ if $p > 2\tau + 2$, as a consequence of Parasiuk's work, [5], see [6], based on Moser's technique; the solution is smooth in ε .

We address the question of the regularity in ε of the latter solutions. From the quoted results only differentiability in ε , to an order related to p , can be deduced.

We consider functions f, f which are finitely differentiable but that have the special form: $f = (f_1, \dots, f_\ell)$, with each f_j of the class $\hat{C}^{(p)}(T^\ell)$ introduced in [7], for some p .

Namely the class $\hat{C}^{(p)}$ consists of the R^ℓ valued functions such that $f(\psi) = \sum_{\nu \in \mathbb{Z}^\ell} f_\nu e^{i\nu \cdot \psi}$, $f_\nu = f_{-\nu}$, with $f_0 = 0$ and for $\nu \neq 0$,

$$f_\nu = \sum_{n \geq p+\ell}^N \frac{P_{2k}^{(n)}(\nu)}{|\nu|^{n+2k}} \tag{1.5}$$

for some $N \geq p + \ell$ and some even polynomials $P_{2k}^{(n)}(\nu)$ of degree $2k$.

For instance we can choose $f_{j\nu} = a_j |\nu|^{-b}$, with $b = p + \ell$ and $\mathbf{a} = (a_1, \dots, a_\ell)$ a vector in R^ℓ . In the following we shall deal explicitly with such a function: the analysis can be trivially extended to the class of functions (1.5).

Hence we discuss (1.3) only in cases in which f or, respectively $\mathbf{f}$ are special even functions. The results that follow do not hold in general for non even functions, see [7] for some simple counterexamples.

A solution of the second of (1.3) will be any pair $\mathbf{h} \in C^{(0)}(T^\ell)$ and $\mathbf{N} \in R^\ell$ that satisfies the identity:

$$-\int_{T^\ell} D\mathbf{F}(\psi) \cdot \mathbf{h}(\psi) d\psi = \int \mathbf{F}(\psi) \cdot (\varepsilon \mathbf{f}(\psi + \mathbf{h}(\psi)) + \mathbf{N}) d\psi \quad (1.6)$$

for all $\mathbf{F} \in C^\infty(T^\ell)$. Similarly we define solutions of the first of (1.3).

In [7],[8] the following theorem is proved.

Theorem: Equations (1.3) admit solutions that are analytic in ε for ε small if f or, respectively, $\mathbf{f}$ are in $\hat{C}^{(p)}(T^\ell)$ and p is large enough. For instance:

- (i) if $D = (\omega_0 \cdot \partial)^2$ the first of (1.3) admits a $C^{(<p-6\tau-3)}(T^\ell)$ solution if $f \in \hat{C}^{(p)}(T^\ell)$ provided $p - 6\tau - 3 > 0$.
- (ii) if $D = \omega_0 \cdot \partial$ the second of (1.3) admits a $C^{(<p-3\tau-2)}(T^\ell)$ solution if $f \in \hat{C}^{(p)}(T^\ell)$ provided $p - 3\tau - 2 > 0$.
- (iii) if $D = 1$ the first of (1.3) admits a $C^{(<p-2)}(T^\ell)$ solution if $f \in \hat{C}^{(p)}(T^\ell)$ provided $p - 2 > 0$.

The analyticity in ε is perhaps unexpected even in the case $D = 1$ since the solutions $\mathbf{h}(\psi)$ are *not* analytic in ψ . Furthermore it is not a general property valid whenever the equations admit a solution, because of the mentioned counterexamples.

One can prove the result because there are remarkable cancellations that we call *ultraviolet cancellations*, for reasons that become clear if one considers the analogy between the KAM problem and field theory problems, see [9]. The cancellations are *not* the ones pointed out by Eliasson, [10], that are specific to the KAM problems and that permit to derive the KAM theorem by a majorant series method: the latter cancellations are called *infrared cancellations* and have been recently revisited in several papers. The papers reproduce Eliasson's work with minor changes or additions but treat the simplest forms of the problem thus providing a simple understanding of the new methods (see for instance [1] which has the advantage of being short but uncompromising with the key difficulties; further developments can be found, *along the same lines*, in [11], [12], [13], [9] where also the relevant references to other papers can be found).

The names used for the two types of cancellations are taken from Quantum Fields theory which leads to problems, and techniques for their solutions, that are very close to the ones used in the theory of the Lindstedt series, see [9].

In §2 we discuss the Lindstedt formal series solution of (1.3): we shall deal with the second of (1.3) with $D = \omega_0 \cdot \partial$ as it is simpler than the first of (1.3) with $D = (\omega_0 \cdot \partial)^2$. Actually the simplest problem is the first equation in (1.3) with $D = 1$ or $D = (-\Delta)^r$: the case $D = 1$ is *non trivial* and it is not chosen to illustrate the analysis only because we also want to make clear that the small divisors can be treated. But the ideas and methods apply to all cases. In §3 we discuss the cancellation mechanism through some simple examples.

Note that the above quoted results of Moser give existence for a wider range of p 's in the case of (1.1) while the results of [5], see [6], give existence in a wider range of p 's if $\ell > 2$, but in a different range if $\ell = 2$ (the results cannot be really compared, even for $\ell = 2$, because $\hat{C}^{(p)} \subset C^{(p)}$).

§2 Lindstedt series.

Consider the second of Eq. (1.3) with $f_{j\nu} = a_j |\nu|^{-b}$ replaced by $f_{j\nu} e^{-\kappa|\nu|}$ and we write $\mathbf{h}(\psi) = \sum_{k=1}^{\infty} \varepsilon^k \mathbf{h}^{(k)}(\psi)$. The parameter κ is taken $\kappa > 0$ and, after computing the coefficients $\mathbf{h}_{\nu}^{(k)}$ of the Fourier series of the function $\mathbf{h}^{(k)}(\psi)$ which will depend on κ , we shall perform the limit $\kappa \rightarrow 0$ (*Abel's summation*). This is done because, contrary to what one could fear, the limit is well defined, if b is large enough, for all k .

It is easy to find the coefficients via a set of rules that associate with $\mathbf{h}^{(k)}$ a family of "Feynman graphs" ϑ and assign to each graph a *value* which is a function of $\nu \in \mathbb{Z}^{\ell}$ so that summing all values, at fixed ν , of the graphs that are associated with $\mathbf{h}^{(k)}$ one gets the Fourier component of order ν of $\mathbf{h}^{(k)}$.

The graphs have the form of a tree graph (a very unusual situation for the Feynman graphs that arise in quantum field theory) like:

Fig.1. A graph ϑ with $p_{v_0} = 2, p_{v_1} = 2, p_{v_2} = 3, p_{v_3} = 2, p_{v_4} = 2$ and $k = 12, \prod p_v! = 2^4 \cdot 6$, and some decorations. The line numbers, distinguishing the lines, and the arrows, pointing at the root, are not shown. The lines length should be 1 but it is drawn of arbitrary size.

We lay down one after the other, on a plane, k pairwise distinct unit segments oriented from one extreme to the other: respectively the *initial point* and the *endpoint* of the oriented segment. The oriented segment will also be called *arrow*, *branch* or *line*. The segments are supposed to be numbered from 1 to k .

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 by leaving free the new branch initial point. The set of initial points of the object thus constructed will be called the set of the graph *nodes* or *vertices*. A graph of order k is therefore a partially ordered set of k nodes with top point the endpoint of the root branch, also called the *root* (which is not a node); in general there will be several “bottom nodes” (at most $k - 1$ and 7 in Fig.1).

We denote by $\leq$ the ordering relation, and say that two nodes v, w are “comparable” if $v < w$ or $w < v$.

With each graph node v we associate an *external momentum* or *mode* which is simply an integer component vector $\nu_v \neq 0$; with the root of the graph (which is not regarded as a node) we associate a label $j = 1, \dots, \ell$.

For each node v , we denote by v' the node immediately following v and by $\lambda_v \equiv v'v$ the branch connecting v to v' , (v will be the initial point and v' the endpoint of λ_v). If v is the node immediately preceding the root r (*highest node*) then we shall write $v' = r$, for uniformity of notation (recall that r is not a node).

We consider “comparable” two lines λ_v, λ_w , if v, w are comparable.

If p_v is the number of branches entering the node v , then each of the p_v branches can be thought as the root branch of a *subgraph* having root at v : the subgraph is uniquely determined by v and one of the p_v nodes w immediately preceding v . Hence if $w' = v$ it can be denoted ϑ_{vw} .

We shall call *equivalent* graphs which can be overlapped by

- (1) changing the angles between branches emerging from the same node, or
 - (2) permuting the subgraphs entering into a node v
- in such a way that all the labels match.

The number of (non equivalent numbered) graphs with k branches is bounded by $4^k k!$, [14].

The “value” of a graph ϑ is now given by

$$\text{Val}(\vartheta) = \prod_{v < r} \frac{(i\nu_{v'} \cdot \mathbf{f}_{\nu_v})}{i\omega_0 \cdot \nu_{\lambda_v}}, \quad (2.2)$$

where

- (1) v' is the node immediately following v in ϑ , and $\lambda_v = v'v$ is the line emerging from v and entering v' ;
- (2) r is the root of the graph, and $i\nu_r$ denotes the unit vector in the j th direction, $i\nu_r = \mathbf{e}_j$, $j = 1, \dots, \ell$;
- (3) ν_v is the *external momentum* associated with the node v , $\nu_{\lambda_v} = \sum_{w \leq v} \nu_w$ is the *momentum* flowing through the line λ_v , and $\nu(\vartheta)$ is the momentum flowing through the line entering the root. *Only external momenta such that $\nu_\lambda \neq 0$ for all λ will be allowed*: the generality of this property was the contribution of Poincaré to the analysis of the Lindstedt–Newcomb series, [15], Ch. IX (T. II).

It can be convenient to introduce the notations

$$g_\lambda \equiv \frac{1}{\omega_0 \cdot \nu_\lambda} , \quad D(\vartheta) = \prod_{\lambda \in \vartheta} g_\lambda , \quad (2.3)$$

so that (2.2) becomes

$$\text{Val}(\vartheta) = D(\vartheta) \prod_{v < r} (\nu_{v'} \cdot f_{\nu_v}) \quad (2.4)$$

g_λ is called the *propagator* of the line λ . Let us denote by $\mathcal{T}(k, \nu)$ the set of non equivalent labeled graphs of order k with $\nu(\vartheta) = \nu$ and $i\nu_r = e_j$; then

$$h_{j\nu}^{(k)} = \frac{1}{k!} \sum_{\vartheta \in \mathcal{T}(k, \nu)} \text{Val}(\vartheta) = \frac{1}{k!} \sum_{\vartheta^0 \in \mathcal{T}^0(k)} W(\vartheta^0, \nu) , \quad (2.5)$$

where $\vartheta = (\vartheta^0, \{\nu_\lambda\})$, if ϑ is a labeled graph, while ϑ^0 is a graph bearing no external momentum labels, $\mathcal{T}^0(k)$ is the set of such graphs of order k , and

$$W(\vartheta^0, \nu) \equiv \sum_{\{\nu_x\}: \nu(\vartheta) = \nu} \text{Val}(\vartheta) . \quad (2.6)$$

§3 The basic ultraviolet cancellation.

We consider the simplest graph that cannot be naively estimated uniformly in the ultraviolet cut off κ : it is the "brusch" graph:

Recalling that $f_\nu = a|\nu|^{-b}$, its value is given by:

$$\frac{e \cdot a}{(\omega_0 \cdot \nu) |\nu_0|^b} \prod_{j=1}^{k-1} \frac{\nu_0 \cdot a}{(\omega_0 \cdot \nu_j) |\nu_j|^b} \quad \nu = \sum_{j=0}^{k-1} \nu_j \quad (3.2)$$

Then we see that the sum of all the contributions of the above type, with fixed k and various ν_j such that $\sum_j \nu_i = \nu$, to $h_\nu^{(k)}$ can be bounded by:

$$C^k |\nu_0|^{k-1} |a|^k |\nu|^\tau |\nu_0|^{-b} \prod_{j=1}^{k-1} |\nu_j|^{\tau-b} \quad (3.3)$$

which, to be summable over ν_j requires that $\mathbf{f}$ is in $\hat{C}^{(>k-1+2\tau)}(T^\ell)$, so that the contributions from the graphs just considered may not make sense for large k no matter how regular f is (i.e. for $k \geq b + 1 - 2\tau - \ell$).

But the problem really arises only if $|\nu_0| \gg |\nu_j|$: in the cases $|\nu_0| \ll |\nu_j|$ we can replace $|\nu_0|$ by $|\nu_j|$ and $|\nu|$ by $k \prod_v |\nu_v|$ in the bounds and, by (3.3), this would require only $f \in \hat{C}^{(>2\tau+1)}$.

Therefore we add, for each ν_j with $|\nu_j| \ll |\nu_0|$, the graph with ν_0, ν_j replaced by $\nu_0 + 2\nu_j, -\nu_j$, and noting that both graphs contribute to the same Fourier component ν of $\mathbf{h}$, we see that this replaces:

$$\frac{(\nu_0 \cdot \mathbf{a})}{(\omega_0 \cdot \nu) |\nu_0|^b (\omega_0 \cdot \nu_j)} \rightarrow \frac{1}{(\omega_0 \cdot \nu)} \left(\frac{(\nu_0 \cdot \mathbf{a})}{|\nu_0|^b (\omega_0 \cdot \nu_j)} - \frac{(\nu_0 + 2\nu_j) \cdot \mathbf{a}}{|\nu_0 + 2\nu_j|^b (\omega_0 \cdot \nu_j)} \right) \quad (3.4)$$

the other factors in (3.2) remain unchanged by the parity properties of $\mathbf{f}$.

More generally if $|\nu_0| \gg |\nu_j|$, $j = 1, \dots, s$ and $|\nu_0| \ll |\nu_j|$, $j > s$ we combine all the graphs with ν_0 replaced by $\nu_0 + \sum_{j=1}^s (1 - \sigma_j) \nu_j$ and with ν_j replaced by $\sigma_j \nu_j$ for $j = 1, \dots, s$ and $\sigma_j = \pm 1$. The sum of the considered graph values can be expressed by an interpolation formula; let us define $\nu(t_{v_0}) = \nu_0 + 2 \sum_{j=1}^s t_j \nu_j$ for $t_{v_0} = (t_1, \dots, t_s)$, then the sum is:

$$\int_1^0 \left(\prod_{j=1}^s dt_j \frac{\partial}{\partial t_j} \right) \left\{ \frac{\mathbf{e} \cdot \mathbf{a}}{\omega_0 \cdot \nu} \frac{1}{|\nu_{v_0}(t_{v_0})|^b} \prod_{j=1}^{k-1} \frac{\nu_{v_0}(t_{v_0}) \cdot \mathbf{a}}{\omega_0 \cdot \nu_j} \frac{1}{|\nu_j|^b} \right\} \quad (3.5)$$

and the *cancellation* is expressed by the fact that in the above interpolation formula there are no contributions from the boundary values $t_j = 1$. It is important to note, also, that the factors containing ω_0 are $\mathbf{t}$ -independent, hence they (i.e. the "small divisors") are *not affected* by the interpolation.

The (3.5) can be bounded as before *but replacing each of the s vectors ν_{v_0} which are large with the corresponding $|\nu_j|$* , and (trivially) by replacing the other $k - 1 - s$ and ν_0 's by the corresponding $|\nu_j|$; i.e. for a suitable $\tilde{C}$ the bound (3.3) is replaced by:

$$\tilde{C}^k |\nu|^\tau |\nu_0|^{-b} \prod_{j=1}^{k-1} |\nu_j|^{\tau+1-b} \quad (3.6)$$

and one needs (to sum over ν_j at fixed ν) only $b > \ell + \tau + 1$, i.e. $f \in \hat{C}^{(>\tau+1)}(T^\ell)$, or $b > \ell + 2\tau + 1$ to sum without fixing ν , (i.e. $f \in \hat{C}^{(2\tau+1)}$).

This is the key "ultraviolet" cancellation introduced and used in [7] and [8]. Clearly there remains quite a lot of work to do:

(1) how does one treat the cases in which the ν_0 and the ν_j are "close"? In such cases $\nu_{v_0}(t_{v_0}) = \nu_0 + 2 \sum_{j=1}^{k-1} t_j \nu_j$ may vanish for some choices of the interpolation variables t_j . This type of situation is a manifestation of rather general class of difficulties known in field theory as *overlapping divergences*: it must clearly be avoided.

(2) how does one treat the cases in which the graph is more elaborated, with many nodes with bifurcations, and therefore many of the above problems arise *simultaneously*?

(3) how does one treat the *infrared* cancellations of [10]? are they obvious even considering the necessity of collecting graphs so that no ultraviolet divergence arises? In principle exhibiting the ultraviolet cancellations may require different *non commuting* collections of terms and this would ruin the analysis (again a "overlapping divergences problem").

In the next section we discuss the basic ideas about (1): this will give an idea also for the solution of the problem (2). The problem in (3) is real and it arises, in fact. However the incompatibility is somewhat trivial and one can eliminate it by giving up performing some ultraviolet cancellations. This leads eventually, see [8], only to replace the minimum value for p that one would naively expect from Eliasson method and his extension of Siegel's small divisor bound ($p > 3\tau + 1$) with $p > 3\tau + 2$.

§4 Overlapping divergences?

Consider the case of fig.(3.1): we set $p_{v_1} = k - 1$ and $w_j = v_j$ and rename v_1 the graph node previously called v_0 and $w_1, \dots, w_{p_{v_1}}$ the $k - 1$ endnodes.

The change in notation is made to adhere to the notation in [8] which is apt to treat also the more structured graphs.

Then we introduce the notion of *scale* of a vector $\nu \in Z^l$ (ν will also be called a "mode"). This is done by simply declaring it of *scale* h if $2^{h-1} \leq |\nu| < 2^h$.

We shall set $\nu_{w_j} \equiv \nu_{\lambda_{w_j}}$, again for ease of reference to [8]. Note that in the graph (3.1) ν_{w_j} and ν_{w_j} accidentally coincide due to the graph simplicity.

Fixed ν and $\{\nu_{\lambda_w}\}_{w \in B_{v_1}}$ let h_{v_1} be the *scale* of ν_{v_1} (it depends on ν and on the $\nu_{\lambda_{w_j}} \equiv \nu_{w_j}$ because $\nu_{v_1} = \nu - \sum \nu_{w_j}$). This means that ν_{v_1} is such that $2^{h_{v_1}-1} \leq |\nu_{v_1}| < 2^{h_{v_1}}$ and we refer to this relation by saying that h_{v_1} is *compatible* with ν_{v_1} .

With the new notations the value of the graph is:

$$\frac{e \cdot a}{|\nu_{v_1}|^b \omega_0 \cdot \nu} \prod_{j=1}^{p_{v_1}} \frac{\nu_j \cdot a}{|\nu_j|^b \omega_0 \cdot \nu_j} \quad \nu = \sum_{j=0}^{k-1} \nu_j \quad (4.1)$$

We say that w_j is *out of order* with respect to v_1 if:

$$2^{h_{v_1}} > 2^o p_{v_1} |\nu_{\lambda_w}| \quad o \approx 5 \quad (4.2)$$

where p_v is the number of branches entering v .

We denote B_{1v_1} the set of nodes $w \in B_{v_1}$ which are out of order with respect to v_1 : this set can be denoted also $B_{1v_1}(\nu, \{\nu_{\lambda_w}\}_{w \in B_{v_1}})$. The number of elements in B_{1v_1} will be denoted $q_v = |B_{1v_1}|$. The notion of w being out of order with respect to v_1 depends on $\{\nu_{\lambda_w}\}_{w \in B_{v_1}}$ and ν .

Given a set $\{\nu_{\lambda_w}\}_{w \in B_{v_1}}$ for all choices of $\sigma_w = \pm 1$ we define the map:

$$U(\{\nu_{\lambda_w}\}_{w \in B_{v_1}}) \equiv (\{\sigma_w \nu_{\lambda_w}\}_{w \in B_{v_1}}) \quad (4.3)$$

and given a set $C \subseteq \mathcal{B}_{v_1}$ we call $\mathcal{U}(C)$ the set of all transformations U such that $\sigma_w = 1$ for $w \notin C$; we also set $\sigma_w = 1$ if $w \notin \mathcal{B}_{1v_1}$.

If $[2^{h-1}, 2^h)$ is a scale interval I_h , $h = 1, 2, \dots$ we call

- the first quarter of I_h the *lower part* $I_h^- = [2^{h-1}, \frac{5}{4}2^{h-1})$ of I_h ,
- the fourth quarter of I_h the *upper part* $I_h^+ = [\frac{7}{8}2^h, 2^h)$ of I_h , and
- the remaining part the *central part* I_h^c .

We shall group branch momenta $\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}$ into collections by proceeding iteratively. *The collections will be built so that in each collection the cancellation discussed in Remark 4.2 above can be exhibited.*

Fixed ν and h choose $\{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}$ such that $|\nu_{v_1}^1| \in I_h^c$: such $\{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}$ is called a *representative*. Given the representative we define:

- the *branch momenta collection* to be the set of the $\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}$ of the form

$$U(\{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}), \quad U \in \mathcal{U}(\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}})); \quad (4.4)$$

- the *external momenta collection*, for $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}))$, to be the set of momenta

$$\nu_w^{1U} = \sigma_w \nu_w, \quad \nu_{v_1}^{1U} = \nu - \sum_{w \in \mathcal{B}_{v_1}} \sigma_w \nu_{\lambda_w}^1, \quad (4.5)$$

Note that the elements of the above constructed external momenta collection need not be necessarily contained in I_h^c (see below).

We consider then another representative $\{\nu_{\lambda_w}^2\}_{w \in \mathcal{B}_{v_1}}$ such that $|\nu_{v_1}^2| \in I_h^c$ and not belonging to the branch momenta collection associated with $\{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}$, if there are any left; and we consider the corresponding branch momenta and external momenta collections as above. We proceed in this way until all the representatives such that ν_1 is in I_h^c , for the given h , have been put into some collection of branch momenta.

We then repeat the above construction with the interval I_h^- replacing the I_h^c , always being careful not to consider representatives $\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}$ that appeared as members of previously constructed collections. It is worth pointing out that not all the external momenta $\nu_{v_1}^U$, $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}))$, are in I_h^- , but they are all in the corridor $I_{h-1}^+ \cup I_h^-$, by (4.2).

Finally we consider the interval I_{h-1}^+ , (if $h = 1$ we simply skip this step). The construction is repeated for such intervals.

Proceeding iteratively in this way starting from $h = 1$ and, after exhausting all the $h = 1$ cases, continuing with the $h = 2, 3, \dots$ cases, we shall have grouped the sets of branch momenta into collections obtainable from a representative $\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}$ by applying the operations $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}))$ to it. *Note that, in this way, when the interval I_{h-1}^+ is considered, all the remaining representatives are such that $|\nu_{v_1}^U| \in I_{h-1}^+$ for all $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}^1\}_{w \in \mathcal{B}_{v_1}}))$.*

Note that the graphs with momenta in each collection are just the graphs involved in the parity cancellation described in the previous section. In fact if U is generated by the signs $\{\sigma_w\}_{w \in \mathcal{B}_v}$, we have:

$$\begin{aligned} \nu_{v_1}^U &= \left(\prod_{w \in \mathcal{B}_{1v_1}} U_{v_1 w}^{\sigma_w} \{\nu_x\} \right)_{v_1} \text{ " = " } \nu + \sum_w (1 - \sigma_w) \nu_w \\ (U(\{\nu_{\lambda_{\tilde{w}}}\}_{\tilde{w} \in \mathcal{B}_{v_1}}))_w &= \sum_{z \leq w} \left(\prod_{\tilde{w} \in \mathcal{B}_{1v_1}} U_{v_1 \tilde{w}}^{\sigma_{\tilde{w}}} \{\nu_x\} \right)_z \text{ " = " } \sigma_w \nu_w \end{aligned} \quad (4.6)$$

where, given the sets of momenta $\{\nu_x\}$ and $\{\nu_{\lambda_{\tilde{w}}}\}$, $(\{\nu_x\})_v$ denotes the external momentum in $\{\nu_x\}$ corresponding to the node v and $(\{\nu_{\lambda_{\tilde{w}}}\})_w$ denotes the branch momentum in $\{\nu_{\lambda_{\tilde{w}}}\}$ corresponding to the branch λ_w . The " = " mean that, in the special case of (3.1) that we treat, the equality between ν_w and ν_{λ_w} implies that the r.h.s. of (4.6) takes the simple form indicated. The general case of more structured graphs is more involved, [8].

Remark. The complexity of the above construction is due to the necessity of avoiding overcountings: *i.e.* overlapping of cancellations whereby the same graph value is used to "compensate" two or more other graph values. In fact it is possible that, for some $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}))$, one has:

$$\mathcal{B}_{1v_1}(\nu, U(\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}})) \neq \mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}), \quad (4.7)$$

because the scale of $\nu_{v_1}^U$ may be $h-1$, while that of ν_{v_1} may be h ; so that if one considered, for instance, I_{h-1}^+ before I_h^- overcountings would be possible, and in fact they would occur.

It is convenient rewriting the sum that we want to estimate, *i.e.* the contribution to $\mathbf{h}^{(k)}$ from the sum of values of graphs like (3.1), as:

$$\sum_{\nu} |\nu|^s \left| \sum_{h_{v_1}} \sum_{\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}}^* \sum_{U \in \mathcal{U}(\mathcal{B}_{1v_1})} \frac{\mathbf{e} \cdot \mathbf{a}}{\omega_0 \cdot \nu |\nu_{v_1}^U|^b} \left(\prod_{w \in \mathcal{B}_{1v_1}} \sigma_w \right) \prod_{j=1}^{p_{v_1}} \frac{\nu_{v_1} \cdot \mathbf{a}}{\omega_0 \cdot \nu_{w_j} |\nu_{w_j}|^b} \right| \quad (4.8)$$

where $\nu_{v_1}(t_{v_1}) = \nu_{v_1} + 2 \sum_{j=1}^{k-1} t_j \nu_j$, $\nu_w \equiv \nu_{\lambda_w}$, and the sum $\sum_{\{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}}}^*$ means sum over the above defined representatives such that ν_{v_1} is "compatible" with h_{v_1} , *i.e.* it has scale h_{v_1} (see lines preceding (4.1)); and we abridge $\mathcal{B}_{1v_1}(\nu, \{\nu_{\lambda_w}\}_{w \in \mathcal{B}_{v_1}})$ by $\mathcal{B}_{1v_1}$ in conformity with the notations introduced after (4.2). The summation over ν is weighted with the weight $|\nu|^s$ in order to obtain also estimates of the s -th derivatives of $\mathbf{h}^{(k)}(\psi)$.

The explicit sum over the scales h_{v_1} is introduced to simplify the bounds analysis that we perform now. Note that $\nu_{v_1}^U$ is, in general, not compatible with h_{v_1} , *i.e.* we are grouping together also terms with different scale label (but the difference in scale is at most one, see (4.11) below).

We can now apply interpolation in (3.5) to the node v and rewrite (4.8) as:

$$\sum_{\nu} |\nu|^s \left| \sum_{h_{v_1}} \sum_{\{\nu_{w_j}\}_{w \in B_{v_1}}}^* \left[\int_1^0 \left(\prod_{w \in B_{v_1}} dt_w \right) \cdot \right. \right. \\ \left. \left. \left(\prod_{w \in B_{v_1}} \frac{\partial}{\partial t_w} \right) \left(\frac{1}{(\omega_0 \cdot \nu) |\nu_{v_1}(t_{v_1})|^b} \prod_{j=1}^{k-1} \frac{\nu_{v_1}(t_{v_1}) \cdot a}{(\omega_0 \cdot \nu_{w_j}) |\nu_{w_j}|^b} \right) \right] \right| \quad (4.10)$$

where if $B_{v_1} = \emptyset$ no interpolation is made; by the definition of nodes out of order, see (4.2), and by the iterative grouping of the representatives it is:

$$2^{h_{v_1}-2} \leq |\nu_{v_1}(t_{v_1})| < 2^{h_{v_1}}, \quad (4.11)$$

so that the interpolation formulae discussed in §3 *can be used* because no singularity (i.e. no division by 0) arises in performing the t_{v_1} -integrations.

It is now easy to check, by computing the derivatives, that each of the terms in (4.10) can be conveniently bounded by its (easily estimated) maximum value (over τ) and the resulting bound has the form obtained in the simple case of §3, (3.6), in which some of the ν_{w_j} were $\ll |\nu_{v_1}|$ and some were $\gg |\nu_{v_1}|$. The ν_w of the nodes out of order play the role of the small ν_{w_j} 's and the others can be used as bounds on ν_{v_1} . *No further care needs to be exercised in the bounds.* The details are not very interesting, see [8].

The extension to more structured graphs is not really more difficult. One has to be careful in interpreting ν_{w_j} as the momentum flowing on the brach $v_1 w_j$, i.e. as the sum of the node momenta of all the nodes above v_1 . See [7] for the detailed algebra that is necessary. Therefore the above analysis gives a rather complete idea of the ultraviolet cancellations mechanism. The infrared cancellations, i.e. the cancellations discussed by Eliasson does not interfere with the above ultraviolet cancellation and the analysis is done in [8]. Note that in the case of the first of (1.3) with $D = (-\Delta)^r$, or the second with $D = (-\Delta)^r L$, there is no need of infrared cancellations (i.e. no small divisors are present).

The above analysis seems to imply that the validity condition for the theorem should be $p > 2\tau + 1$ instead of $3\tau + 2$, §1. The "extra" τ is due to the fact that the product of the small divisors, in general graphs "without resonances" (see [8]) is essentially estimated, following [10] (see also [7]) by $\prod_v |\nu_v|^{3\tau}$, instead of the $\prod_v |\nu_v|^{2\tau}$ that occurs in the simple example chosen here. The "extra" 1 is due to the fact that general graphs may contain "resonances" (see [10], and §3 in [8]) which interfere with the ultraviolet cancellations causing an extra loss of one unit in the differentiability condition.

Acknowledgements. This review is part of the research program of the European Network on: "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460: it includes the text of a conference delivered at the meeting "Recent Advances in partial differenzial equations and applications", Venezia, june 10-14, 1996 and describes the results in [7], [8].

References

- [1] G. Gallavotti: *Twistless KAM tori*, Communications in Mathematical Physics, **164**, 145–156 (1994).
- [2] J. Moser: *Convergent series expansions for quasi periodic motions*, *Mathematische Annalen*, **169**, 136–176 (1967).
- [3] G. Gallavotti: Classical Mechanics and Renormalization Group, Lectures notes, 1983 Erice School, Ed. A. Wightman, G. Velo, Plenum, N.Y. (1985).
- [4] J. Moser: *A rapidly convergent iteration method and nonlinear differential equations II*, *Annali della Scuola Normale Sup. di Pisa*, **20**, 499–535, (1966).
- [5] I.S. Parasyuk: *On instability zones of the Schrödinger equation with a quasi periodic potential*, *Ukr. Mat. Zh.* **30**, 70–78 (1985).
- [6] L.A. Pastur, A. Figotin: *Spectra of random and almost-periodic operators*, *Grundlehren der Math. Wissenschaften* **297**, Springer, Berlin (1992).
- [7] F. Bonetto, G. Gallavotti, G. Gentile, V. Mastropietro: *Lindstedt series, ultraviolet divergences and Moser's theorem*, Preprint IHES/P/95/102 (1995).
- [8] F. Bonetto, G. Gallavotti, G. Gentile, V. Mastropietro: *Quasi linear flows on tori: regularity of their linearization*. in mp_arc@math.utexas.edu, #96-251 and chao-dyn@xyz.lanl.gov #9605017.
- [9] G. Gallavotti, G. Gentile, V. Mastropietro: *Field theory and KAM tori*, Mathematical Physics Electronic Journal, **1**, (5): 1–13 (1995), (<http://mpej@math.utexas.edu>).
- [10] H. Eliasson: *Absolutely convergent series expansions for quasi-periodic motions*, *Mathematical Physics Electronic Journal*, **2**, 33 pp, 1996, <http://mpej@math.utexas.edu>.
- [11] G. Gallavotti, G. Gentile: *Majorant series method for KAM tori*, *Ergodic theory and applications*, **15**, p. 857–869, 1995.
- [12] G. Gentile, V. Mastropietro: *KAM theorem revisited*, *Physica D*, **90**, 225–234 (1996).
- [13] G. Gentile, V. Mastropietro: *Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications*, *Reviews of Mathematical Physics*, **8**, 393–444 (1996).
- [14] F. Harary, E. Palmer: *Graphical enumeration*, Academic Press, N.Y., 1973.
- [15] Poincaré, H.: *Les Méthodes Nouvelles de la Mécanique Céleste*, Gauthiers Villars, Paris, 1893.

Internet: Authors' preprints also freely downloadable (latest revision-correction, postscript version) from

<http://chimera.roma1.infn.it>

<http://www.math.rutgers.edu/~giovanni>

Mathematical Physics Preprints (mirror) pages.

e-mail: giovanni@ipparco.roma1.infn.it