# Methods in the theory of quasi periodic motions

Giovanni Gallavotti*

* Fisica, Università di Roma, P.le Moro 2, 00185 Roma

*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 discussed, with attention to the cancellation mechanisms that make regularity possible.*

## §*1 The Hamilton Jacobi equation.*

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

1.1

$$\begin{aligned}(\underline{\omega}_0\cdot\underline{\partial})^2\,\underline{h}(\underline{\psi}) &= \varepsilon\,(\underline{\partial} f)(\underline{\psi}+\underline{h}(\underline{\psi}))\\ (\underline{\omega}_0\cdot\underline{\partial})\,\underline{h}(\underline{\psi}) &= \varepsilon\,\underline{f}(\underline{\psi}+\underline{h}(\underline{\psi}))+\underline{N}\end{aligned}\qquad(1.1)$$

where $\underline{\psi}$ is a point on the $\ell$–dimensional torus, $\underline{h}(\underline{\psi})$ a $R^\ell$–valued function, $f$, $\underline{f}$ are functions on $T^\ell$, $\underline{N}$ is a vector in $R^\ell$ and $\underline{\omega}_0$ is a diophantine vector in $R^\ell$ such that, considering the integer components lattice $Z^\ell$, there are constants $C,\tau>0$ with:

1.2

$$C|\underline{\omega}_0\cdot\underline{\nu}|\ge|\underline{\nu}|^{-\tau}\qquad\forall\underline{\nu}\in Z^\ell,\underline{\nu}\ne\underline{0}\qquad(1.2)$$

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

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), and the second corresponds to the KAM tori for perturbations of non resonant harmonic oscillators: in the latter case not only $\underline{h}$ is unknown but also $\underline{N}$, 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, [M1].

Here we shall consider a wider class of equations:

1.3

$$\begin{aligned}&(1)\qquad D\,\underline{h}(\underline{\psi}) = \varepsilon\,(\underline{\partial} f)(\underline{\psi}+\underline{h}(\underline{\psi}))\\ &(2)\qquad D\,\underline{h}(\underline{\psi}) = \varepsilon\,\underline{f}(\underline{\psi}+\underline{h}(\underline{\psi}))+\underline{N}\end{aligned}\qquad(1.3)$$

where $D$ is a pseudodifferential operator like, respectively:

1.4

$$\begin{array}{llll}(1) & D = (\underline{\omega}_0 \cdot \underline{\partial})^2, & \text{or} & D = (-\Delta)^r \quad r \ge 0 \\ (2) & D = (\underline{\omega}_0 \cdot \underline{\partial}), & \text{or} & D = (-\Delta)^r L \quad r \ge 0 \end{array} \tag{1.4}$$

where $\Delta$ is the Laplace operator and $L$ is a linear elliptic operator ("Dirac operator": for instance $\underline{\gamma} \cdot \underline{\partial}$ where $\underline{\gamma}$ 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 $r \ge 0$ do have a solution $\underline{h}$ or, respectively, $\underline{h}$, $\underline{N}$ for small $\varepsilon$ if $\underline{f}$ is regular enough. This remains true, but it is far less easy, in the cases in which $D$ is a power of $(\underline{\omega}_0 \cdot \underline{\partial})$:

• The classical results for $D = (\underline{\omega}_0 \cdot \underline{\partial})^2$ that hold when $f$ is assumed analytic on $T^\ell$ give analyticity in $\varepsilon$, $\underline{\psi}$ of the solution $\underline{h}$ and $\underline{N}$, for $\varepsilon$ small, [M1].
• 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 $\varepsilon$ small. The solution is also $p-1$ times differentiable in $\varepsilon$.
• In the case $D = (\underline{\omega}_0 \cdot \underline{\partial})^2$ and $f \in C^{(p)}(T^\ell)$ it admits a solution in class $C^{(p-4\tau+2)}$ if $p > 4\tau + 2$, as a consequence of Moser' s work [M2], where the contraction method is replaced by Nash's implicit functions theorem. The solution is also smooth in $\varepsilon$.
• In the case $D = \underline{\omega}_0 \cdot \underline{\partial}$ 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, [Pa], see [PF], based on Moser's technique; the solution is smooth in $\varepsilon$.

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

We consider functions $f$, $\underline{f}$ which are finitely differentiable but that have the special form: $\underline{f} = (f_1, \ldots, f_\ell)$, with each $f_j$ of the class $\hat{C}^{(p)}(T^\ell)$ introduced in [BGGM1], for some $p$.

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

1.5

$$f_{\underline{\nu}} = \sum_{n \ge p+\ell}^{N} \frac{c_n + d_n (-1)^{||\underline{\nu}||}}{|\underline{\nu}|^n} \tag{1.5}$$

for some $N \ge p + \ell$ and some constants $c_n^{(j)}$, $d_n^{(j)}$.

For instance we can choose $f_{j\underline{\nu}} = a_j |\underline{\nu}|^{-b}$, with $b = p + \ell$ and $a_j \in R$; then define $\sup_{j=1,\ldots,\ell} |a_j| = a$ and $\underline{a} = (a_1, \ldots, a_\ell)$. 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 $\underline{f}$ are special even functions.* The results that follow do not hold in general for non even functions, see [BGGM1] for some simple counterexamples.

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

1.6

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

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

In [BGGM1],[BGGM2] the following theorem is proved.

**Theorem:** *Equations (1.3) admit solutions that are analytic in $\varepsilon$ for $\varepsilon$ small if $f$ or, respectively, $\underline{f}$ are in $\hat{C}^{(p)}(T^\ell)$ and $p$ is large enough. For instance:*
*(i) if $D = (\underline{\omega}_0 \cdot \underline{\partial})^2$ 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 $D = \underline{\omega}_0 \cdot \underline{\partial}$ the second of (1.3) admits a $C^{(<p-3\tau-2)}(T^\ell)$ solution if $f \in \hat{C}^{(p)}(T^\ell)$ provided $p - 3\tau - 2 > 0$.*
*(ii) if $D = 1$ the first of (1.3) admits a $C^{(<p-3)}(T^\ell)$ solution if $f \in \hat{C}^{(p)}(T^\ell)$ provided $p - 3 > 0$.*

The analyticity in $\varepsilon$ is perhaps unexpected even in the case $D = 1$ since the solutions $\underline{h}(\underline{\psi})$ are *not* analytic. 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 [GGM]. The cancellations *are not* the ones pointed out by Eliasson, [E], 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 but treat the simplest forms of the problem thus providing a simple understanding of the new methods (see for instance [G1] which has the advantage of being short but uncompromising with the key difficulties; further developments can be found, *along the same lines*, in [GG], [GM1], [GM2], [GGM] 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 [GGM].

In §2 we discuss the Linstedt formal series solution of (1.3): we shall deal with the second of (1.3) with $D = \underline{\omega}_0 \cdot \underline{\partial}$ as it is simpler than the first of (1.3) with $D = (\underline{\omega}_0 \cdot \underline{\partial})^2$. 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 for the analysis only because we 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 [Pa], see [PF], 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 for $\ell = 2$ because $\hat{C}^{(p)} \subset C^{(p)}$).

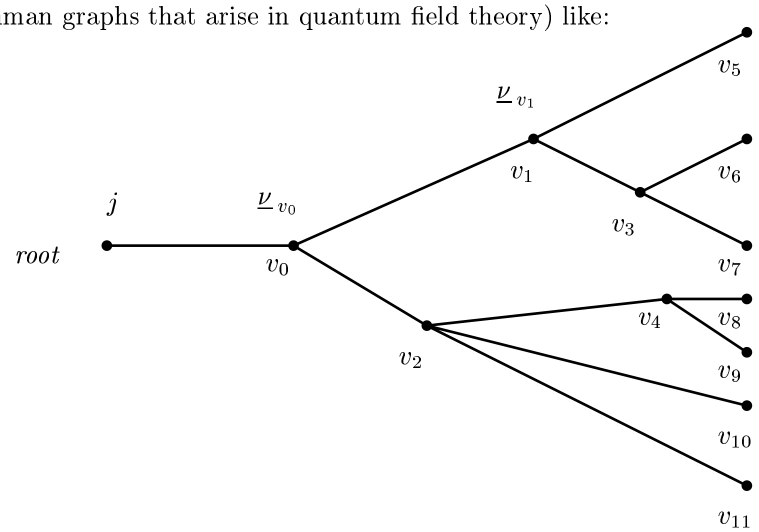
*§2 Lindstedt series.*

Consider the second of Eq. (1.3) with $f_{j\underline{\nu}} = a_j |\underline{\nu}|^{-b}$ replaced by $f_{j\underline{\nu}}\, e^{-\kappa|\underline{\nu}|}$ and we write $\underline{h}(\underline{\psi}) = \sum_{k=1}^\infty \varepsilon^k\, \underline{h}^{(k)}(\underline{\psi})$. The parameter $\kappa$ is taken $\kappa > 0$, and, after computing the coefficients $\underline{h}^{(k)}_{\underline{\nu}}$ of the Forurier series of the function $\underline{h}^{(k)}(\underline{\psi})$ which will depend on $\kappa$, we shall perform the limit $\kappa \to 0$ (*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 $\underline{h}^{(k)}$ a family of "Feynman graphs" $\vartheta$ and to each graph a *value* which is a function of $\underline{\nu} \in Z^\ell$ so that summing all values at fixed $\underline{\nu}$ of the graphs that are associated with $\underline{h}^{(k)}$ one gets the Fourier component of order $\underline{\nu}$ of $\underline{h}^{(k)}$.

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

2.1 

(2.1)

**Fig.1**. A graph $\vartheta$ 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 $k-1$).

We denote by $\le$ 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 $\underline{\nu}_v \ne \mathbf{0}$; with the root of the graph (which is not regarded as a node) we associate a label $j = 1, \ldots, \ell$.

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 $\lambda_v$). 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 $\lambda_v$, $\lambda_w$, if $v$, $w$ are such.

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 $\vartheta_{vw}$.

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 $4^k k!$, [HP].

The "value" of a graph $\vartheta$ is now given by

2.2 $$\mathrm{Val}\,(\vartheta)=\prod_{v<r}\frac{(i\,\underline{\nu}_{v'}\cdot\underline{f}_{\,\underline{\nu}_v})}{i\,\underline{\omega}_0\cdot\underline{\nu}_{\lambda_v}}\;, \tag{2.2}$$

where
(1) $v'$ is the node immediately following $v$ in $\vartheta$, and $\lambda_v=v'v$ is the line emerging from $v$ and entering $v'$;
(2) $r$ is the *root* of the graph, and $i\,\underline{\nu}_r$ denotes the unit vector in the $j$th direction, $i\,\underline{\nu}_r=\underline{e}_j$, $j=1,\ldots,\ell$;
(3) $\underline{\nu}_v$ is the *external momentum* associated with the *node* $v$, $\underline{\nu}_{\lambda_v}=\sum_{w\le v}\underline{\nu}_w$ is the *momentum* flowing through the line $\lambda_v$, and $\underline{\nu}(\vartheta)$ is the momentum flowing through the line entering the root. *Only external momenta such that* $\underline{\nu}_\lambda\ne\underline{0}$ *for all* $\lambda$ *will be allowed*: the generality of this property was the contribution of Poincaré to the analysis of the Lindstedt–Newcomb series, [P], Ch. IX (T. II).

It can be convenient to introduce the notations

2.3 $$g_\lambda\equiv\frac{1}{\underline{\omega}_0\cdot\underline{\nu}_\lambda}\;,\qquad D(\vartheta)=\prod_{\lambda\in\vartheta}g_\lambda\;, \tag{2.3}$$

so that (2.2) becomes

2.4 $$\mathrm{Val}\,(\vartheta)=D(\vartheta)\prod_{v<r}\Big(\underline{\nu}_{v'}\cdot\underline{f}_{\,\underline{\nu}_v}\Big) \tag{2.4}$$

$g_\lambda$ is called the *propagator* of the line $\lambda$. Let us denote by $\mathcal{T}(k,\underline{\nu})$ the set of non equivalent labeled graphs of order $k$ with $\underline{\nu}(\vartheta)=\underline{\nu}$ and $i\,\underline{\nu}_r=\underline{e}_j$; then

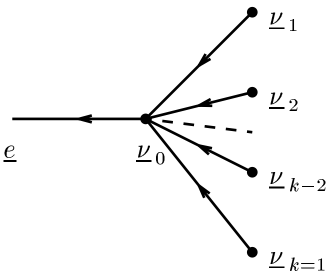
2.5 $$h^{(k)}_{j\,\underline{\nu}}=\frac{1}{k!}\sum_{\vartheta\in\mathcal{T}(k,\underline{\nu})}\mathrm{Val}\,(\vartheta)=\frac{1}{k!}\sum_{\vartheta^0\in\mathcal{T}^0(k)}W(\vartheta^0,\underline{\nu})\;, \tag{2.5}$$

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

2.6 $$W(\vartheta^0,\underline{\nu})\equiv\sum_{\{\underline{\nu}_x\}:\,\underline{\nu}(\vartheta)=\underline{\nu}}\mathrm{Val}\,(\vartheta)\;. \tag{2.6}$$

§*3 The basic ultraviolet cancellation.*

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

3.1 

(3.1)

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

3.2
$$\frac{\underline{e}\cdot\underline{a}}{(\underline{\omega}_0\cdot\underline{\nu}_0)|\underline{\nu}|^b}\prod_{j=1}^{k-1}\frac{\underline{\nu}_0\cdot\underline{a}}{(\underline{\omega}_0\cdot\underline{\nu}_j)|\underline{\nu}|^b}\qquad\qquad \underline{\nu}=\sum_{j=0}^{k-1}\underline{\nu}_j \tag{3.2}$$

Then we see that the bound on the sum of all the contributions of the above type, with fixed $k$ and various $\underline{\nu}_j$ such that $\sum_j \underline{\nu}_i = \underline{\nu}$, to $\underline{h}^{(k)}_{\underline{\nu}}$ can be bounded by:

3.3
$$C^k|\underline{\nu}_0|^{k-1}|\underline{a}|^k|\underline{\nu}_0|^{\tau-b}\prod_{j=1}^{k-1}|\underline{\nu}_j|^{\tau-b} \tag{3.3}$$

which, to be summable over $\underline{\nu}_j$ requires that $\underline{f}$ is in $\hat{C}^{(>k-1+\tau)}(T^\ell)$ and also $f\in\hat{C}^{(>\tau)}(T^\ell)$, so thhat the contributions from the graphs just considered may not make sense for large $k$ *no matter how regular $f$ is*.

But the problem really arises only if $|\underline{\nu}_0|\gg|\underline{\nu}_j|$: in the cases $|\underline{\nu}_0|\ll|\underline{\nu}_j|$ we can replace $|\underline{\nu}_0|$ by $|\underline{\nu}_j|$ in the bounds and, by (3.3), this would require only $f\in\hat{C}^{(>\tau+2)}$.

Therefore we collect for each $\underline{\nu}_j$, with $|\underline{\nu}_j|\ll|\underline{\nu}_0|$, the graph with $\underline{\nu}_0,\underline{\nu}_j$ replaced by $\underline{\nu}_0+2\underline{\nu}_j,-\underline{\nu}_j$, and noting that both graphs contribute to the same Fourier component $\underline{\nu}$ of $\underline{h}$, we see that this replaces:

3.4
$$\begin{aligned}&\frac{(\underline{\nu}_0\cdot\underline{a})}{(\underline{\omega}_0\cdot\underline{\nu})|\underline{\nu}_0|^b(\underline{\omega}_0\cdot\nu_j)}\to\\ &\frac{1}{(\underline{\omega}_0\cdot\underline{\nu})}\Big(\frac{(\underline{\nu}_0\cdot\underline{a})}{|\underline{\nu}_0|^b(\underline{\omega}_0\cdot\nu_j)}-\frac{(\underline{\nu}_0+2\underline{\nu}_j)\cdot\underline{a}}{|\underline{\nu}_0+2\underline{\nu}_j|^b(\underline{\omega}_0\cdot\underline{\nu}_j)}\Big)\end{aligned} \tag{3.4}$$

the rest remains unchanged by the parity properties of $\underline{f}$.

More generally if $|\underline{\nu}_0|\gg|\underline{\nu}_j|$, $j=1,\ldots,s$ and $|\underline{\nu}_0|\ll|\underline{\nu}_j|$, $j>s$ we combine all the graphs with $\underline{\nu}_0$ replaced by $\underline{\nu}_0+2\sum_{j=1}^s\sigma_j\underline{\nu}_j$ and with $\underline{\nu}_j$ replaced by $\sigma_j\underline{\nu}_j$ for $j=1,\ldots,s$ and $\sigma_j=\pm1$. The sum of the considered graph values can be expressed by an interpolation formula; let us define $\underline{\nu}(\underline{t}_{v_0})=\underline{\nu}_0+2\sum_{j=1}^s t_j\underline{\nu}_j$ fot $\underline{t}_{v_0}=(t_1,\ldots,t_s)$, then the sum is:

3.5

$$\int_0^1 \Big(\prod_{j=1}^{s} dt_j\,\frac{\partial}{\partial_{t_j}}\Big)\Big\{\frac{\underline{e}\cdot\underline{a}}{\underline{\omega}_0\cdot\underline{\nu}_{v_0}(\underline{t}_{v_0})}\frac{1}{|\underline{\nu}_{v_0}(\underline{t}_{v_0})|^b}
\prod_{j=1}^{k-1}\frac{\underline{\nu}_{v_0}(\underline{t}_{v_0})\cdot\underline{a}}{\underline{\omega}_0\cdot\underline{\nu}_j}\frac{1}{|\underline{\nu}_j|^b}\Big\} \tag{3.5}$$

and the *cancellation* is expressed by the fact that in the above interpolation formula there are no contributions from the boundary values $t_j = 0$.

The (3.5) can be bounded as before *but replacing each of the $s$ vectors $\underline{\nu}_{v_0}$ which are large with the corresponding* $|\underline{\nu}_j|$, and (trivially) by replacing the other $k-1-s$ and $\underline{\nu}_0$'s by the correponding $|\underline{\nu}_j|$; *i.e.* for a suitable $\tilde{C}$ the bound (3.3) is replaced by:

3.6

$$\tilde{C}^k|\underline{\nu}_0|^{\tau-b}\prod_{j=1}^{k-1}|\underline{\nu}_j|^{\tau+1-b} \tag{3.6}$$

and one needs (to sum over $\underline{\nu}_j$) only $b > \ell+\tau+1$, *i.e.* $f \in \hat{C}^{(>\tau+1)}(T^\ell)$.

This is the key "ultraviolet" cancellation introduced and used in [BGGM1] and [BGGM2]. Clearly there remains quite a lot of work to do:

(1) how does one treat the cases in which the $\underline{\nu}_0$ and the $\underline{\nu}_j$ are "close"? In such cases $\underline{\nu}_{v_0}(\underline{t}_{v_0} = \underline{\nu}_0 + 2\sum_{j=1}^{k-1} t_j\,\underline{\nu}_j$ 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 [E]? are they obvious even considering the necessity of collecting the 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 eventualy 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 node of the graph that was previouly called $v_0$ and $w_1,\ldots,w_{p_{v_1}}$ the $k-1$ endnodes.

The change in notation is made to adhere to the notation in [BGGM2] which is apt to treat also the more structered graphs.

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

We shall set $\underline{\nu}_{w_j} \equiv \underline{\nu}_{\lambda_{w_j}}$, again for ease of reference to [BGGM2]. Note that in the graph (3.1) $\underline{\nu}_{w_j}$ and $\underline{\nu}_{w_j}$ accidentally coincide because of the simplicity opf the graph.

Fixed $\underline{\nu}$ and $\{\underline{\nu}_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}$ let $h_{v_1}$ be *the scale of* $\underline{\nu}_{v_1}$ (it depends on $\underline{\nu}$ and on the $\underline{\nu}_{\lambda_{w_j}}$ because $\underline{\nu}_{v_1} = \underline{\nu} - \sum \underline{\nu}_{w_j}$). This means that $\underline{\nu}_{v_1}$ is such that $2^{h_{v_1}-1} \le |\underline{\nu}_{v_1}| < 2^{h_{v_1}}$ and we refer to this relation by saying thet $h_{v_1}$ is *compatible* with $\underline{\nu}_1$.

With the new notations the value of the graph is:

4.1
$$\frac{\underline{e}\cdot\underline{a}}{|\underline{\nu}_{v_1}|^b\,\underline{\omega}_0\cdot\underline{\nu}}\prod_{j=1}^{p_{v_1}}\frac{\underline{\nu}_1\cdot\underline{a}}{|\underline{\nu}_j|^b\,\underline{\omega}_0\cdot\underline{\nu}_j}\qquad\qquad \underline{\nu}=\sum_{j=0}^{k-1}\underline{\nu}_j \tag{4.1}$$

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

4.2
$$2^{h_{v_1}} > 2^o p_{v_1}|\underline{\nu}_{\lambda_w}| \qquad o=5 \tag{4.2}$$

where $p_v$ is the number of branches entering $v$.

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

Given a set $\{\underline{\nu}_{\lambda_w}\}_{w\in B_{v_1}}$ for all choices of $\sigma_w=\pm1$ we define the transformation:

4.3
$$U(\{\underline{\nu}_{\lambda_w}\}_{w\in B_{v_1}}) \equiv \{\sigma_w\,\underline{\nu}_{\lambda_w}\}_{w\in B_{v_1}}) \tag{4.3}$$

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 $w\notin C$.

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

- the first quarter of $I_h$ the *lower part* $I_h^- = [2^{h-1},\frac54 2^{h-1})$ of $I_h$,
- the fourth quarter of $I_h$ the *upper part* $I_h^+ = [\frac78 2^h, 2^h)$ of $I_h$, and
- the remaining part the *central part* $I_h^c$.

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

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

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

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

- the *external momenta collection* to be the set of momenta

4.5
$$\underline{\nu}^{1U}_{v_1} = \underline{\nu} - \sum_{w\in\mathcal{B}_{v_1}}\sigma_w\,\underline{\nu}^1_{\lambda_w},\qquad \text{for } U\in\mathcal{U}(\mathcal{B}_{1v_1}(\underline{\nu},\{\underline{\nu}^1_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}})) \tag{4.5}$$

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

We consider then another representative $\{\underline{\nu}^2_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}$ such that $|\underline{\nu}^2_{v_1}| \in I^c_h$ *and* not belonging to the branch momenta collection associated with $\{\underline{\nu}^1_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}$, 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 $\underline{\nu}_1$ is in $I^c_h$, 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 $I^c_h$, always being careful not to consider representatives $\{\underline{\nu}_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}$ that appeared as members of previously constructed collections. It is worth pointing out that not all the external momenta $\underline{\nu}^U_{v_1}$, $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\underline{\nu}, \{\underline{\nu}^1_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}))$, 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 $h = 2, 3\ldots$ cases, we shall have grouped the sets of branch momenta into collections obtainable from a representative $\{\underline{\nu}_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}$ by applying the operations $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\underline{\nu}, \{\underline{\nu}^1_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}))$ to it. *Note that, in this way, when the interval $I^+_{h-1}$ is considered, all the remaining representatives are such that $|\underline{\nu}^U_{v_1}| \in I^+_{h-1}$ for all $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\underline{\nu}, \{\underline{\nu}^1_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}))$.*

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:

4.6
$$\begin{aligned}
&\underline{\nu}^U_{v_1} = \Big(\prod_{w\in\mathcal{B}_{1v_1}} U^{\sigma_w}_{v_1 w}\{\underline{\nu}_x\}\Big)_{v_1} \;" = "\; \underline{\nu} - \sum_w \sigma_w \underline{\nu}_w \\
&(U(\{\underline{\nu}_{\lambda_{\tilde w}}\}_{\tilde w\in B_{v_1}}))_w = \sum_{z\le w} \Big(\prod_{\tilde w\in\mathcal{B}_{1v_1}} U^{\sigma_{\tilde w}}_{v_1\tilde w}\{\underline{\nu}_x\}\Big)_z \;" = "\; \sigma_w \underline{\nu}_w
\end{aligned} \tag{4.6}$$

where, given the sets of momenta $\{\underline{\nu}_x\}$ and $\{\underline{\nu}_{\lambda_{\tilde w}}\}$, $(\{\underline{\nu}_x\})_v$ denotes the external momentum in $\{\underline{\nu}_x\}$ corresponding to the node $v$ and $(\{\underline{\nu}_{\lambda_{\tilde w}}\})_w$ denotes the branch momentum in $\{\underline{\nu}_{\lambda_{\tilde w}}\}$ corresponding to the branch $\lambda_w$. The " = " mean that in the special case of (3.1) that we treat the equality between $\underline{\nu}_w$ and $\underline{\nu}_{\lambda_w}$ the r.h.s. of (4.6) take the simple form indicated.

**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 $U \in \mathcal{U}(\mathcal{B}_{1v_1}(\underline{\nu}, \{\underline{\nu}_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}})$, one has:

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

because the scale of $\underline{\nu}^U_{v_1}$ may be $h-1$, while that of $\underline{\nu}_{v_1}$ 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.

A convenient way to rewrite the sum that we want to estimate, *i.e.* the contribution to $\underline{h}^{(k)}$ form the sum of the values of the graphs like (3.1) is the following:

4.8

$$\sum_{\underline{\nu}}|\underline{\nu}|^s\Big|\sum_{h_{v_1}}\overset{*}{\sum_{\{\underline{\nu}_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}}}\sum_{U\in\mathcal{U}(B_{1v_1})}\frac{\underline{e}\cdot\underline{a}}{\underline{\omega}_0\cdot\underline{\nu}\,|\underline{\nu}^U_{v_1}|^b}\Big(\prod_{w\in\mathcal{B}_{1v_1}}\sigma_w\Big)\prod_{j=1}^{p_{v_1}}\frac{\underline{\nu}_{v_1}\cdot\underline{a}}{\underline{\omega}_0\cdot\underline{\nu}_{w_j}|\underline{\nu}_{w_j}|^b} \tag{4.8}$$

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

The explicit sum over the scales $h_{v_1}$ is introduced to simplify the bounds analysis that we perform now. Note that $\underline{\nu}^U_{v_1}$ 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 (4.10) below).

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

4.10

$$\sum_{\underline{\nu}}|\underline{\nu}|^s\Big|\sum_{h_{v_1}}\overset{*}{\sum_{\{\underline{\nu}_{\lambda_w}\}_{w\in\mathcal{B}_{v_1}}}}\Big[\Big(\prod_{w\in\mathcal{B}_{1v_1}}\int_1^0 dt_w\Big)\cdot \cdot\Big(\prod_{w\in\mathcal{B}_{1v_1}}\frac{\partial}{\partial t_w}\Big)\Big(\frac{1}{(\underline{\omega}_0\cdot\underline{\nu})\,|\underline{\nu}_{v_1}(\underline{t}_{v_1})|^b}\prod_{j=1}^{k-1}\frac{\underline{\nu}_{v_1}(\underline{t}_{v_1})\cdot\underline{a}}{(\underline{\omega}_0\cdot\underline{\nu}_{w_j})|\underline{\nu}_{w_j}|^b}\Big) \tag{4.9}$$

where if $\mathcal{B}_{1v_1}=\emptyset$ no interpolation is made; and we note that by the definition of nodes out of order and by the iterative grouping of the representatives:

4.11

$$2^{h_{v_1}-2}\le|\underline{\nu}_{v_1}(\underline{t}_{v_1})|<2^{h_{v_1}}\ , \tag{4.10}$$

so that the interpolation formulae discussed in §3 *can be used* because no singularity (*i.e.* no division by 0) arises in performing the $\underline{t}_{v_1}$–integrations.

It is now easy to check, by computing the derivatives, that each of the terms in (4.9) can be conveniently bounded by its (easily estimated) maximum value and the resulting bound has the form obtained in the simple case of §3, (3.6), in which some of the $\underline{\nu}_{w_j}$ were $\ll|\underline{\nu}_{v_1}|$ and some were $\gg|\underline{\nu}_{v_1}|$. The $\underline{\nu}_w$ of the nodes out of order play the role of the small $\underline{\nu}_{w_j}$'s and the others can be used as bounds on $\underline{\nu}_{v_1}$. *No further care needs to be exercized in the bounds.* The details are not very interesting, see [BGGM2].

The extension to more structured graphs is not really more difficult. One has to be careful in interpreting $\underline{\nu}_{w_j}$ as the momentum flowing on the brach $v_1w_j$, *i.e.* as the sum of the node momenta of all the nodes above $v_1$. See [BGGM1] for the detailed algebra that is necessary. Therefore the above analysis gives a rater 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 [BGGM2]. 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).

**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 [BGGM1], [BGGM2].

## References

[BGGM1] F. Bonetto, G. Gallavotti, G. Gentile, V. Mastropietro: *Lindstedt series, ultraviolet divergences and Moser's theorem*, Preprint IHES/P/95/102 (1995).

[BGGM2] 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.

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

[G1] G. Gallavotti: *Twistless KAM tori*, Communications in Mathematical Physics, **164**, 145–156 (1994).

[GG] G. Gallavotti, G. Gentile: *Majorant series method for KAM tori*, Ergodic theory and applications, **15**, p. 857–869, 1995.

[GM1] G. Gentile, V. Mastropietro: *KAM theorem revisited*, Physica D, **90**, 225–234 (1996).

[GM2] 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, , – , 1996.

[GGM] 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).

[HP] F. Harary, E. Palmer: *Graphical enumeration*, Academic Press, New York, 1973.

[M1] J. Moser: *Convergent series expansions for quasi periodic motions*, *Mathematische Annalen*, **169**, 136-176 (1967).

[M2] J. Moser: *A rapidly convergent iteration method and nonlinear differential equations II*, Annali della Scuola Normale Superiore di Pisa Serie III, **20**, 499–535, (1966).

[P] Poincaré, H.: *Les Méthodes Nouvelles de la Mécanique Céleste*, Gauthiers Villars, Paris, 1893.

[Pa] I.S. Parasyuk: *On instability zones of the Schrödinger equation with a quasi periodic potential*, Ukr. Mat. Zh. **30**, 70-78 (1985).

[PF] L.A. Pastur, A. Figotin: *Spectra of random and almost-periodic operators*, Grundlehren der Mathematischen Wissenschaften **297**, Springer, Berlin (1992).

*Internet access: All the Author's quoted preprints can be found and freely downloaded (latest postscript version including corrections of misprints and errors) at:* `http://chimera.roma1.infn.it` *in the Mathematical Physics Preprints page; mirror at* `http://www.math.rutgers.edu/ giovanni`.

*e-mail address of author: giovanni@ipparco.roma1.infn.it*