Diagrammatic Methods in Classical Perturbation Theory

Guido Gentile

Dipartimento di Matematica, Università di Roma Tre, Roma, I-00146, Italy.
E-mail: gentile@mat.uniroma3.it

Article outline

Glossary

Dynamical system
Let $W \subseteq \mathbb{R}^N$ be an open set and $f : W \times \mathbb{R} \to \mathbb{R}^N$ be a smooth function. The ordinary differential equation $\dot{x} = f(x,t)$ on W defines a continuous dynamical system. A discrete dynamical system on W is defined by a map $x \to x' = F(x)$, with F depending smoothly on x.

Hamiltonian system
Let $\mathcal{A} \subseteq \mathbb{R}^d$ be an open set and $\mathcal{H} : \mathcal{A} \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ be a smooth function ($\mathcal{A} \times \mathbb{R}^d$ is called the phase space). Consider the system of ordinary differential equations $\dot{q}_k = \partial \mathcal{H}(q,p,t)/\partial p_k$, $\dot{p}_k = -\partial \mathcal{H}(q,p,t)/\partial q_k$, for $k = 1, \dots, d$. The equations are called Hamilton equations, and $\mathcal{H}$ is called a Hamiltonian function. A dynamical system described by Hamilton equations is called a Hamiltonian system.

Integrable system
A Hamiltonian system is called integrable if there exists a system of coordinates $(\alpha, A) \in \mathbb{T}^d \times \mathbb{R}^d$, called angle-action variables, such that in these coordinates the motion is $(\alpha, A) \to (\alpha + \omega(A)t, A)$, for some smooth function $\omega(A)$. Hence in these coordinates the Hamiltonian function $\mathcal{H}$ depends only on the action variables, $\mathcal{H} = \mathcal{H}_0(A)$.

Invariant torus
Given a continuous dynamical system we say that the motion occurs on an invariant d-torus if it takes

place on a d-dimensional manifold and its position on the manifold is identified through a coordinate in $\mathbb{T}^d$. In an integrable Hamiltonian system all phase space is filled by invariant tori. In a quasi-integrable system the KAM theorem states that most of the invariant tori persist under perturbation, in the sense that the relative Lebesgue measure of the fraction of phase space filled by invariant tori tends to 1 as the perturbation tends to disappear. The persisting invariant tori are slight deformations of the unperturbed invariant tori.

Quasi-integrable system

A quasi-integrable system is a Hamiltonian system described by a Hamiltonian function of the form $\mathcal{H} = \mathcal{H}_0(A) + \varepsilon f(\alpha, A)$, with (α, A) angle-action variables, ε a small real parameter and f periodic in its arguments α.

Quasi-periodic motion

Consider the motion $\alpha \to \alpha + \omega t$ on $\mathbb{T}^2$, with $\omega = (\omega_1, \omega_2)$. If ω_1/ω_2 is rational, the motion is periodic, that is there exists $T > 0$ such that $\omega_1 T = \omega_2 T = 0 \mod 2\pi$. If ω_1/ω_2 is irrational, the motion never returns to its initial value. On the other hand it densely fills $\mathbb{T}^2$, in the sense that it comes arbitrarily close to any point of $\mathbb{T}^2$. We say in that case that the motion is quasi-periodic. The definition extends to $\mathbb{T}^d$, $d > 2$: a linear motion $\alpha \to \alpha + \omega t$ on $\mathbb{T}^d$ is quasi-periodic if the components of ω are rationally independent, that is if $\omega \cdot \nu = \omega_1\nu_1 + \ldots + \omega_d\nu_d = 0$ for $\nu \in \mathbb{Z}^d$ if and only if $\nu = 0$ ($a \cdot b$ is the standard scalar product between the two vectors a, b). More generally we say that a motion on a manifold is quasi-periodic if, in suitable coordinates, it can be described as a linear quasi-periodic motion. The vector ω is usually called frequency or rotation vector.

Renormalisation group

By renormalisation group one denotes the set of techniques and concepts used to study problems where there are some scale invariance properties. The basic mechanism consists in considering equations depending on some parameters and defining some transformations on the equations, including a suitable rescaling, such that after the transformation the equations can be expressed, up to irrelevant corrections, in the same form as before but with new values for the parameters.

Torus

The 1-torus $\mathbb{T}$ is defined as $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, that is the set of real numbers defined modulo 2π (this means that x is identified with y if $x - y$ is a multiple of 2π). So it is the natural domain of an angle. One defines the d-torus $\mathbb{T}^d$ as a product of d 1-tori, that is $\mathbb{T}^d = \mathbb{T} \times \ldots \times \mathbb{T}$. For instance one can imagine $\mathbb{T}^2$ as a square with the opposite sides glued together.

Tree

A graph is a collection of points, called nodes, and of lines which connect the nodes. A walk on the graph is a sequence of lines such that any two successive lines in the sequence share a node; a walk is nontrivial if it contains at least one line. A tree is a planar graph with no closed loops, that is, such that there is no nontrivial walk connecting any node to itself. An oriented tree is a tree with a special node such that all lines of the tree are oriented toward that node. If we add a further oriented line connecting the special node to another point, called the root, we obtain a rooted tree (see Figure 1 in Section 4).

1 Definition of the Subject and Its Importance

Recursive equations naturally arise whenever a dynamical system is considered in the regime of perturbation theory; for an introductory article on perturbation theory see Gallavotti (2008). A classical example is provided by Celestial Mechanics, where perturbation series, known as Lindstedt series, are widely used; see Gallavotti (2006) and Celletti (2008).

A typical problem in Celestial Mechanics is to study formal solutions of given ordinary differential

equations in the form of expansions in a suitable small parameter, the perturbation parameter. In the case of quasi-periodic solutions, the study of the series, in particular of its convergence, is made difficult by the presence of the small divisors – which will be defined later on. Under some non-resonance condition on the frequency vector, one can show that the series are well-defined to any order. The first proof of such a property was given by Poincaré (1892-1899), even if the convergence of the series remained an open problem up to the advent of KAM theory – an account can be found in Gallavotti (1983) and in Arnold *et al.* (1988); see also Chierchia (2008). KAM is an acronym standing for Kolmogorov (1954), Arnold (1963) and Moser (1962), who proved in the middle of last century the persistence of most of invariant tori for quasi-integrable systems.

Kolmogorov and Arnold proofs apply to analytic Hamiltonian systems, while Moser's approach deals also with the differentiable case; the smoothness condition on the Hamiltonian function was therafter improved by Pöschel (1982). In the analytic case, the persisting tori turn out to be analytic in the perturbation parameter, as explicitly showed by Moser (1967). In particular, this means that the perturbation series not only are well-defined, but also converge. However, a systematic analysis with diagrammatic techniques started only recently after the pioneering, fundamental works by Eliasson (1996) and Gallavotti (1994), and were subsequently extended to many other problems with small divisors, including dynamical systems with infinitely many degrees of freedom, such as nonlinear partial differential equations, and non-Hamiltonian systems. Some of these extensions will be discussed in Section 8.

From a technical point of view, the diagrammatic techniques used in classical perturbation theory are strongly reminiscent of the Feynman diagrams used in quantum field theory: this was first pointed out by Gallavotti (1994). Also the multiscale analysis used to control the small divisors is typical of renormalisation group techniques, which have been successfully used in problems of quantum field theory, statistical mechanics and classical mechanics; see Gallavotti (2001) and Gentile & Mastropietro (2002) for some reviews.

Note that there exist other renormalisation group approaches to the study of dynamical systems, and of KAM-like problems in particular, different from that outlined in this article. By confining ourselves to the framework of problems of KAM-type, we can mention the paper by Bricmont *et al.* (1999), which also stressed the similarity of the technique with quantum field theory, and the so called dynamical renormalisation group method – see MacKay (1993) – which recently produced rigorous proofs of persistence of quasi-periodic solutions; see for instance Koch (1997) and Khanin *et al.* (2007).

2 Introduction

Consider the ordinary differential equation on $\mathbb{R}^d$

$$Du = G(u) + \varepsilon F(u), \tag{2.1}$$

where D is a pseudo-differential operator and G, F are real analytic functions. Assume that the equation (2.1) admits a solution $u^{(0)}(t)$ for $\varepsilon = 0$, that is $Du^{(0)} = G(u^{(0)})$. The problem we are interested in is to investigate whether there exists a solution of (2.1) which reduces to $u^{(0)}$ as $\varepsilon \to 0$. For simplicity assume $G = 0$ in the following.

The first attempt one can try is to look for solutions in the form of power series in ε,

$$u(t) = \sum_{k=0}^{\infty} \varepsilon^k u^{(k)}(t), \tag{2.2}$$

which, inserted into (2.1), when equating the left and right hand sides order by order, gives the list of recursive equations $Du^{(0)} = 0$, $Du^{(1)} = F(u^{(0)})$, $Du^{(2)} = \partial_u F(u^{(0)})\, u^{(1)}$, and so on. In general to order

$k \geq 1$ one has

$$Du^{(k)} = \sum_{s=0}^{k-1} \frac{1}{s!} \partial_u^s F(u^{(0)}) \sum_{\substack{k_1+\ldots+k_s=k-1 \\ k_i \geq 1}} u^{(k_1)} \ldots u^{(k_s)}, \tag{2.3}$$

where $\partial_u^s F$, the s-th derivative of F, is a tensor with $s+1$ indices (s must be contracted with the vectors $u^{(k_1)}, \ldots, u^{(k_s)}$), and the term with $s = 0$ in the sum has to be interpreted as $F(u^{(0)})$ and appears only for $k = 1$.

For instance for $F(u) = u^3$ the first orders give

$$\begin{aligned} Du^{(1)} &= u^{(0)3}, \\ Du^{(2)} &= 3u^{(0)2}\, u^{(1)}, \\ Du^{(3)} &= 3u^{(0)2}\, u^{(2)} + 3u^{(0)}\, u^{(1)2}, \\ Du^{(4)} &= 3u^{(0)2}\, u^{(3)} + 6u^{(0)}\, u^{(1)}\, u^{(2)} + u^{(1)3}, \end{aligned} \tag{2.4}$$

as is easy to check.

If the operator D can be inverted then the recursions (2.3) provide an algorithm to compute the functions $u^{(k)}(t)$. In that case we say that (2.2) defines a *formal power series*: by this we mean that the functions $u^{(k)}(t)$ are well-defined for all $k \geq 0$. Of course, even if this can be obtained, there is still the issue of the convergence of the series that must be dealt with.

3 Examples

In this section we consider a few paradigmatic examples of dynamical systems which can be described by equations of the form (2.1).

3.1 A class of quasi-integrable Hamiltonian systems

Consider the Hamiltonian system described by the Hamiltonian function

$$\mathcal{H}(\alpha, A) = \frac{1}{2} A^2 + \varepsilon f(\alpha), \tag{3.1}$$

where $(\alpha, A) \in \mathbb{T}^d \times \mathbb{R}^d$ are angle-action variables, with $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, f is a real analytic function, 2π-periodic in each of its arguments, and $A^2 = A \cdot A$, if (here and henceforth) $\cdot$ denotes the standard scalar product in $\mathbb{R}^d$, that is $a \cdot b = a_1 b_1 + \ldots + a_d b_d$. Assume also for simplicity that f is a trigonometric polynomial of degree N.

The corresponding Hamilton equations are (we shorten $\partial_x = \partial/\partial_x$)

$$\dot{\alpha} = \partial_A \mathcal{H}(\alpha, A) = A, \qquad \dot{A} = -\partial_\alpha \mathcal{H}(\alpha, A) = -\varepsilon \partial_\alpha f(\alpha),$$

which can be written as an equation involving only the angle variables:

$$\ddot{\alpha} = -\varepsilon \partial_\alpha f(\alpha), \tag{3.2}$$

which is of the form (2.1) with $u = \alpha$, $G = 0$, $F = -\partial_\alpha f$, and $D = \mathrm{d}^2/\mathrm{d}t^2$.

For $\varepsilon = 0$, $\alpha^{(0)}(t) = \alpha_0 + \omega t$ is a solution of (3.2) for any choice of $\alpha_0 \in \mathbb{T}^d$ and $\omega \in \mathbb{R}^d$. Take for simplicity $\alpha_0 = 0$: we shall see that this choice makes sense. We say that for $\varepsilon = 0$ the Hamiltonian function (3.1) describes a system of d *rotators*.

We call ω the *frequency vector*, and we say that ω is irrational if its components are rationally independent, that is if $\omega \cdot \nu = 0$ for $\nu \in \mathbb{Z}^d$ if and only if $\nu = 0$. For irrational ω the solution $\alpha^{(0)}(t)$ describes a *quasi-periodic motion* with frequency vector ω, and it densely fills $\mathbb{T}^d$.

Then (2.3) becomes

$$\ddot{\alpha}^{(k)} = -\left[\varepsilon f(\alpha)\right]^{(k)} := -\sum_{s=0}^{k-1} \frac{1}{s!} \partial_\alpha^{s+1} f(\omega t) \sum_{\substack{k_1+\ldots+k_s=k-1 \\ k_i \geq 1}} \alpha^{(k_1)} \ldots \alpha^{(k_s)}. \tag{3.3}$$

We look for a quasi-periodic solution of (3.2), that is a solution of the form $\alpha(t) = \omega t + h(\omega t)$, with $h(\omega t) = O(\varepsilon)$. We call h the *conjugation function*, as it "conjugates" (that is, maps) the perturbed solution $\alpha(t)$ to the unperturbed solution ωt. In terms of the function h (3.2) becomes

$$\ddot{h} = -\varepsilon \partial_\alpha f(\omega t + h), \tag{3.4}$$

where ∂_α denotes derivative with respect to the argument. Then (3.4) can be more conveniently written in Fourier space, where the operator D acts as a multiplication operator.

If we write

$$h(\omega t) = \sum_{\nu \in \mathbb{Z}^d} \mathrm{e}^{i\omega\cdot\nu t} h_\nu, \qquad h_\nu = \sum_{k=1}^{\infty} \varepsilon^k h_\nu^{(k)}, \tag{3.5}$$

and insert (3.5) into (3.4) we obtain

$$(\omega\cdot\nu)^2 h_\nu^{(k)} = \left[\varepsilon\partial_\alpha f(\alpha)\right]_\nu^{(k)} := \sum_{s=0}^{k-1} \sum_{\substack{k_1+\ldots+k_s=k-1 \\ k_i\geq 1}} \sum_{\substack{\nu_0+\nu_1+\ldots+\nu_s=\nu \\ \nu_i\in\mathbb{Z}^n}} \frac{1}{s!} (i\nu_0)^{s+1} f_{\nu_0} h_{\nu_1}^{(k_1)} \ldots h_{\nu_s}^{(k_s)}. \tag{3.6}$$

These equations are well-defined to all orders provided $\left[\varepsilon\partial_\alpha f(\alpha)\right]_\nu^{(k)} = 0$ for all ν such that $\omega\cdot\nu = 0$. If ω is an irrational vector we need $\left[\varepsilon\partial_\alpha f(\alpha)\right]_0^{(k)} = 0$ for the equations to be well-defined. In that case the coefficients $h_0^{(k)}$ are left undetermined, and we can fix them arbitrarily to vanish (which is a convenient choice).

We shall see that under some condition on ω a quasi-periodic solution $\alpha(t)$ exists, and densely fills a d-dimensional manifold. The analysis carried out above for $\alpha_0 = 0$ can be repeated unchanged for all values of $\alpha_0 \in \mathbb{T}^d$: α_0 represents the *initial phase* of the solution, and by varying α_0 we cover all the manifold. Such a manifold can be parameterised in terms of α_0, so it represents an invariant torus for the perturbed system.

3.2 A simplified model with no small divisors

Consider the same equation as (3.4) with $D = \mathrm{d}^2/\mathrm{d}t^2$ replaced by -1, that is

$$h = \varepsilon\partial_\alpha f(\omega t + h). \tag{3.7}$$

Of course in this case we no longer have a differential equation; still, we can look again for quasi-periodic solutions $h(\omega t) = O(\varepsilon)$ with frequency vector ω. In such a case in Fourier space we have

$$h_\nu^{(k)} = \left[\varepsilon\partial_\alpha f(\alpha)\right]_\nu^{(k)} := \sum_{s=0}^{k-1} \sum_{\substack{k_1+\ldots+k_s=k-1 \\ k_i\geq 1}} \sum_{\substack{\nu_0+\nu_1+\ldots+\nu_s=\nu \\ \nu_i\in\mathbb{Z}^n}} \frac{1}{s!} (i\nu_0)^{s+1} f_{\nu_0} h_{\nu_1}^{(k_1)} \ldots h_{\nu_s}^{(k_s)}. \tag{3.8}$$

For instance if $d = 1$ and $f(\alpha) = \cos\alpha$ the equation, which is known as the *Kepler equation*, can be explicitly solved by the Lagrange inversion theorem (Wintner, 1941), and gives

$$h_\nu^{(k)} = \begin{cases} -i\dfrac{(-1)^{k+(\nu-k)/2}}{2^k}\dfrac{\nu^{k-1}}{((k-\nu)/2)!((k+\nu)/2)!}, & |\nu| \le k,\ \nu+k \text{ even}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.9}$$

We shall show in Section 5 that a different derivation can be provided by using the forthcoming diagrammatic techniques.

3.3 The standard map

Consider the finite difference equation

$$D\alpha = -\varepsilon\sin\alpha, \tag{3.10}$$

on $\mathbb{T}$, where now D is defined by

$$D\alpha(\psi) := 2\alpha(\psi) - \alpha(\psi+\omega) - \alpha(\psi-\omega). \tag{3.11}$$

By writing $\alpha = \psi + h(\psi)$, (3.10) becomes

$$Dh = -\varepsilon\sin(\psi+h), \tag{3.12}$$

which is the functional equation that must be solved by the *conjugation function* of the *standard map*

$$\begin{cases} x' = x + y + \varepsilon\sin x, \\ y' = y + \varepsilon\sin x. \end{cases} \tag{3.13}$$

In other words, by writing $x = \psi + h(\psi)$ and $y = \omega + h(\psi) - h(\psi-\omega)$, with (ψ,ω) solving (3.13) for $\varepsilon = 0$, that is $(\psi',\omega') = (\psi+\omega,\omega)$, we obtain a closed-form equation for h, which is exactly (3.12).

In Fourier space the operator D acts as $D\colon \mathrm{e}^{i\nu\psi} \to 4\sin^2(\omega\nu/2)\,\mathrm{e}^{i\nu\psi}$, so that, by expanding h according to (3.5), we can write (3.12) as

$$h_\nu^{(k)} = \frac{1}{4\sin^2(\omega\nu/2)}\sum_{s=0}^{k-1}\sum_{\substack{k_1+\ldots+k_s=k-1\\ k_i\ge 1}}\sum_{\substack{\nu_0+\nu_1+\ldots+\nu_s=\nu\\ \nu_i\in\mathbb{Z}}}\frac{1}{s!}(i\nu_0)^{s+1}f_{\nu_0}h_{\nu_1}^{(k_1)}\ldots h_{\nu_s}^{(k_s)}, \tag{3.14}$$

where $\nu_0 = \pm 1$ and $f_{\pm 1} = 1/2$.

Note that (3.13) is a discrete dynamical system. However, when passing to Fourier space, the equation (3.14) acquires the same form as for the continuous dynamical systems previously considered, simply with a different kernel for D. In particular if we replace D with 1 we recover the Kepler equation.

The number ω is called the *rotation number*. We say that ω is irrational if the vector $(2\pi,\omega)$ is irrational according to the previous definition.

4 Trees and Graphical Representation

Take ω to be irrational. We study the recursive equations

$$\begin{cases} h_\nu^{(k)} = g(\omega\cdot\nu)\,[\varepsilon\partial_\alpha f(\alpha)]_\nu^{(k)}, & \nu \ne 0, \\ [\varepsilon\partial_\alpha f(\alpha)]_0^{(k)} = 0, & \nu = 0, \end{cases} \tag{4.1}$$

where the form of g depends on the particular model we are investigating. Hence one has either $g(\omega\cdot\nu) = (\omega\cdot\nu)^{-2}$ or $g(\omega\cdot\nu) = 1$ or $g(\omega\cdot\nu) = (2\sin(\omega\nu/2))^{-2}$ according to models described in Section 3.

For $\nu \neq 0$ we have equations which express the coefficients $h_\nu^{(k)}$, $\nu \in \mathbb{Z}^d$, in terms of the coefficients $h_\nu^{(k')}$, $\nu \in \mathbb{Z}^d$, with $k' < k$, provided the equations for $\nu = 0$ are satisfied for all $k \geq 1$. Recursive equations, such as (4.1), naturally lead to a graphical representation in terms of trees.

4.1 Trees

A connected *graph* $\mathcal{G}$ is a collection of points (nodes) and lines connecting all of them. Denote with $N(\mathcal{G})$ and $L(\mathcal{G})$ the set of nodes and the set of lines, respectively. A path between two nodes is the minimal subset of $L(\mathcal{G})$ connecting the two nodes. A graph is planar if it can be drawn in a plane without graph lines crossing.

A *tree* is a planar graph $\mathcal{G}$ containing no closed loops. Consider a tree $\mathcal{G}$ with a single special node v_0: this introduces a natural partial ordering on the set of lines and nodes, and one can imagine that each line carries an arrow pointing toward the node v_0. We add an extra oriented line ℓ_0 exiting the special node v_0; the added line will be called the *root line* and the point it enters (which is not a node) will be called the *root* of the tree. In this way we obtain a *tree* θ defined by $N(\theta) = N(\mathcal{G})$ and $L(\theta) = L(\mathcal{G}) \cup \ell_0$. A *labelled tree* is a rooted tree θ together with a label function defined on the sets $L(\theta)$ and $N(\theta)$.

We call *equivalent* two rooted trees which can be transformed into each other by continuously deforming the lines in the plane in such a way that the lines do not cross each other. We can extend the notion of equivalence also to labelled trees, by considering equivalent two labelled trees if they can be transformed into each other in such a way that the labels also match. In the following we shall deal mostly with nonequivalent labelled trees: for simplicity, where no confusion can arise, we call them just trees.

Given two nodes $v, w \in N(\theta)$, we say that $w \prec v$ if v is on the path connecting w to the root line. We can identify a line ℓ through the node v it exits by writing $\ell = \ell_v$.

We call *internal nodes* the nodes such that there is at least one line entering them, and *end-points* the nodes which have no entering line. We denote with $L(\theta)$, $V(\theta)$ and $E(\theta)$ the set of lines, internal nodes and end-points, respectively. Of course $N(\theta) = V(\theta) \cup E(\theta)$.

The number of unlabelled trees with k nodes (and hence with k lines) is bounded by 2^{2k}, which is a bound on the number of random walks with $2k$ steps (Gentile & Mastropietro, 2001).

For each node v denote by $S(v)$ the set of the lines entering v and set $s_v = |S(v)|$. Hence $s_v = 0$ if v is an end-node, and $s_v \geq 1$ if v is an internal node. One has

$$\sum_{v\in N(\theta)} s_v = \sum_{v\in V(\theta)} s_v = k - 1; \tag{4.2}$$

this can be easily checked by induction on the order of the tree. An example of unlabelled tree is represented in Figure 1.

For further details on graphs and trees we refer to the literature; cf. for instance Harary (1969).

4.2 Labels and diagrammatic rules

We associate with each node $v \in N(\theta)$ a *mode* label $\nu_v \in \mathbb{Z}^d$, and with each line $\ell \in L(\theta)$ a *momentum* label $\nu_\ell \in \mathbb{Z}^d$, with the constraint

$$\nu_{\ell_v} = \sum_{\substack{w\in N(\theta)\\ w\preceq v}} \nu_w = \nu_v + \sum_{\ell\in S(v)} \nu_\ell, \tag{4.3}$$

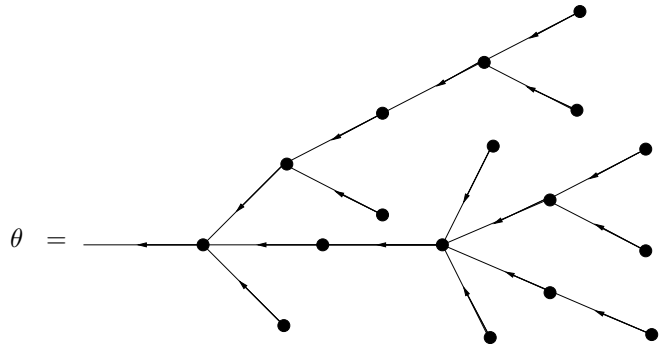

Figure 1: An unlabelled tree with 17 nodes.

which represents a *conservation rule* for each node.

Call $\mathcal{T}_{k,\nu}$ the set of all trees θ with k nodes and momentum ν associated with the root line. We call k and ν the order and the momentum of θ, respectively.

We want to show that trees naturally arise when studying the equations (4.1). Let $h_\nu^{(k)}$ be represented with the graph element in Figure 2 as a line with label ν exiting from a ball with label (k).

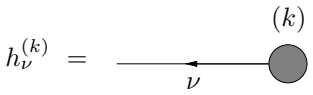

Figure 2: Graph element.

Then we can represent (4.1) graphically as depicted in Figure 3. Simply represent each factor $h_{\nu_i}^{(k_i)}$ on the right hand side as a graph element according to Figure 2. The lines of all such graph elements enter the same node v_0. This is a graphical expedient to recall the conservation rule: the momentum ν of the root line is the sum of the mode label ν_0 of the node v_0 plus the sum of the momenta of the lines entering v_0.

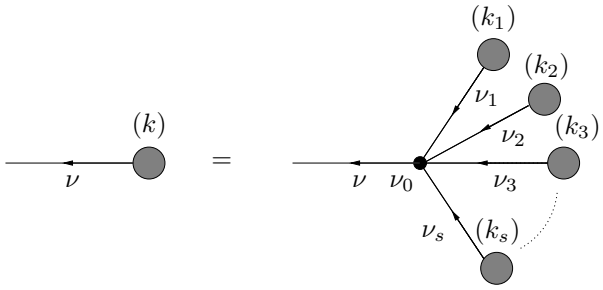

Figure 3: Graphical representation of the recursive equations.

The first few orders $k \leq 4$ are as depicted in Figure 4. For each node the conservation rule (4.3) holds: for instance for $k = 2$ one has $\nu = \nu_1 + \nu_2$, for $k = 3$ one has $\nu = \nu_1 + \nu_{\ell_1}$ and $\nu_{\ell_1} = \nu_2 + \nu_3$ in the first tree and $\nu = \nu_1 + \nu_2 + \nu_3$ in the second tree, and so on. Moreover one has to sum over all possible choices of the labels ν_v, $v \in N(\theta)$, which sum up to ν.

Given any tree $\theta \in \mathcal{T}_{k,\nu}$ we associate with each node $v \in N(\theta)$ a *node factor* F_v and with each line

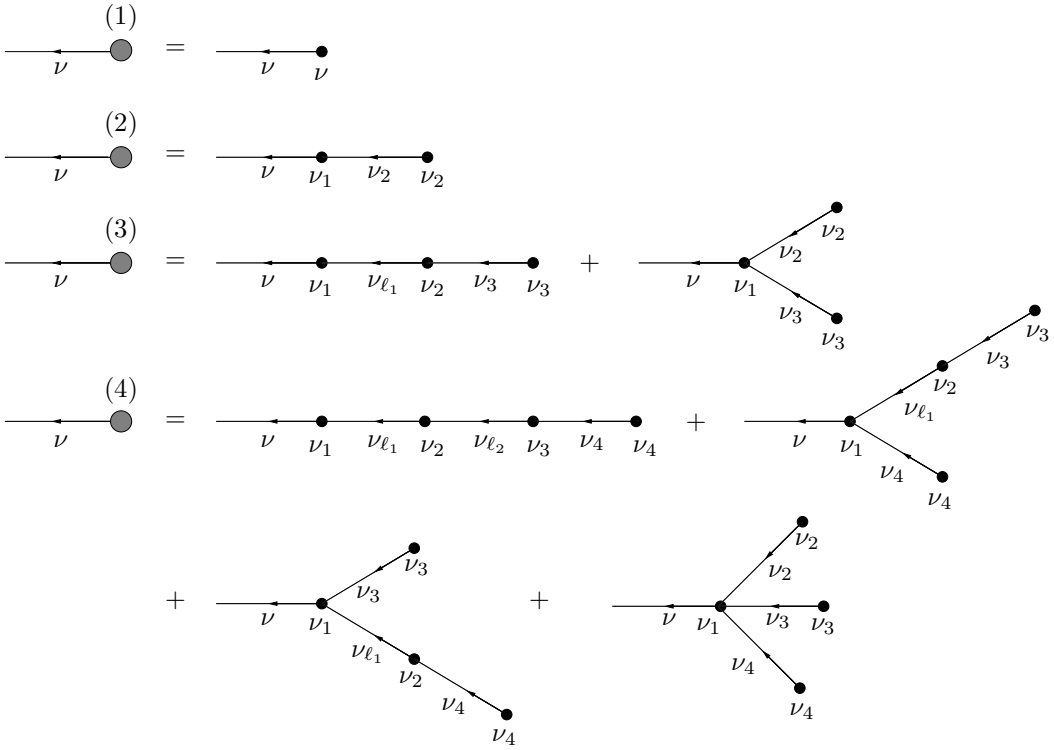

Figure 4: Trees of lower orders.

$\ell \in L(\theta)$ a *propagator* g_ℓ, by setting

$$F_v := \frac{1}{s_v!}(i\nu_v)^{s_v+1} f_{\nu_v}, \qquad g_\ell := g(\omega \cdot \nu_\ell), \tag{4.4}$$

and define the *value* of the tree θ as

$$\mathrm{Val}(\theta) := \Big(\prod_{v \in N(\theta)} F_v \Big) \Big(\prod_{\ell \in L(\theta)} g_\ell \Big). \tag{4.5}$$

The propagators g_ℓ are scalars, whereas each F_v is a tensor with $s_v + 1$ indices, which can be associated with the $s_v + 1$ lines entering or exiting v. In (4.5) the indices of the tensors F_v must be contracted: this means that if a node v is connected to a node v' by a line ℓ then the indices of F_v and $F_{v'}$ associated with ℓ are equal to each other, and eventually one has to sum over all the indices except that associated with the root line.

For instance the value of the tree in Figure 4 contributing to $h_\nu^{(2)}$ is given by

$$\mathrm{Val}(\theta) = (i\nu_1)^2 f_{\nu_1}\, (i\nu_2)\, f_{\nu_2}\, g(\omega \cdot \nu)\, g(\omega \cdot \nu_2),$$

with $\nu_1 + \nu_2 = \nu$, while the value of the last tree in Figure 4 contributing to $h_\nu^{(4)}$ is given by

$$\mathrm{Val}(\theta) = \frac{(i\nu_1)^4}{3!} f_{\nu_1}\, (i\nu_2)\, f_{\nu_2}\, (i\nu_3)\, f_{\nu_3}\, (i\nu_4)\, f_{\nu_4}\, g(\omega \cdot \nu)\, g(\omega \cdot \nu_2)\, g(\omega \cdot \nu_3)\, g(\omega \cdot \nu_4),$$

with $\nu_1 + \nu_2 + \nu_3 + \nu_4 = \nu$.

It is straightforward to prove that one can write

$$h_\nu^{(k)} = \sum_{\theta \in \mathcal{T}_{k,\nu}} \mathrm{Val}(\theta), \qquad \nu \neq 0, \quad k \geq 1. \tag{4.6}$$

This follows from the fact that the recursive equations (4.1) can be graphically represented through Figure 3: one iterates the graphical representation of Figure 3 until only graph elements of order $k = 1$ appear, and if θ is of order 1 (cf. Figure 4) then $\mathrm{Val}(\theta) = (i\nu)\, f_\nu\, g(\omega \cdot \nu)$.

Each line $\ell \in L(\theta)$ can be seen as the root line of the tree consisting of all nodes and lines preceding ℓ. The choice $h_0^{(k)} = 0$ for all $k \geq 1$ implies that no line can have zero momentum: in other words we have $\nu_\ell \neq 0$ for all $\ell \in L(\theta)$.

Therefore in order to prove that (3.5) with $h_\nu^{(k)}$ given by (4.6) solves formally, that is order by order, the equations (4.1), we have only to check that $[\varepsilon \partial_\alpha f(\omega t + h(\omega t))]_0^{(k)} = 0$ for all $k \geq 1$.

If we define $g_\ell = 1$ for $\nu_\ell = 0$, then also the second relation in (4.1) can be graphically represented as in Figure 3 by setting $\nu = 0$ and requiring $h_0^{(k)} = 0$, which yields that the sum of the values of all trees on the right hand side must vanish. Note that this is not an equation to solve, but just an identity that has to be checked to hold at all orders.

For instance for $k = 2$ (the case $k = 1$ is trivial) the identity $[\varepsilon \partial_\alpha f(\omega t + h(\omega t))]_0^{(2)} = 0$ reads (cf. the second line in Figure 4)

$$\sum_{\nu_1 + \nu_2 = 0} (i\nu_1)^2\, f_{\nu_1}\, (i\nu_2)\, f_{\nu_2}\, g(\omega \cdot \nu_2) = 0,$$

which is found to be satisfied because the propagators are even in their arguments.

Such a cancellation can be graphically interpreted as follows. Consider the tree with mode labels ν_1 and ν_2, with $\nu_1 + \nu_2 = 0$: its value is $(i\nu_1)^2\, f_{\nu_1}\, (i\nu_2)\, f_{\nu_2}\, g(\omega \cdot \nu_2)$. One can detach the root line from the node with mode label ν_1 and attach it to the node with mode label ν_2, and reverse the arrow of the other line so that it points toward the new root line. In this way we obtain a new tree (cf. Figure 5): the value of the new tree is $(i\nu_1)\, f_{\nu_1}\, (i\nu_2)^2\, f_{\nu_2}\, g(\omega \cdot \nu_1)$, where $g(\omega \cdot \nu_1) = g(-\omega \cdot \nu_2) = g(\omega \cdot \nu_2)$, so that the values of the two trees contain a common factor $(i\nu_1)\, f_{\nu_1}\, (i\nu_2)\, f_{\nu_2}\, g(\omega \cdot \nu_2)$ times an extra factor which is $(i\nu_1)$ for the first tree and $(i\nu_2)$ for the second tree. Hence the sum of the two values gives zero.

Figure 5: Trees to be considered together to prove that $[\varepsilon \partial f(\alpha)]_0^{(2)} = 0$.

The cancellation mechanism described above can be generalised to all orders. Given a tree θ one considers all trees which can be obtained by detaching the root line and attaching to the other nodes of the tree, and by reversing the arrows of the lines (when needed) to make them point toward the root line. Then one sums together the values of all the trees so obtained: such values contain a common factor times a factor $i\nu_v$, if v is the node which the root line exits (the only nontrivial part of the proof is to check that the combinatorial factors match each other: we refer to Gentile & Mastropietro (1996) for details). Hence the sum gives zero, as the sum of all the mode labels vanishes.

For instance for $k = 3$ the cancellation operates by considering the three trees in Figure 5: such trees can be considered to be obtained from each other by shifting the root line and consistently reversing the arrows of the lines.

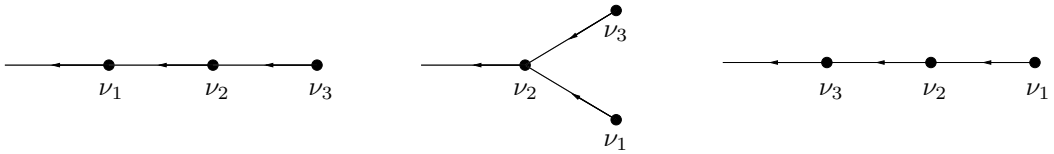

Figure 6: Trees to be considered together to prove that $[\varepsilon \partial f(\alpha)]_0^{(3)} = 0$.

In such a case the combinatorial factors of the node factors are different, because in the second tree the node factor associated with the node with mode label ν_2 contains a factor 1/2: on the other hand if $\nu_1 \neq \nu_3$ there are two nonequivalent trees with that shape (with the labels ν_1 and ν_3 exchanged between themselves), whereas if $\nu_1 = \nu_3$ there is only one such tree, but then the first and third trees are equivalent, so that only one of them must be counted. Thus, by using that $\nu_1 + \nu_2 + \nu_3 = 0$ – which implies $g(\omega \cdot (\nu_2 + \nu_3)) = g(-\omega \cdot \nu_1)$ and $g(\omega \cdot (\nu_1 + \nu_2)) = g(-\omega \cdot \nu_3)$, – in all cases we find that the sum of the values of the trees gives a common factor $(i\nu_1) f_{\nu_1} (i\nu_2)^2 f_{\nu_2} (i\nu_3) f_{\nu_3} g(\omega \cdot \nu_3) g(\omega \cdot \nu_1)$ times a factor 1 or 1/2 times $i(\nu_1 + \nu_2 + \nu_3)$, and hence vanishes: once more the property that g is even is crucial.

5 Small Divisors

We want to study the convergence properties of the series

$$h(\omega t) = \sum_{\nu \in \mathbb{Z}^d} \mathrm{e}^{i\omega \cdot \nu t} h_\nu, \qquad h_\nu = \sum_{k=1}^{\infty} \varepsilon^k h_\nu^{(k)}, \tag{5.1}$$

which has been shown to be well-defined as a formal power series for the models considered in Section 3.

Recall that the number of unlabelled trees of order k is bounded by 2^{2k}. To sum over the labels we can confine ourselves to the mode labels, as the momenta are uniquely determined by the mode labels. If f is a trigonometric polynomial of degree N, that is $f_\nu = 0$ for all ν such that $|\nu| := |\nu_1| + \ldots + |\nu_d| > N$, we have that $h_\nu^{(k)} = 0$ for all $|\nu| > kN$ (which can be easily proved by induction), and moreover we can bound the sum over the mode labels of any tree of order k by $(2N+1)^{dk}$. Finally we can bound

$$\prod_{v \in N(\theta)} |\nu_v|^{s_{v+1}} \leq \prod_{v \in N(\theta)} N^{s_{v+1}} \leq N^{2k}, \tag{5.2}$$

because of (4.2).

For the model (3.7), where $g_\ell = 1$ in (4.4), we can bound

$$\left| h_\nu^{(k)} \right| \leq \sum_{\nu \in \mathbb{Z}^d} \left| h_\nu^{(k)} \right| \leq 2^{2k} (2N+1)^{dk} N^{2k} \Phi^k, \qquad \Phi = \max_{|\nu| \leq N} |f_\nu|, \tag{5.3}$$

which shows that the series (5.1) converges for ε small enough, more precisely for $|\varepsilon| < \varepsilon_0$, with

$$\varepsilon_0 := C_0 (4N^2 \Phi (2N+1)^d)^{-1}, \tag{5.4}$$

where $C_0 = 1$. Hence the function $h(\omega t)$ in that case is analytic in ε. For $d = 1$ and $f(\alpha) = \cos \alpha$, we can easily provide an exact expression for the coefficients $h_\nu^{(k)}$: all the computational difficulties reduce to a combinatorial check, which can be found in Gentile & van Erp (2005), and the formula (3.9) is recovered.

However for the models where $g_\ell \neq 1$, the situation is much more involved: the propagators can be arbitrarily close to zero for ν large enough. This is the so-called *small divisor problem*. The series (5.1) is formally well-defined, assuming only an irrationality condition on ω. But to prove the convergence of the series, we need a stronger condition. For instance one can require the *standard Diophantine condition*

$$|\omega \cdot \nu| > \frac{\gamma}{|\nu|^\tau} \qquad \forall \nu \neq 0, \tag{5.5}$$

for suitable positive constants γ and τ. For fixed $\tau > d - 1$, the sets of vectors which satisfy (5.5) for some constant $\gamma > 0$ has full Lebesgue measure in $\mathbb{R}^d$ (Gallavotti, 1983). We can also impose a weaker condition, known as the *Bryuno condition*, which can be expressed by requiring

$$\mathcal{B}(\omega) := \sum_{k=0}^{\infty} \frac{1}{2^k} \log \frac{1}{\min_{0 < |\nu| \leq 2^k} |\omega \cdot \nu|} < \infty. \tag{5.6}$$

We call *Diophantine vectors* and *Bryuno vectors* the vectors satisfying (5.5) and (5.6), respectively. Note that any Diophantine vector satisfies (5.6).

If we assume only analyticity on f – that is, we remove the assumption that f be a trigonometric polynomial, – then we need a Diophantine condition, such as (5.5) or (5.6), also to show that the series (5.1) are well-defined as formal power series: the condition is needed, as one can easily check, in order to sum over the mode labels.

In the case of the standard map $|\omega \cdot \nu|$ must be replaced with $\min_{p\in\mathbb{Z}} |\omega\nu - p|$, and the function $\mathcal{B}(\omega)$ can be expressed in terms of the *best approximants* of the number ω (Davie, 1994).

For the models with small divisors considered in Section 3 convergence of the series (5.1) can be proved for ω satisfying (5.5) or (5.6). But this requires more work, and, in particular, a detailed discussion of the product of propagators in (4.6). In the following we shall consider the case of Diophantine vectors, by referring to the bibliography for the Bryuno vectors (cf. Section 8.3).

6 Multiscale Analysis

Consider explicitly the case $g(\omega \cdot \nu) = (\omega \cdot \nu)^{-2}$ in (4.1) and ω satisfying (5.5). We can introduce a label characterising the size of the propagator: we say that $\nu \in \mathbb{Z}^d$ is on *scale*

$$\begin{cases} n \geq 1, & \text{if } 2^{-n}\gamma \leq |\omega \cdot \nu| < 2^{-(n-1)}\gamma, \\ n = 0, & \text{if } \gamma \leq |\omega \cdot \nu|, \end{cases} \tag{6.1}$$

where γ is the constant appearing in (5.5), and we say that a line ℓ has a scale label $n_\ell = n$ if ν_ℓ is on scale n. If $\mathfrak{N}_n(\theta)$ denotes the number of lines $\ell \in L(\theta)$ with scale $n_\ell = n$, then we can bound in (4.5)

$$\Big| \prod_{\ell\in L(\theta)} g_\ell \Big| \leq \gamma^{-2k} \prod_{n=0}^{\infty} 2^{2n\mathfrak{N}_n(\theta)}, \tag{6.2}$$

so that the problem is reduced to bounding $\mathfrak{N}_n(\theta)$.

The product of propagators gives problems when the small divisors "accumulate". To make more precise the idea of accumulation we introduce the notion of cluster. Once all lines of a tree θ have been given their scale labels, for any $n \geq 0$ we can identify the maximal connected sets of lines with scale not larger than n. If at least one among such lines has scale equal to n we say that the set is a *cluster* on scale n. For instance consider the tree in Figure 1, and assign the mode labels to the nodes: this uniquely fixes the momenta, and hence the scale labels, of the lines. Suppose that we have found the scales as in Figure 7 (a). Then a cluster decomposition as in Figure 7 (b) follows. Given a cluster T call $L(T)$ the set of lines of θ contained in T, and denote by $N(T)$ the set of nodes connected by such lines. We define $k_T = |N(T)|$ the order of the cluster T.

Any cluster has either one or no exiting line, and can have an arbitrary number of entering lines. We call *self-energy clusters* the clusters which have one exiting line and only one entering line and are such that both lines have the same momentum. This means that if T is a self-energy cluster and ℓ_1 and ℓ_2 are the lines entering and exiting T, respectively, then $\nu_{\ell_1} = \nu_{\ell_2}$, so that

$$\sum_{v\in N(T)} \nu_v = 0. \tag{6.3}$$

By definition of scale, both lines ℓ_1 and ℓ_2 have the same scale, say n, and that, by definition of cluster, one has $n_\ell < n$ for all $\ell \in L(T)$.

We define the *value* of the self-energy cluster T whose entering line has momentum ν as the matrix

$$\mathcal{V}_T(\omega \cdot \nu) := \Big(\prod_{v\in N(T)} F_v \Big) \Big(\prod_{\ell\in L(T)} g_\ell \Big), \tag{6.4}$$

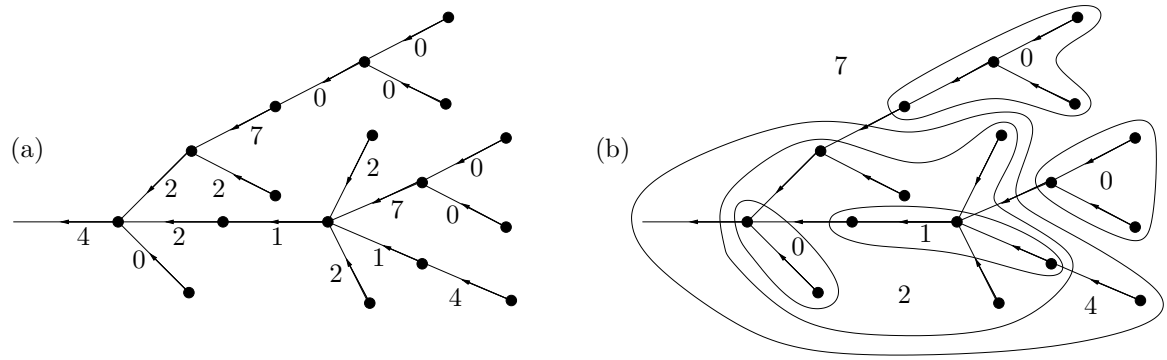

Figure 7: Example of clusters with their scales.

where all the indices of the node factors must be contracted except those associated with the line ℓ_1 entering T and with the line ℓ_2 exiting T.

We can extend the notion of self-energy clusters also to a single node, by saying that v is a self-energy cluster if $s_v = 1$ and the line entering v has the same momentum as the exiting line. In that case (6.4) has to be interpreted as $\mathcal{V}_T(\omega \cdot \nu) = F_v$: in particular it is independent of $\omega \cdot \nu$.

The simplest self-energy cluster one can think of consists of only one node v, but then (6.3) implies $\nu_v = 0$, so that $F_v = 0$, see (4.4), hence the corresponding value is zero. Thus the simplest non-trivial self-energy clusters contain at least two nodes, and are represented by the clusters T_1, T_2 and T_3 of Figure 8. In all cases one has $\nu_1 + \nu_2 = 0$. By using the definition (6.4) one has $\mathcal{V}_{T_1}(x) = (i\nu_1)^2 f_{\nu_1}(i\nu_2)^2 f_{\nu_2} g(\omega \cdot \nu_2 + x)$ and $\mathcal{V}_{T_2}(x) = \mathcal{V}_{T_3}(x) = (i\nu_1)^3 f_{\nu_1}(i\nu_2) f_{\nu_2} g(\omega \cdot \nu_2)/2$, with $x = \omega \cdot \nu$. Hence for $x = 0$ the sum of the three values gives zero. It is not difficult to see that also the first derivatives of the values of all self-energy clusters of order 2 sum up to zero, that is the sum of the values of all possible self-energy clusters of order $k = 2$ gives zero up to order x^2. (The only *caveat* is that, if we want to derive $\mathcal{V}_T(x)$, the sharp multiscale decomposition in (6.1) can be a little annoying, hence it can be more convenient to replace it with a smooth decomposition through C^∞ compact support functions; we refer to the literature for details). This property generalises to all orders k, and the underlying cancellation mechanism is essentially the same that ensures the validity of the second relation in (4.1).

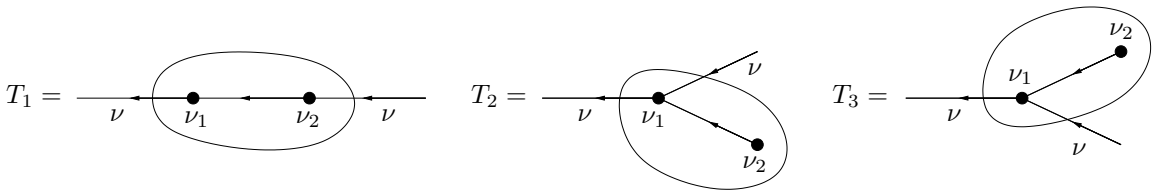

Figure 8: Self-energy clusters of order $k = 2$.

The reason why it is important to introduce the self-energy clusters is that if we could neglect them then the product of small divisors would be controlled. Indeed, let us denote by $\mathfrak{R}_n(\theta)$ the number of lines on scale n which do exit a self-energy cluster, and set $\mathfrak{N}^*_n(\theta) = \mathfrak{N}_n(\theta) - \mathfrak{R}_n(\theta)$. Then an important result, known as the *Siegel-Bryuno lemma*, is that

$$\mathfrak{N}^*_n(\theta) \le c\, 2^{-n/\tau} k, \tag{6.5}$$

for some constant c, where k is the order of θ and τ is the Diophantine exponent in (5.5).

The bound (6.5) follows from the fact that if $\mathfrak{N}^*_n(\theta) \neq 0$ then $\mathfrak{N}^*_n(\theta) \le E(n,k) := 2Nk2^{-(n-2)/\tau} - 1$, which can be proved by induction on k as follows. Given a tree θ let ℓ_0 be its root line, let $\ell_1, \dots, \ell_m$,

$m \geq 0$, be the lines on scales $\geq n$ which are the closest to ℓ_0, and let $\theta_1, \dots, \theta_m$ the trees with root lines $\ell_1, \dots, \ell_m$, respectively (cf. Figure 9 – note that by construction all lines ℓ in the subgraph T have scales $n_\ell < n$, so that if $n_{\ell_0} \geq n$ then T is necessarily a cluster). If either ℓ_0 is not on scale n or it is on scale n but exits a self-energy cluster then $\mathfrak{N}_n^*(\theta) = \mathfrak{N}_n^*(\theta_1) + \ldots + \mathfrak{N}_n^*(\theta_m)$ and the bound $\mathfrak{N}_n^*(\theta) \leq E(n,k)$ follows by the inductive hypothesis. If ℓ_0 does not exit a self-energy cluster and $n_{\ell_0} = n$ then $\mathfrak{N}_n^*(\theta) = 1 + \mathfrak{N}_n^*(\theta_1) + \ldots + \mathfrak{N}_n^*(\theta_m)$, and the lines $\ell_1, \dots, \ell_m$ enter a cluster T with $k_T = k - (k_1 + \ldots + k_m)$, where $k_1, \dots, k_m$ are the orders of $\theta_1, \dots, \theta_m$, respectively. If $m \geq 2$ the bound $\mathfrak{N}_n^*(\theta) \leq E(n,k)$ follows once more by the inductive hypothesis. If $m = 0$ then $\mathfrak{N}_n^*(\theta) = 1$; on the other hand for ℓ_0 to be on scale $n_{\ell_0} = n$ one must have $|\omega \cdot \nu_{\ell_0}| < 2^{-n+1}\gamma$ (see (6.1)), which, by the Diophantine condition (5.5), implies $Nk \geq |\nu_{\ell_0}| > 2^{(n-1)/\tau}$, hence $E(n,k) > 1$. If $m = 1$ call ν_1 and ν_2 the momenta of the lines ℓ_0 and ℓ_1, respectively. By construction T cannot be a self-energy cluster, hence $\nu_1 \neq \nu_2$, so that, by the Diophantine condition (5.5),

$$2^{-n+2}\gamma \geq |\omega \cdot \nu_1| + |\omega \cdot \nu_2| \geq |\omega \cdot (\nu_1 - \nu_2)| > \frac{\gamma}{|\nu_1 - \nu_2|^\tau}, \tag{6.6}$$

because $n_{\ell_0} = n$ and $n_{\ell_1} \geq n$. Thus, one has

$$Nk_T \geq \sum_{v \in N(T)} |\nu_v| \geq |\nu_1 - \nu_2| > 2^{(n-2)/\tau}, \tag{6.7}$$

hence T must contain "many nodes". In particular, one finds also in this case $\mathfrak{N}_n^*(\theta) = 1 + \mathfrak{N}_n^*(\theta_1) \leq 1 + E(n,k_1) \leq 1 + E(n,k) - E(n,k_T) \leq E(n,k)$, where we have used that $E(n,k_T) \geq 1$ by (6.7).

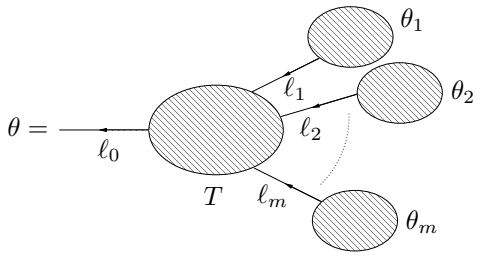

Figure 9: Construction for the proof of the Siegel-Bryuno lemma.

The argument above shows that small divisors can accumulate only by allowing self-energy clusters. That accumulation really occurs is shown by the example in Figure 10, where a tree θ of order k containing a chain of p self-energy clusters is depicted. Assume for simplicity that $k/3$ is an integer: then if $p = k/3$ the subtree θ_1 with root line ℓ is of order $k/3$. If the line ℓ entering the rightmost self-energy cluster T_p has momentum ν, also the lines exiting the p self-energy clusters have the same momentum ν. Suppose that $|\nu| \approx Nk/3$ and $|\omega \cdot \nu| \approx \gamma/|\nu|^\tau$ (this is certainly possible for some ν). Then the value of the tree θ grows like $a_1^k (k!)^{a_2}$, for some constants a_1 and a_2: a bound of this kind prevents the convergence of the perturbation series (5.1).

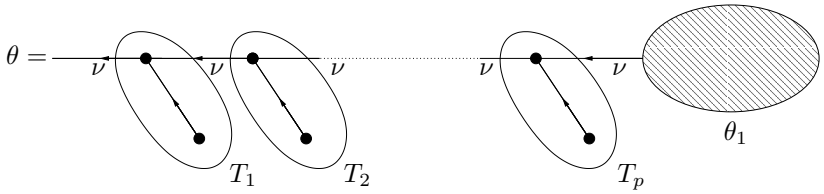

Figure 10: Example of accumulation of small divisors because of the self-energy clusters.

If no self-energy clusters could occur (so that $\mathfrak{R}_n(\theta) = 0$) the Siegel-Bryuno lemma would allow us to bound in (6.2)

$$\prod_{n=0}^{\infty} 2^{2n\mathfrak{N}_n(\theta)} = \prod_{n=0}^{\infty} 2^{2n\mathfrak{N}^*_n(\theta)} \leq \exp\left(C_1 k \sum_{n=0}^{\infty} n 2^{-n/\tau}\right) \leq C_2^k, \tag{6.8}$$

for suitable constants C_1 and C_2. In that case convergence of the series for $|\varepsilon| < \varepsilon_0$ would follow, with ε_0 defined as in (5.2) with $C_0 = \gamma^2/C_2$. However, there are self-energy clusters and they produce factorials, as the example in Figure 10 shows, so that we have to deal with them.

7 Resummation

Let us come back to the equation (3.6). If we expand $g(\omega \cdot \nu)[\varepsilon \partial_\alpha f(\alpha)]^{(k)}_\nu$ in trees according to the diagrammatic rules described in Section 4, we can distinguish between contributions in which the root line exits a self-energy cluster T, that we can write as

$$g(\omega \cdot \nu) \sum_{T : k_T < k} \mathcal{V}_T(\omega \cdot \nu)\, h^{(k-k_T)}_\nu, \tag{7.1}$$

and all the other contributions, that we denote by $[\varepsilon \partial_\alpha f(\alpha)]^{(k)*}_\nu$. We can shift the contributions (7.1) to the left hand side of (3.6) and divide by $g(\omega \cdot \nu)$, so to obtain

$$D(\omega \cdot \nu)\, h^{(k)}_\nu - \sum_{T : k_T < k} \mathcal{V}_T(\omega \cdot \nu)\, h^{(k-k_T)}_\nu = [\varepsilon \partial_\alpha f(\omega t + h)]^{(k)*}_\nu. \tag{7.2}$$

where $D(\omega \cdot \nu) = 1/g(\omega \cdot \nu) = 1/(\omega \cdot \nu)^2$. By summing over k and setting

$$M(\omega \cdot \nu; \varepsilon) = \sum_{k=1}^{\infty} \varepsilon^k \sum_{T : k_T = k} \mathcal{V}_T(\omega \cdot \nu), \tag{7.3}$$

then (7.2) gives

$$\mathcal{D}(\omega \cdot \nu)\, h_\nu = [\varepsilon \partial_\alpha f(\omega t + h)]^*_\nu, \qquad \mathcal{D}(\omega \cdot \nu) := D(\omega \cdot \nu) - M(\omega \cdot \nu; \varepsilon). \tag{7.4}$$

The motivation for proceeding in this way is that, at the price of changing the operator D into $\mathcal{D}$, hence of changing the propagators, lines exiting self-energy clusters no longer appear. Therefore, in the tree expansion of the right hand side of the equation, we have eliminated the self-energy clusters, that is the source of the problem of accumulation of small divisors.

Unfortunately the procedure described above has a problem: $M(\omega \cdot \nu; \varepsilon)$ itself is a sum of self-energy clusters, which can still contain some other self-energy clusters on lower scales. So finding a good bound for $M(\omega \cdot \nu; \varepsilon)$ could have the same problems as for the values of the trees.

To deal with such a difficulty we modify the prescription by proceeding recursively, in the following sense. Let us start from the momenta ν which are on scale $n = 0$. Since there are no self-energy clusters with exiting line on scale $n = 0$, for such ν one has $M(\omega \cdot \nu; \varepsilon) = 0$. Next, we consider the momenta ν which are on scale $n = 1$, and we can write (7.4), where now all self-energy clusters T whose values contribute to $M(\omega \cdot \nu; \varepsilon)$ cannot contain any self-energy clusters, because the lines $\ell \in T$ are on scale $n_\ell = 0$. Then, we consider the momenta ν which are on scale $n = 2$: again all the self-energy clusters contributing to $M(\omega \cdot \nu; \varepsilon)$ do not contain any self-energy clusters, because the lines on scale $n = 0, 1$ cannot exit self-energy clusters by the construction of the previous step, and so on. The conclusion is that we have obtained a different expansion for $h(\omega t)$, that we call a *resummed series*,

$$h(\omega t) = \sum_{\nu \in \mathbb{Z}^d} \mathrm{e}^{i\omega \cdot \nu t} h_\nu, \qquad h_\nu = \sum_{k=1}^{\infty} \varepsilon^k h^{[k]}_\nu(\varepsilon), \tag{7.5}$$

where the self-energy clusters do not appear any more in the tree expansion and the propagators must be defined recursively: the propagator g_ℓ of a line ℓ on scale $n_\ell = n$ and momentum $\nu_\ell = \nu$ is the matrix

$$g_\ell := g^{[n]}(\omega \cdot \nu; \varepsilon) = \Big(D(\omega \cdot \nu) - \mathcal{M}^{[n]}(\omega \cdot \nu; \varepsilon) \Big)^{-1}, \tag{7.6}$$

with

$$\mathcal{M}^{[n]}(\omega \cdot \nu; \varepsilon) := \sum_{T : n_T < n} \varepsilon^{k_T} \mathcal{V}_T(\omega \cdot \nu), \tag{7.7}$$

where the value $\mathcal{V}_T(\omega \cdot \nu)$ is written in accord with (6.4), with all the lines $\ell' \in L(T)$ on scales $n_{\ell'} < n$ and the corresponding propagators $g_{\ell'}$ expressed in terms of matrices $\mathcal{M}^{[n_{\ell'}]}(\omega \cdot \nu_{\ell'}; \varepsilon)$ as in (7.6). By construction, the new propagators depend on ε, so that the coefficients $h_\nu^{[k]}(\varepsilon)$ depend explicitly on ε: hence (7.5) is not a power series expansion.

The coefficients $h_\nu^{[k]}(\varepsilon)$ still admit a tree expansion

$$h_\nu^{[k]}(\varepsilon) = \sum_{\theta \in \mathcal{T}^{\mathcal{R}}_{k,\nu}} \mathrm{Val}(\theta), \qquad \mathrm{Val}(\theta) := \Big(\prod_{v \in N(\theta)} F_v \Big) \Big(\prod_{\ell \in L(\theta)} g^{[n_\ell]}(\omega \cdot \nu_\ell; \varepsilon) \Big) \qquad \nu \neq 0, \quad k \geq 1, \tag{7.8}$$

which replaces (4.6). In particular $\mathcal{T}^{\mathcal{R}}_{k,\nu}$ is defined as the set of *renormalised trees* of order k and momentum ν, where "renormalised" means that the trees do not contain any self-energy clusters. Now the propagators are matrices: the contractions of the indices in (7.8) yields that if ℓ connects v to v' the indices of F_v and $F_{v'}$ associated with ℓ must be equal to the colum and row indices of g_ℓ, respectively.

Since for any tree $\theta \in \mathcal{T}^{\mathcal{R}}_{k,\nu}$ one has $\mathfrak{N}_n(\theta) = \mathfrak{N}^*_n(\theta)$, we can bound the product of propagators according to (6.6) *provided the propagators on scale n can still be bounded proportionally to* 2^{2n}. This is certainly not obvious, because of the extra term $\mathcal{M}^{[n]}(\omega \cdot \nu; \varepsilon)$ appearing in (7.6).

It is a remarkable cancellation that $\mathcal{M}^{[n]}(x; \varepsilon)$ vanishes up to second order, that is $\mathcal{M}^{[n]}(x; \varepsilon) = O(x^2)$, so that, by taking into account also that $\mathcal{V}_T(x) \neq 0$ requires $k_T \geq 2$, we can write, for some constant C,

$$\mathcal{M}^{[n]}(x; \varepsilon) = \varepsilon^2 x^2 \overline{\mathcal{M}}^{[n]}(x; \varepsilon), \qquad \Big\| \overline{\mathcal{M}}^{[n]}(x; \varepsilon) \Big\| \leq C, \tag{7.9}$$

where $\| \cdot \|$ denotes – say – the uniform norm. The cancellation leading to (7.9) can be proved as discussed in Section 6 – where it has been explicitly proved for $k = 2$ in the absence of resummation. Now the propagators are matrices, but one can prove (by induction on n) that $\mathcal{M}^{[n]}(x; \varepsilon) = (\mathcal{M}^{[n]}(-x; \varepsilon))^T = (\mathcal{M}^{[n]}(x; \varepsilon))^\dagger$, with T and $\dagger$ denoting transposition and adjointness, respectively, and this is enough to see that the same cancellation mechanism applies (Gallavotti *et al.*, 2004).

Thus (7.9) implies that

$$\Big\| g^{[n]}(x; \varepsilon) \Big\| = \Big\| \Big(x^2 - \varepsilon^2 x^2 \overline{\mathcal{M}}^{[n]}(x; \varepsilon) \Big)^{-1} \Big\| \leq \frac{2}{x^2}, \tag{7.10}$$

and the same argument as used in Section 5 implies that the series in (7.5) for h_ν converges for $|\varepsilon| < \varepsilon_0$, where ε_0 is defined as in (5.2) with $C_0 = 2\gamma^2/C_2$. Moreover (7.9) also implies that $h(\omega t)$ is analytic in ε (notwithstanding that the expansion is not a power expansion), so that we can say *a posteriori* that the original power series (5.1) also converges.

8 Generalisations

The case where the function $G(u)$ in (2.1) does not vanish is notationally more involved, and it is explictly discussed in Gentile (2006a).

Extensions of the results described above to Hamiltonian functions more general than (3.1) require only some notational complications, and can be found in Gentile & Mastropietro (1996) for anisochronous systems and in Bartuccelli & Gentile (2002) for isochronous systems. The case (3.1) with f a real analytic function was first discussed by using trees in Chierchia & Falcolini (1994).

8.1 Lower-dimensional tori

A tree formalism for hyperbolic lower-dimensional tori was introduced and studied in Gallavotti (1994b), Gentile (1995) and Gallavotti, Gentile & Mastropietro (1999), for systems consisting of a set of rotators interacting with a pendulum. In that case also the stable and unstable manifolds (*whiskers*) of the tori were studied with diagrammatic techniques.

For Hamiltonian functions of the form (3.1) solutions of the form (7.5) describe quasi-periodic motions with frequency vector ω on a d-dimensional manifold. If we fix t, say $t = 0$, and keep α_0 as a parameter, we obtain a parameterisation of the manifold in terms of $\alpha_0 \in \mathbb{T}^d$, hence the manifold is an invariant torus. We say that the torus is a *maximal torus*.

We can study the problem of persistence of *lower-dimensional tori* by considering unperturbed solutions $\alpha(t) = \alpha_0 + \omega t$, where the components of ω are rationally dependent. For instance we can imagine that there exist s linearly independent vectors $\widehat{\nu}_1, \dots, \widehat{\nu}_s \in \mathbb{Z}^d$ such that $\omega \cdot \widehat{\nu}_k = 0$ for $k = 1, \dots, s$. In that case, we say that the unperturbed torus is a *resonant* torus of order s. We can imagine performing a linear change of variables which transforms the frequency vector into a new vector, that we still denote with ω, such that $\omega = (\bar{\omega}, 0)$, where $\bar{\omega} \in \mathbb{R}^r$ and $0 \in \mathbb{R}^s$, with $r+s = d$, and $\bar{\omega}$ is irrational. This naturally suggests that we write $\alpha = (\bar{\alpha}, \beta)$, with $\bar{\alpha} \in \mathbb{T}^r$ and $\beta \in \mathbb{T}^s$. More generally, here and henceforth in this subsection for any vector $v \in \mathbb{R}^d$ we denote by $\bar{v}$ the vector in $\mathbb{R}^r$ whose components are the first r components of v. For instance for the initial phase α_0 we write $\alpha_0 = (\bar{\alpha}_0, \beta_0)$.

In general the resonant torus is destroyed by the perturbation, and only some lower-dimensional tori persist under perturbation. We assume on $\bar{\omega}$ one of the Diophantine conditions of Section 5 in $\mathbb{R}^r$, for instance $|\bar{\omega} \cdot \bar{\nu}| > \gamma/|\bar{\nu}|^\tau$ for all $\bar{\nu} \in \mathbb{Z}^r$, $\bar{\nu} \neq 0$.

To prove the existence of a maximal torus for the system with Hamiltonian function (3.1) we needed no condition on the perturbation f. On the contrary to prove existence of lower-dimensional tori, we need some *non-degeneracy condition*: by defining

$$f_0(\beta) = \int_{\mathbb{T}^r} \frac{\mathrm{d}\bar{\alpha}}{(2\pi)^r} f(\bar{\alpha}, \beta), \tag{8.1}$$

if $\partial_\beta f_0(\beta_*) = 0$ for some β_* then we assume that the matrix $\partial^2_\beta f_0(\beta_*)$ is positive definite (more generally one could assume it to be non-singular, that is $\det \partial^2_\beta f_0(\beta_*) \neq 0$).

The formal analysis can be carried out as in Section 4, with the only difference being that now the compatibility conditions $[\varepsilon \partial_\alpha f(\alpha)]^{(k)}_\nu = 0$ have to be imposed for all ν such that $\bar{\nu} = 0$, because $\omega \cdot \nu = \bar{\omega} \cdot \bar{\nu}$. It turns out to be an identity only for the first r components (this is trivial for $k = 1$, whereas it requires some work for $k > 1$). For the last s components, for $k = 1$ it reads $\partial_\beta f_0(\beta_0) = 0$, hence it fixes β_0 to be a stationary point for $f_0(\beta)$, while for higher values of k it fixes the corrections of higher order of these values (to do this we need the non-degeneracy condition). Thus, we are free to choose only $\bar{\alpha}_0$ as a free parameter, since the last s components of α_0 have to be fixed.

Clusters and self-energy clusters are defined as in Section 6. Note that only the first r components $\bar{\nu}$ of the momenta ν intervene in the definition of the scales – again because $\omega \cdot \nu = \bar{\omega} \cdot \bar{\nu}$. In particular, in the definition of self-energy clusters, in (6.3) we must replace ν_v with $\bar{\nu}_v$. Thus, already to first order the value of a self-energy cluster can be non-zero: for $k_T = 1$, that is for T containing only a node v with mode label $(\bar{\nu}_v, \tilde{\nu}_v) = (0, \tilde{\nu}_v)$, the matrix $\mathcal{V}_T(x; \varepsilon)$ is of the form

$$\mathcal{V}_T(x) = \begin{pmatrix} 0 & 0 \\ 0 & b_{\tilde{\nu}_v} \end{pmatrix}, \qquad (b_{\tilde{\nu}_v})_{i,j} = \mathrm{e}^{i\tilde{\nu}_v \cdot \beta_0} (i\tilde{\nu}_{v,i})(i\tilde{\nu}_{v,j}) f_{(0,\tilde{\nu}_v)}, \tag{8.2}$$

with $i, j = r+1, \dots, d$. If we sum over $\tilde\nu_v \in \mathbb{Z}^s$ and multiply times ε, we obtain

$$M_0 := \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon B \end{pmatrix}, \qquad B_{i,j} = \sum_{\tilde\nu \in \mathbb{Z}^s} \mathrm{e}^{i\tilde\nu\cdot\beta_0} (i\tilde\nu_i)(i\tilde\nu_j) f_{(0,\tilde\nu)} = \partial_{\beta_i}\partial_{\beta_j} f_0(\beta_0). \tag{8.3}$$

The $s\times s$ block B is non-zero: in fact, the non-degeneracy condition yields that it is invertible.

To higher orders one finds that the matrix $\mathcal{M}^{[n]}(x;\varepsilon)$, with $x = \omega\cdot\nu = \bar\omega\cdot\bar\nu$, is a self-adjoint matrix and $\mathcal{M}^{[n]}(x;\varepsilon) = (\mathcal{M}^{[n]}(-x;\varepsilon))^T$, as in the case of maximal tori. Moreover the corresponding eigenvalues $\lambda_i^{[n]}(x;\varepsilon)$ satisfy $\lambda_i^{[n]}(x;\varepsilon) = O(\varepsilon^2 x^2)$ for $i = 1,\dots,r$ and $\lambda_i^{[n]}(x;\varepsilon) = O(\varepsilon)$ for $i = r+1,\dots,d$; this property is not trivial because of the off-diagonal blocks (which in general do not vanish at orders $k \ge 2$), and to prove it one has to use the self-adjointness of the matrix $\mathcal{M}^{[n]}(x;\varepsilon)$. More precisely one has $\lambda_i^{[n]}(x;\varepsilon) = \varepsilon a_i + O(\varepsilon^2)$ for $i > r$, where $a_{r+1},\dots,a_d$ are the s eigenvalues of the matrix B in (8.3). From this point on the discussion proceeds in a very different way according to the sign of ε (recall that we are assuming that $a_i > 0$ for all $i > r$).

For $\varepsilon < 0$ one has $\lambda_i^{[n]}(x;\varepsilon) = a_i\varepsilon + O(\varepsilon^2) < 0$ for $i > r$, so that we can bound the last s eigenvalues of $x^2 - \mathcal{M}^{[n]}(x;\varepsilon)$ with x^2, and the first r with $x^2/2$ by the same argument as in Section 7. Hence we obtain easily the convergence of the series (7.5); of course, analyticity at the origin is prevented because of the condition $\varepsilon < 0$. We say in that case that the lower-dimensional tori are *hyperbolic*. We refer to Gallavotti & Gentile (2002) and Gallavotti *et al.* (2004) for details.

The case of *elliptic* lower-dimensional tori – that is $\varepsilon > 0$ when $a_i > 0$ for all $i > r$ – is more difficult. Essentially the idea is as follows (we only sketch the strategy: the details can be found in Gentile & Gallavotti (2005)). One has to define the scales recursively, by using a variant, first introduced in Gentile (2003), of the resummation technique described in Section 7. We say that ν is on scale 0 if $|\bar\omega\cdot\bar\nu| \ge \gamma$ and on scale $[\ge 1]$ otherwise: for ν on scale 0 we write (7.4) with $M = M_0$, as given in (8.3). This defines the propagators of the lines ℓ on scale $n = 0$ as

$$g_\ell = g^{[0]}(\omega\cdot\nu_\ell) = \left((\bar\omega\cdot\bar\nu_\ell)^2 - M_0\right)^{-1}. \tag{8.4}$$

Denote by λ_i the eigenvalues of M_0: given ν on scale $[\ge 1]$ we say that ν is on scale 1 if $2^{-1}\gamma \le \min_{i=1,\dots,d}\sqrt{|(\bar\omega\cdot\bar\nu)^2 - \lambda_i|}$, and on scale $[\ge 2]$ if $\min_{i=1,\dots,d}\sqrt{|(\bar\omega\cdot\bar\nu)^2 - \lambda_i|} < 2^{-1}\gamma$. For ν on scale 1 we write (7.4) with M replaced by $\mathcal{M}^{[0]}(\omega\cdot\bar\nu;\varepsilon)$, which is given by M_0 plus the sum of the values of all self-energy clusters T on scale $n_T = 0$. Then the propagators of the lines ℓ on scale $n_\ell = 1$ is defined as

$$g_\ell = g^{[1]}(\omega\cdot\nu_\ell) = \left((\bar\omega\cdot\bar\nu_\ell)^2 - \mathcal{M}^{[0]}(\bar\omega\cdot\bar\nu_\ell;\varepsilon)\right)^{-1}. \tag{8.5}$$

Call $\lambda_i^{[n]}(x;\varepsilon)$ the eigenvalues of $\mathcal{M}^{[n]}(x;\varepsilon)$: given ν on scale $[\ge 2]$ we say that ν is on scale 2 if $2^{-2}\gamma \le \min_{i=1,\dots,d}\sqrt{|(\bar\omega\cdot\bar\nu)^2 - \lambda_i^{[0]}(\bar\omega\cdot\bar\nu;\varepsilon)|}$, and on scale $[\ge 3]$ if $\min_{i=1,\dots,d}\sqrt{|(\bar\omega\cdot\bar\nu)^2 - \lambda_i^{[0]}(\bar\omega\cdot\bar\nu;\varepsilon)|} < 2^{-2}\gamma$. For ν on scale 2 we write (7.4) with M replaced by $\mathcal{M}^{[1]}(\bar\omega\cdot\bar\nu;\varepsilon)$, which is given by $\mathcal{M}^{[0]}(\bar\omega\cdot\bar\nu;\varepsilon)$ plus the sum of the values of all self-energy clusters T on scale $n_T = 1$. Thus, the propagators of the lines ℓ on scale $n_\ell = 2$ will be defined as

$$g_\ell = g^{[2]}(\omega\cdot\nu_\ell) = \left((\bar\omega\cdot\bar\nu_\ell)^2 - \mathcal{M}^{[1]}(\bar\omega\cdot\bar\nu_\ell;\varepsilon)\right)^{-1}, \tag{8.6}$$

and so on. The propagators are self-adjoint matrices, hence their norms can be bounded through the corresponding eigenvalues. In order to proceed as in Sections 6 and 7 we need some Diophantine conditions on these eigenvalues. We can assume for some $\tau' > \tau$

$$\left|\,|\bar\omega\cdot\bar\nu| - \sqrt{|\lambda_i^{[n]}(\bar\omega\cdot\bar\nu;\varepsilon)|}\,\right| > \frac{\gamma}{|\bar\nu|^{\tau'}} \qquad \forall \bar\nu \ne 0, \tag{8.7}$$

for all $i = 1, \ldots, d$ and $n \geq 0$. These are known as the *first Melnikov conditions*.

Unfortunately, things do not proceed so plainly. In order to prove a bound like (6.5), possibly with a different τ' replacing τ, we need to compare the propagators of the lines entering and exiting clusters T which are not self-energy clusters. This requires replacing (6.6) with

$$2^{-n+2}\gamma \geq \left| \bar{\omega} \cdot (\bar{\nu}_1 - \bar{\nu}_2) \pm \sqrt{|\lambda_i^{[n]}(\bar{\omega} \cdot \bar{\nu}_1; \varepsilon)|} \pm \sqrt{|\lambda_j^{[n]}(\bar{\omega} \cdot \bar{\nu}_2; \varepsilon)|} \right| > \frac{\gamma}{|\bar{\nu}_1 - \bar{\nu}_2|^{\tau'}}, \tag{8.8}$$

for all $i, j = 1, \ldots, d$ and choices of the signs $\pm$, and hence introduces further Diophantine conditions, known as the *second Melnikov conditions*.

The conditions in (8.8) turn out to be too many, because for all $n \geq 0$ and all $\bar{\nu} \in \mathbb{Z}^r$ such that $\bar{\nu} = \bar{\nu}_1 - \bar{\nu}_2$ there are infinitely many conditions to be considered, one per pair $(\bar{\nu}_1, \bar{\nu}_2)$. However we can impose both the conditions (8.7) and (8.8) not for the eigenvalues $\lambda_i^{[n]}(\bar{\omega} \cdot \bar{\nu}; \varepsilon)$, but for some quantities $\underline{\lambda}_i^{[n]}(\varepsilon)$ independent of $\bar{\nu}$ and then use the smoothness of the eigenvalues in x to control $(\bar{\omega} \cdot \bar{\nu})^2 - \lambda_i^{[n]}(\bar{\omega} \cdot \bar{\nu}; \varepsilon)$ in terms of $(\bar{\omega} \cdot \bar{\nu})^2 - \underline{\lambda}_i^{[n]}(\varepsilon)$. Eventually, beside the Diophantine condition on $\bar{\omega}$, we have to impose the Melnikov conditions

$$\left| |\bar{\omega} \cdot \bar{\nu}| - \sqrt{|\underline{\lambda}_i^{[n]}(\varepsilon)|} \right| > \frac{\gamma}{|\bar{\nu}|^{\tau'}}, \qquad \left| |\bar{\omega} \cdot \bar{\nu}| \pm \sqrt{|\underline{\lambda}_i^{[n]}(\varepsilon)|} \pm \sqrt{|\underline{\lambda}_j^{[n]}(\varepsilon)|} \right| > \frac{\gamma}{|\bar{\nu}|^{\tau'}}, \tag{8.9}$$

for all $\bar{\nu} \neq 0$ and all $n \geq 0$. Each condition in (8.9) leads us to eliminate a small interval of values of ε. For the values of ε which are left define $h(\omega t)$ according to (7.5) and (7.8), with the new definition of the propagators. If ε is small enough, say $|\varepsilon| < \varepsilon_0$, then the series (7.5) converges. Denote by $\mathfrak{E} \subset [0, \varepsilon_0]$ the set of values of ε for which the conditions (8.9) are satisfied. One can prove that $\mathfrak{E}$ is a *Cantor set*, that is a perfect, nowhere dense set. Moreover $\mathfrak{E}$ has large relative Lebesgue measure, in the sense that

$$\lim_{\varepsilon \to 0} \frac{\text{meas}(\mathfrak{E} \cap [0, \varepsilon])}{\varepsilon} = 1, \tag{8.10}$$

provided τ' in (8.9) is large enough with respect to τ. The property (8.10) yields that, notwithstanding that we are eliminating infinitely many intervals, the measure of the union of all these intervals is small.

If $a_i < 0$ for all $i > r$ we reason in the same way, simply exchanging the role of positive and negative ε. On the contrary if $a_i = 0$ for some $i > r$, the problem becomes much more difficult. For instance if $s = 1$ and $a_{r+1} = 0$, then in general perturbation theory in ε is not possible, not even at a formal level. However, under some conditions, one can still construct fractional series in ε, and prove that the series can be resummed (Gallavotti, Gentile & Giuliani, 2006).

8.2 Other ordinary differential equations

The formalism described above extends to other models, such as skew-product systems (Gentile, 2006b) and systems with strong damping in the presence of a quasi-periodic forcing term (Gentile, Bartuccelli & Deane, 2005).

As an example of skew-product system one can consider the linear differential equation $\dot{x} = (\lambda A + \varepsilon f(\omega t))\, x$ on $\mathrm{SL}(2, \mathbb{R})$, where $\lambda \in \mathbb{R}$, ε is a small real parameter, $\omega \in \mathbb{R}^n$ is an irrational vector, and $A, f \in \mathfrak{sl}(2, \mathbb{R})$, with A is a constant matrix and f an analytic function periodic in its arguments. Trees for skew-products were considered by Iserles and Nørsett (1999), but they used expansions in time, hence not suited for the study of global properties, such as quasi-periodicity.

Quasi-periodically forced one-dimensional systems with strong damping are described by the ordinary differential equations $\ddot{x} + \gamma \dot{x} + g(x) = f(\omega t)$, where $x \in \mathbb{R}$, $\varepsilon = 1/\gamma$ is a small real parameter, $\omega \in \mathbb{R}^n$ is irrational, and f, g are analytic functions (g is the "force"), with f periodic in its arguments.

We refer to the bibliography for details and results on the existence of quasi-periodic solutions.

8.3 Bryuno vectors

The diagrammatic methods can be used to prove the any unperturbed maximal torus with frequency vector which is a Bryuno vector persists under perturbation for ε small enough (Gentile, 2007). One could speculate whether the Bryuno condition (5.6) is optimal. In general the problem is open. However, in the case of the standard map – see Section 3.3, – one can prove (Davie, 1994; Berretti & Gentile, 2001b) that, by considering the radius of convergence ε_0 of the perturbation series as a function of ω, say $\varepsilon_0 = \rho_0(\omega)$, then there exists a universal constant C such that

$$|\log \rho_0(\omega) + 2\mathcal{B}(\omega)| \le C. \tag{8.11}$$

In particular this yields that the invariant curve with rotation number ω persists under perturbation if and only if ω is a Bryuno number. The proof of (8.11) requires the study of a more refined cancellation than that discussed in Section 7. We refer to Berretti & Gentile (2001b) and Gentile (2006a) for details.

Extensions to Bryuno vectors for lower-dimensional tori can also be found in Gentile (2007). For the models considered in Section 8.2 we refer to Gentile (2006b) and Gentile, Bartuccelli & Deane (2006).

8.4 Partial differential equations

Existence of quasi-periodic solutions in systems described by one-dimensional nonlinear partial differential equations (finite-dimensional tori in infinite-dimensional systems) was first studied by Kuksin (1993), Craig and Wayne (1993) and Bourgain (1998). In these systems, even the case of periodic solutions yields small divisors, and hence requires a multiscale analysis. The study of persistence of periodic solutions for nonlinear Schrödinger equations and nonlinear wave equations, with the techniques discussed here, can be found in Gentile & Mastropietro (2004), Gentile, Mastropietro & Procesi (2005), and Gentile & Procesi (2006).

The models are still described by (2.1), with $G(u) = 0$, but now D is given by $D = \partial_t^2 - \Delta + \mu$ in the case of the wave equation and by $D = i\partial_t - \Delta + \mu$ in the case of the Schrödinger equation, where Δ is the Laplacian and $\mu \in \mathbb{R}$. In dimension 1, one has $\Delta = \partial_x^2$. If we look for periodic solutions with frequency ω it can be convenient to pass to Fourier space, where the operator D acts as

$$D : \mathrm{e}^{i\omega nt + imx} \to \left(-\omega^2 n^2 + m^2 + \mu\right) \mathrm{e}^{i\omega nt + imx}, \tag{8.12}$$

for the wave equation; a similar expression holds for the Schrödinger equation. Therefore the kernel of D can be arbitrarily close to zero for n and m large enough.

Then one can consider, say, (2.1) for $x \in [0, \pi]$ and $F(u) = u^3$, with Dirichlet boundary conditions $u(0) = u(\pi) = 0$, and study the existence of periodic solutions with frequency ω close to some of the unperturbed frequencies. We refer to the cited bibliography for results and proofs.

9 Conclusions, and Future Directions

The diagrammatic techniques described above have been applied also in cases where no small divisors appear; cf. Berretti & Gentile (1999) and Gentile, Bartuccelli & Deane (2007). Of course, such problems are much easier from a technical point of view, and can be considered as propaedeutic examples to become familiar with the tree formalism. Also the study of lower-dimensional tori becomes easy for $r = 1$ (periodic solutions): in that case one has $|\bar\omega \cdot \bar\nu| \ge |\bar\omega|$ for all $\bar\nu \ne 0$, so that the product of the propagators is bounded by $|\bar\omega|^{-2k}$, and one can proceed as in Section 5 to obtain analyticity of the solutions.

In the case of hyperbolic lower-dimensional tori, if ω is a two-dimensional Diophantine vector of constant type (that is, with $\tau = 1$) the conjugation function h can be proved to be Borel summable (Costin *et al.*, 2007). Analogous considerations hold for the one-dimensional systems in the presence of

friction and of a quasiperiodic forcing term described in Section 8.2; in that case one has Borel summability also for one-dimensional ω, that is for periodic forcing (Gentile, Bartuccelli & Deane, 2006). It would be interesting to investigate whether Borel summability could be obtained for higher values of τ.

Recently existence of finite-dimensional tori in the nonlinear Schrödinger equation in higher dimensions was proved by Bourgain (2005). It would be interesting to investigate how far the diagrammatic techniques extend to deal with such higher dimensional generalisations. The main problem is that (the analogues of) the second Melnikov conditions in (8.9) cannot be imposed.

In certain cases the tree formalism was extended to non-analytic systems, such as some quasi-integrable systems of the form (3.1) with f in a class of C^p functions for some finite p (Bonetto *et al.*, 1998a and 1998b). However, up to exceptional cases, the method described here seems to be intrinsically suited in cases in which the vector fields are analytic. The reason is that in order to exploit the expansion (2.3), we need that F be infinitely many times differentiable and we need a bound on the derivatives. It is a remarkable property that the perturbation series can be given a meaning also in cases where the solutions are not analytic in ε.

An advantage of the diagrammatic method is that it allows rather detailed information about the solutions, hence it could be more convenient than other techniques to study problems where the underlying structure is not known or too poor to exploit general abstract arguments.

Another advantage is the following. If one is interested not only in proving the existence of the solutions, but also in explicitly constructing them with any prefixed precision, this requires performing analytical or numerical computations with arbitrarily high accuracy. Then high perturbation orders have to be reached, and the easiest and most direct way to proceed is just through perturbation theory: so the approach illustrated here allows a unified treatment for both theoretical investigations and computational ones.

The resummation technique described in Section 7 can also be used for computational purposes. With respect to the naive power series expansion it can reduce the computation time required to approximate the solution within a prefixed precision. It can also provide accurate information on the analyticity properties of the solution. For instance, for the Kepler equation, Levi-Civita at the beginning of the last century described a resummation rule (see Levi-Civita, 1954), which gives immediately the radius of convergence of the perturbation series. Of course, in the case of small divisor problems, everything becomes much more complicated.

10 Bibliography

10.1 Books and Reviews

1. V.I. Arnold, V.V. Kozlov, A.I. Neĭshtadt (1988) *Dynamical systems. III*, Encyclopaedia of Mathematical Sciences Vol. 3, Springer, Berlin.

2. A. Berretti, G. Gentile (2001a) *Renormalization group and field theoretic techniques for the analysis of the Lindstedt series*, Regul. Chaotic Dyn. **6**, no. 4, 389–420.

3. A. Celletti (2008) *Perturbation theory in celestial mechanics*, Encyclopedia of Complexity and System Science.

4. L. Chierchia (2008) *Kolmogorov-Arnold-Moser (KAM) theory*, Encyclopedia of Complexity and System Science.

5. G. Gallavotti (1983) *The elements of mechanics*, Texts and Monographs in Physics, Springer, New York.

6. G. Gallavotti (2001) *Renormalization group in statistical mechanics and mechanics: gauge symmetries and vanishing beta functions*, Renormalization group theory in the new millennium, III, Phys. Rep. **352**, no. 4-6, 251–272.

7. G. Gallavotti, F. Bonetto, G. Gentile (2004) *Aspects of ergodic, qualitative and statistical theory of motion*, Texts and Monographs in Physics, Springer, Berlin.

8. G. Gallavotti (2006) *Classical mechanics*, Encyclopedia of Mathematical Physics, Vol. 1, 1-32, Eds. J.-P. Françoise, G.L. Naber and T. Sh. Tsun, Elsevier, Oxford.

9. G. Gallavotti (2008) *Perturbation theory*, Encyclopedia of Complexity and System Science.

10. G. Gentile (1999) *Diagrammatic techniques in perturbations theory, and applications*, Symmetry and perturbation theory (Rome, 1998), 59–78, World Science, River Edge, NJ.

11. G. Gentile (2006a) *Brjuno numbers and dynamical systems*, Frontiers in number theory, physics, and geometry. I, 585–599, Springer, Berlin.

12. G. Gentile (2006b) *Diagrammatic techniques in perturbation theory*, Encyclopedia of Mathematical Physics, Vol. 2, 54-60, Eds. J.-P. Françoise, G.L. Naber and T. Sh. Tsun, Elsevier, Oxford.

13. G. Gentile, V. Mastropietro (2001) *Renormalization group for one-dimensional fermions. A review on mathematical results*, Renormalization group theory in the new millennium, III, Phys. Rep. **352**, no. 4-6, 273–437.

14. F. Harary (1969) *Graph theory*, Addison-Wesley, Reading.

10.2 Primary literature

15. V.I. Arnold (1963) *Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian* (Russian), Uspehi Mat. Nauk **18**, no. 5 (113), 13–40.

16. M.V. Bartuccelli, G. Gentile (2002) *Lindstedt series for perturbations of isochronous systems: a review of the general theory*, Rev. Math. Phys. **14**, no. 2, 121–171.

17. A. Berretti, G. Gentile (1999) *Scaling properties for the radius of convergence of a Lindstedt series: the standard map*, J. Math. Pures Appl. **78**, no. 2, 159–176.

18. A. Berretti, G. Gentile (2000) *Scaling properties for the radius of convergence of a Lindstedt series: generalized standard maps*, J. Math. Pures Appl. **79**, no. 7, 691-713.

19. A. Berretti, G. Gentile (2001b) *Bryuno function and the standard map*, Comm. Math. Phys. **220**, no. 3, 623–656.

20. F. Bonetto, G. Gallavotti, G. Gentile, V. Mastropietro (1998a) *Quasi linear flows on tori: regularity of their linearization*, Comm. Math. Phys. **192**, no. 3, 707-736.

21. F. Bonetto, G. Gallavotti, G. Gentile, V. Mastropietro (1998b) *Lindstedt series, ultraviolet divergences and Moser's theorem*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **26**, no. 3, 545-593.

22. J. Bourgain (1998) *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Ann. of Math. **148**, no. 2, 363-439.

23. J. Bourgain (2005) *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, NJ.

24. J. Bricmont, K. Gawędzki, A. Kupiainen, *KAM theorem and quantum field theory*, Comm. Math. Phys. **201** (1999), no. 3, 699-727.

25. L. Chierchia, C. Falcolini (1994) *A direct proof of a theorem by Kolmogorov in Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **21**, no. 4, 541–593.

26. O. Costin, G. Gallavotti, G. Gentile, A. Giuliani (2007) *Borel summability and Lindstedt series*, Comm. Math. Phys. **269**, no. 1, 175–193.

27. W. Craig, C.E. Wayne (1993) *Newton's method and periodic solutions of nonlinear wave equations*, Comm. Pure Appl. Math. **46**, no. 11, 1409–1498.

28. A.M. Davie (1994) *The critical function for the semistandard map*, Nonlinearity **7**, no. 1, 219–229.

29. L.H. Eliasson (1996) *Absolutely convergent series expansions for quasi periodic motions*, Math. Phys. Electron. J. **2**, Paper 4, 33 pp. (electronic).

30. G. Gallavotti (1994a) *Twistless KAM tori*, Comm. Math. Phys. **164**, no. 1, 145–156.

31. G. Gallavotti (1994b) *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review*, Rev. Math. Phys. **6**, no. 3, 343–411.

32. G. Gallavotti, G. Gentile (2002), *Hyperbolic low-dimensional invariant tori and summations of divergent series*, Comm. Math. Phys. **227**, no. 3, 421–460.

33. G. Gallavotti, G. Gentile, A. Giuliani (2006) *Fractional Lindstedt series*, J. Math. Phys. **47**, no. 1, 012702, 33 pp.

34. G. Gallavotti, G. Gentile, V. Mastropietro (1999) *A field theory approach to Lindstedt series for hyperbolic tori in three time scales problems*, J. Math. Phys. **40**, no. 12, 6430–647.

35. G. Gentile (1995) *Whiskered tori with prefixed frequencies and Lyapunov spectrum*, Dynam. Stability Systems **10**, no. 3, 269–308.

36. G. Gentile (2003) *Quasi-periodic solutions for two-level systems*, Comm. Math. Phys. **242**, no. 1-2, 221–250.

37. G. Gentile (2006c) *Resummation of perturbation series and reducibility for Bryuno skew-product flows*, J. Stat. Phys. **125**, no. 2, 321–361.

38. G. Gentile (2007) *Degenerate lower-dimensional tori under the Bryuno condition*, Ergodic Theory Dynam. Systems **27**, no. 2, 427–457.

39. G. Gentile, M.V. Bartuccelli, J.H.B. Deane (2005) *Summation of divergent series and Borel summability for strongly dissipative differential equations with periodic or quasiperiodic forcing terms*, J. Math. Phys. **46**, no. 6, 062704, 20 pp.

40. G. Gentile, M.V. Bartuccelli, J.H.B. Deane (2006) *Quasiperiodic attractors, Borel summability and the Bryuno condition for strongly dissipative systems*, J. Math. Phys. **47**, no. 7, 072702, 10 pp.

41. G. Gentile, M.V. Bartuccelli, J.H.B. Deane (2007) *Bifurcation curves of subharmonic solutions and Melnikov theory under degeneracies*, Rev. Math. Phys. **19**, no. 3, 307–348.

42. G. Gentile, D.A. Cortez, J.C.A. Barata (2005) *Stability for quasi-periodically perturbed Hill's equations*, Comm. Math. Phys. **260**, no. 2, 403-443.

43. G. Gentile, G. Gallavotti (2005), *Degenerate elliptic tori*, Comm. Math. Phys. **257**, no. 2, 319–362.

44. G. Gentile, V. Mastropietro (1996) *Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications*, Rev. Math. Phys. **8**, no. 3, 393–444.

45. G. Gentile, V. Mastropietro (2004) *Construction of periodic solutions of nonlinear wave equations with Dirichlet boundary conditions by the Lindstedt series method*, J. Math. Pures Appl. **83**, no. 8, 1019–1065.

46. G. Gentile, V. Mastropietro, M. Procesi (2005) *Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions*, Comm. Math. Phys. **256**, no. 2, 437–490.

47. G. Gentile, M. Procesi (2006) *Conservation of resonant periodic solutions for the one-dimensional nonlinear Schrödinger equation*, Comm. Math. Phys. **262**, no. 3, 533–553.

48. G. Gentile, T.S. van Erp (2005) *Breakdown of Lindstedt expansion for chaotic maps*, J. Math. Phys. **46**, no. 10, 102702, 20 pp.

49. A. Iserles, S.P. Nørsett (1999) *On the solution of linear differential equations in Lie groups*, R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357**, no. 1754, 983–1019.

50. K. Khanin, J. Lopes Dias, J. Marklof (2007) *Multidimensional continued fractions, dynamical renormalization and KAM theory*, Comm. Math. Phys. **270**, no. 1, 197–231.

51. H. Koch (1999) *A renormalization group for Hamiltonians, with applications to KAM tori*, Ergodic Theory Dynam. Systems **19**, no. 2, 475–521.

52. A.N. Kolmogorov (1954) *On conservation of conditionally periodic motions for a small change in Hamilton's function* (Russian), Dokl. Akad. Nauk SSSR **98**, 527–530.

53. S.B. Kuksin (1993) *Nearly integrable infinite-dimensional Hamiltonian systems*, Lecture Notes in Mathematics,

1556. Springer-Verlag, Berlin.

54. T. Levi-Civita (1954) *Opere matematiche. Memorie e note*, Zanichelli, Bologna.

55. R.S. MacKay (1993) *Renormalisation in area-preserving maps*, Advanced Series in Nonlinear Dynamics, 6. World Scientific Publishing, River Edge, NJ.

56. J. Moser (1962) *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1962**, 1–20.

57. J. Moser (1967) *Convergent series expansions for quasi-periodic motions*, Math. Ann. **169**, 136–176.

58. H. Poincaré (1892-1899) *Les méthodes nouvelles de la mécanique céleste*, Gauthier-Villars, Paris.

59. J. Pöschel (1982) *Integrability of Hamiltonian systems on Cantor sets*, Comm. Pure Appl. Math. **35**, no. 5, 653–696.

60. A. Wintner (1941) *The analytic foundations of celestial mechanics*, Princeton Universisty Press, Princeton.