# Quasi-periodic motions in dynamical systems. Review of a renormalisation group approach

**Guido Gentile**

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

**Abstract**

Power series expansions naturally arise whenever solutions of ordinary differential equations are studied in the regime of perturbation theory. In the case of quasi-periodic solutions the issue of convergence of the series is plagued of the so-called small divisor problem. In this paper we review a method recently introduced to deal with such a problem, based on renormalisation group ideas and multiscale techniques. Applications to both quasi-integrable Hamiltonian systems (KAM theory) and non-Hamiltonian dissipative systems are discussed. The method is also suited to situations in which the perturbation series diverges and a resummation procedure can be envisaged, leading to a solution which is not analytic in the perturbation parameter: we consider explicitly examples of solutions which are only $C^\infty$ in the perturbation parameter, or even defined on a Cantor set.

## 1 Introduction

Consider ordinary differential equations of the form

$$D_\varepsilon u = \varepsilon F(u, \underline{\omega} t), \tag{1.1}$$

where $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, $\varepsilon \in \mathbb{R}$ and $\underline{\omega} \in \mathbb{R}^m$ are parameters, called respectively the *perturbation parameter* and the *frequency vector of the forcing*, $F : \mathcal{A} \times \mathbb{T}^m \to \mathbb{R}^n$ is a real analytic function, with $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ and $\mathcal{A} \subset \mathbb{R}^n$ an open set, and $D_\varepsilon$ is a differential operator possibly depending on $\varepsilon$,

$$D_\varepsilon = \partial_t^{i_0} + a_1 \varepsilon \partial_t^{i_1} + a_2 \varepsilon^2 \partial_t^{i_2} + \dots, \qquad i_0, i_1, i_2, \dots \in \mathbb{N}. \tag{1.2}$$

In particular we shall consider explicitly the cases $D_\varepsilon = \partial_t^2$ and $D_\varepsilon = \partial_t + \varepsilon \partial_t^2$. The case $m = 0$ is allowed and corresponds to $F = F(u)$. If $m \geq 1$, we can assume without loss of generality that the vector $\underline{\omega}$ has rationally independent components.

We also assume that for $\varepsilon = 0$ the *unperturbed equation*

$$D_0 u = 0 \tag{1.3}$$

admits a quasi-periodic solution $u_0 = u_0(\boldsymbol{\omega} t)$, with $\boldsymbol{\omega} \in \mathbb{R}^p$, $p \geq 0$, possibly trivial (that is $p = 0$, which gives a constant). For instance, if $u$ is an angle, $u \in \mathbb{T}^n$, and $D_\varepsilon = \partial_t^2$, one has $u_0 = c_0 + \Omega t$, with $c_0, \Omega \in \mathbb{R}^n$; up to a linear change of coordinates, we can always write $\Omega = (\boldsymbol{\omega}, 0, \dots, 0)$, such that the vector $\boldsymbol{\omega} \in \mathbb{R}^p$, $p \leq n$, has rationally independent components. If $u \in \mathbb{R}^n$ and, say, $D_\varepsilon = \partial_t + \varepsilon \partial_t^2$, one has $u_0 = c_0$, where $c_0 \in \mathbb{R}^n$ is a constant vector.

We are interested in quasi-periodic solutions to (1.1) with *rotation vector* (or *frequency vector*) $\omega = (\boldsymbol{\omega}, \underline{\omega}) \in \mathbb{R}^d$, $d = p + m$, that is solutions of the form $u = u(\omega t, \varepsilon)$, with

$$u(\psi, \varepsilon) = \sum_{\nu \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i}\nu \cdot \psi} u_\nu(\varepsilon), \tag{1.4}$$

where $\nu \cdot \psi = \nu_1\psi_1 + \ldots + \nu_d\psi_d$ denotes the standard scalar product. For $m = 0$ we have $d = p$ and $\omega = \boldsymbol{\omega}$, while for $p = 0$ we have $d = m$ and $\omega = \underline{\omega}$. A suitable Diophantine condition will be assumed on $\omega$, for instance the *standard Diophantine condition*

$$|\omega \cdot \nu| > \gamma|\nu|^{-\tau} \quad \forall \nu \in \mathbb{Z}^d \setminus \{0\}, \tag{1.5}$$

where $\gamma > 0$ and $\tau \geq d-1$ are the *Diophantine constant* and the *Diophantine exponent*, respectively, and $|\nu| = |\nu_1| + \ldots + |\nu_d|$. Weaker Diophantine conditions could be considered – see Section 12.1.

The operator $D_\varepsilon$ acts as a multiplication operator in Fourier space, that is

$$(D_\varepsilon u)_\nu = \delta(\omega \cdot \nu, \varepsilon)\, u_\nu, \tag{1.6}$$

with $\delta(x, \varepsilon) = \delta_0(x) + \varepsilon\delta_1(x) + \varepsilon^2\delta_2(x) + \ldots$. For instance, if $D_\varepsilon = \partial_t^2$ then $\delta(\omega\cdot\nu, \varepsilon) = \delta_0(\omega\cdot\nu) = (\mathrm{i}\omega\cdot\nu)^2$, if $D_\varepsilon = \partial_t + \varepsilon\partial_t^2$ then $\delta(\omega \cdot \nu, \varepsilon) = \mathrm{i}\omega \cdot \nu\,(1 + \mathrm{i}\varepsilon\omega \cdot \nu)$, and so on: we shall explicitly take $\delta(\omega \cdot \nu, \varepsilon)$ to be a polynomial in $\varepsilon$. In other words $D_\varepsilon$ can be expanded as

$$D_\varepsilon = \sum_{k=0}^{k_0} \varepsilon^k D^{(k)}, \qquad D^{(0)} = D_0, \tag{1.7}$$

with $k_0 \in \mathbb{N}$ and $D^{(k)} = a_k\partial_t^{i_k}$, so that $\delta_k(\omega\cdot\nu) = a_k(\mathrm{i}\omega\cdot\nu)^{i_k}$, for $1 \leq k \leq k_0$. The Diophantine condition (1.5) implies

$$|\delta_0(\omega \cdot \nu)| \geq \gamma_0|\nu|^{-\tau_0}, \qquad \gamma_0 = \gamma^{i_0}, \quad \tau_0 = i_0\tau. \tag{1.8}$$

The problem we address here is to find a quasi-periodic solution $u(\omega t, \varepsilon)$ to the full equation (1.1), which continues the unperturbed solution $u_0(\boldsymbol{\omega}t)$, that is such that it reduces to $u_0(\boldsymbol{\omega}t)$ as $\varepsilon \to 0$. This means that we look for results holding for small values of the parameter $\varepsilon$. Hence (1.1) can be seen as a *perturbation* of the equation (1.3), and this explains why $\varepsilon$ is called the perturbation parameter.

More precisely we shall be interested in both the existence and stability of such quasi-periodic solutions. In particular, our analysis accounts for the KAM theory for quasi-integrable Hamiltonian systems (in a special case) and the existence of quasi-periodic attractors for strongly dissipative quasi-periodically forced one-dimensional systems. We can also consider discrete systems, as opposite to the continuous ones such as (1.1). We shall see that in both cases the existence of a quasi-periodic solution for the dynamical system is reduced to existence of a solution for a suitable functional equation – see Section 3.2 for more insight. Extensions to more general systems will be briefly discussed in Section 12.

The method we shall follow uses renormalisation group ideas, and is based on techniques of multiscale analysis which are typical of quantum field theory. The method is widely inspired to the original work of Eliasson [29] and, even more, to its reinterpretation given by Gallavotti [32]. The deep analogy with quantum field theory was stressed and used to full extend in subsequent papers; see for instance [46, 47].

For other renormalisation group approaches existing in the literature to the same kind of problems considered in this review see for instance [65, 17, 60, 59].

# 2 Perturbation theory and formal solutions

As first attempt, we look for quasi-periodic solutions (1.4) to (1.1) in the form of *formal power series* in the perturbation parameter $\varepsilon$,

$$u(\psi, \varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k u^{(k)}(\psi), \qquad u^{(k)}(\psi) = \sum_{\nu\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}\nu\cdot\psi} u_\nu^{(k)}, \quad k \geq 1, \tag{2.1}$$

where $u^{(0)} = u_0$ is such that $D_0u_0 = 0$. The power series expansion (2.1) will be referred to as the *perturbation series* for the quasi-periodic solution. Perturbation series have been widely studied in the

literature, especially in connection with problems of celestial mechanics [71]. They are sometimes called the *Lindstedt series* or *Lindstedt-Newcomb series*, from the name of the astronomers who first studied them in a systematical way.

If we expand also $D_\varepsilon$ according to (1.7), we obtain to all orders $k \geq 1$

$$D_0 u^{(k)} = - \sum_{p=1}^{\min\{k,k_0\}} D^{(p)} u^{(k-p)} + [F(u, \underline{\omega} t)]^{(k-1)} , \tag{2.2}$$

where $[F(u, \underline{\omega} t)]^{(k)}$ means that we take the Taylor expansion of the function $u \to F(u, \cdot)$, then we expand $u$ in powers of $\varepsilon$ according to (2.1), and we keep the coefficient of $\varepsilon^k$, that is

$$[F(u, \underline{\omega} t)]^{(k)} = \sum_{s=0}^{\infty} \frac{1}{s!} \partial_u^s F(u_0(\boldsymbol{\omega} t), \underline{\omega} t) \sum_{\substack{k_1, \ldots, k_s \geq 1 \\ k_1 + \ldots + k_s = k}} u^{(k_1)} \ldots u^{(k_s)}, \tag{2.3}$$

which for $k = 0$ reads $[F(u, \underline{\omega} t)]^{(0)} = F(u_0(\boldsymbol{\omega} t), \underline{\omega} t)$.

If $u_0 = \boldsymbol{c}_0 + \boldsymbol{\omega} t$ (and hence $p = n$), we expand

$$F(u_0, \underline{\omega} t) = F(\boldsymbol{c}_0 + \boldsymbol{\omega} t, \underline{\omega} t) = \sum_{\nu_0 \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i} \boldsymbol{\nu}_0 \cdot \boldsymbol{c}_0} \mathrm{e}^{\mathrm{i} \nu_0 \cdot \omega t} F_{\nu_0}, \tag{2.4}$$

where $\nu_0 = (\boldsymbol{\nu}_0, \underline{\nu}_0)$, so as to obtain in Fourier space

$$\delta_0 (\omega \cdot \nu)\, u_\nu^{(k)} = - \sum_{p=1}^{\min\{k,k_0\}} \delta_p (\omega \cdot \nu)\, u_\nu^{(k-p)} + [F(u, \underline{\omega} t)]_\nu^{(k-1)} , \tag{2.5}$$

with

$$[F(u, \underline{\omega} t)]_\nu^{(k)} = \sum_{s=0}^{\infty} \frac{1}{s!} \sum_{\substack{\nu_0, \nu_1, \ldots, \nu_s \in \mathbb{Z}^d \\ \nu_0 + \nu_1 + \ldots + \nu_s = \nu}} (\mathrm{i} \boldsymbol{\nu})^s \, F_{\nu_0} \mathrm{e}^{\mathrm{i} \boldsymbol{\nu}_0 \cdot \boldsymbol{c}_0} \sum_{\substack{k_1, \ldots, k_s \geq 1 \\ k_1 + \ldots + k_s = k}} u_{\nu_1}^{(k_1)} \ldots u_{\nu_s}^{(k_s)}. \tag{2.6}$$

By the analyticity assumption on the function $F$ one has $|F_\nu| \leq \Xi_0 \mathrm{e}^{-\xi |\nu|}$ for suitable constants $\xi, \Xi_0 > 0$.

If $u_0 = c_0$ one has $\omega = \underline{\omega}$. In that case we Fourier expand $F(u, \underline{\omega} t)$ only in the argument $\underline{\omega} t$,

$$F(u, \underline{\omega} t) = \sum_{\underline{\nu}_0 \in \mathbb{Z}^m} \mathrm{e}^{\mathrm{i} \underline{\nu}_0 \cdot \underline{\omega} t} F_{\underline{\nu}_0}(u), \tag{2.7}$$

and we still obtain (2.5), but with $\nu = \underline{\nu}$ and

$$[F(u, \underline{\omega} t)]_{\underline{\nu}}^{(k)} = \sum_{s=0}^{\infty} \frac{1}{s!} \sum_{\substack{\underline{\nu}_0, \underline{\nu}_1, \ldots, \underline{\nu}_s \in \mathbb{Z}^m \\ \underline{\nu}_0 + \underline{\nu}_1 + \ldots + \underline{\nu}_s = \underline{\nu}}} \partial_u^s F_{\underline{\nu}_0}(c_0) \sum_{\substack{k_1, \ldots, k_s \geq 1 \\ k_1 + \ldots + k_s = k}} u_{\underline{\nu}_1}^{(k_1)} \ldots u_{\underline{\nu}_s}^{(k_s)}. \tag{2.8}$$

More generally, one expands

$$\partial_u^s F(u_0, \underline{\omega} t) = \sum_{\nu_0 \in \mathbb{Z}^d} \mathrm{e}^{\mathrm{i} \nu_0 \cdot \omega t} \mathcal{F}_{s, \nu_0}, \tag{2.9}$$

with coefficients $\mathcal{F}_{s,\nu}$ bounded as $|\mathcal{F}_{s,\nu}| \leq s! \Xi_0 \Xi_1^s \mathrm{e}^{-\xi |\nu|}$ for suitable constants $\xi, \Xi_0, \Xi_1 > 0$, so that

$$[F(u, \underline{\omega} t)]_\nu^{(k)} = \sum_{s=0}^{\infty} \frac{1}{s!} \sum_{\substack{\nu_0, \nu_1, \ldots, \nu_s \in \mathbb{Z}^d \\ \nu_0 + \nu_1 + \ldots + \nu_s = \nu}} \mathcal{F}_{s, \nu_0} \sum_{\substack{k_1, \ldots, k_s \geq 1 \\ k_1 + \ldots + k_s = k}} u_{\nu_1}^{(k_1)} \ldots u_{\nu_s}^{(k_s)}. \tag{2.10}$$

Then, to all order $k \geq 1$ one obtains

$$u_\nu^{(k)} = \frac{1}{\delta_0(\omega\cdot\nu)} \left( - \sum_{p=1}^{\min\{k,k_0\}} \delta_p(\omega\cdot\nu)\, u_\nu^{(k-p)} + [F(u,\underline{\omega}t)]_\nu^{(k-1)} \right), \qquad \nu \neq 0, \tag{2.11}$$

provided the *compatibility condition*

$$0 = [F(u,\underline{\omega}t)]_0^{(k)}\,, \tag{2.12}$$

holds for all $k \geq 0$. Equations (2.11) formally provide a recursive definition of the coefficients $u_\nu^{(k)}$, $\nu \neq 0$, in terms of the coefficients $u_{\nu'}^{(k')}$ of lower orders $k' < k$. Indeed, the Diophantine condition (1.5) ensures that no denominator can be zero, and the sum over the order and Fourier labels can be easily performed (as we shall check explicitly later on – see Section 4). Thus, if the compatibility condition turns out to be satisfied to all orders for a suitable choice of the coefficients $u_0^{(k)}$, one has an algorithm which allows to construct iteratively all the coefficients of the series (2.1). In that case, we say that the equations (1.1) are *formally solvable*.

# 3 Examples

In this section we review some examples of physically relevant dynamical systems which can be written in the form (1.1). Therefore, for all such systems the problem of existence of formal quasi-periodic solutions is reduced to that of showing that (2.11) and (2.12) can be recursively solved. The strategy that we shall follow in the next sections will be to show first that a formal power series solves those equations order by order, and, then, to study the convergence of the series for $\varepsilon$ small enough. We shall be able to prove that either the series converges or it can be suitably resummed so as to give a well-defined function which solves the equation (1.1).

## 3.1 Quasi-integrable Hamiltonian systems: maximal tori

Consider the Hamiltonian

$$\mathcal{H}(\boldsymbol{\alpha},\boldsymbol{A}) = \frac{1}{2}\boldsymbol{A}^2 + \varepsilon f(\boldsymbol{\alpha}), \qquad \boldsymbol{A}^2 = \boldsymbol{A}\cdot\boldsymbol{A} = A_1^2 + \ldots + A_n^2, \tag{3.1}$$

where $(\boldsymbol{\alpha},\boldsymbol{A}) \in \mathbb{T}^n \times \mathbb{R}^n$ are *angle-action* coordinates, and $f\colon \mathbb{T}^n \to \mathbb{R}$ is a real analytic function.

More generally we could consider Hamiltonians of the form

$$\mathcal{H}(\boldsymbol{\alpha},\boldsymbol{A}) = \mathcal{H}_0(\boldsymbol{A}) + \varepsilon f(\boldsymbol{\alpha},\boldsymbol{A}), \tag{3.2}$$

where $\mathcal{H}_0$ and $f$ are both real analytic functions in $\mathbb{T}^n \times \mathcal{A}$, $\mathcal{A} \subset \mathbb{R}^n$ being an open set, with $\mathcal{H}_0$ convex and $f$ periodic in the angles $\boldsymbol{\alpha}$. Existence and properties of quasi-periodic solutions for *quasi-integrable systems*, that is systems described by Hamiltonians of the form (3.2), are the content of *KAM theory* [61, 2, 68]; see [3] for a review.

The advantage of taking the Hamiltonian (3.1) is that the corresponding Hamilton equations lead to a closed equation for $\boldsymbol{\alpha}$,

$$\ddot{\boldsymbol{\alpha}} = -\varepsilon\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}), \tag{3.3}$$

which is the form (1.1) with $u = \boldsymbol{\alpha}$, $m = 0$, and $D_\varepsilon = \partial_t^2$ (so that $\delta_0(\omega\cdot\nu) = (\mathrm{i}\omega\cdot\nu)^2$, while $\delta_k(\omega\cdot\nu) = 0$ for all $k \geq 1$).

We are interested in quasi-periodic solutions of the form (1.4) with $d = p = n$ (and $\omega = \boldsymbol{\omega}$). For $\varepsilon = 0$ we take as unperturbed solution $u_0(\boldsymbol{\omega}t) = \boldsymbol{\alpha}_0 + \boldsymbol{\omega}t$, with $\boldsymbol{\alpha}_0 \in \mathbb{T}^d$ arbitrary and $\omega = \boldsymbol{\omega} \in \mathbb{R}^d$ satisfying the Diophantine condition (1.5).

If we write, in agreement with (2.1),

$$\boldsymbol{\alpha}(\boldsymbol{\omega}t,\varepsilon)=\boldsymbol{\alpha}_0+\boldsymbol{\omega}t+\sum_{k=1}^{\infty}\varepsilon^k\sum_{\boldsymbol{\nu}\in\mathbb{Z}^n}\mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\boldsymbol{\nu}t}\boldsymbol{\alpha}_{\boldsymbol{\nu}}^{(k)},\tag{3.4}$$

we find to all orders $k\geq 1$ – see (2.11) and (2.12) –

$$\boldsymbol{\alpha}_{\boldsymbol{\nu}}^{(k)}=\frac{1}{(\omega\cdot\boldsymbol{\nu})^2}\left[\partial_{\boldsymbol{\alpha}}f(\boldsymbol{\alpha})\right]_{\boldsymbol{\nu}}^{(k-1)},\qquad\boldsymbol{\nu}\neq\mathbf{0},\tag{3.5a}$$

$$\mathbf{0}=\left[\partial_{\boldsymbol{\alpha}}f(\boldsymbol{\alpha})\right]_{\mathbf{0}}^{(k-1)},\tag{3.5b}$$

where

$$\left[\partial_{\boldsymbol{\alpha}}f(\boldsymbol{\alpha})\right]_{\boldsymbol{\nu}}^{(k)}=\sum_{s=0}^{\infty}\frac{1}{s!}\sum_{\substack{\boldsymbol{\nu}_0,\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s\in\mathbb{Z}^n\\ \boldsymbol{\nu}_0+\boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_s=\boldsymbol{\nu}}}(\mathrm{i}\boldsymbol{\nu}_0)^{s+1}f_{\boldsymbol{\nu}_0}\sum_{\substack{k_1,\ldots,k_s\geq 1\\ k_1+\ldots+k_s=k}}\boldsymbol{\alpha}_{\boldsymbol{\nu}_1}^{(k_1)}\ldots\boldsymbol{\alpha}_{\boldsymbol{\nu}_s}^{(k_s)}.\tag{3.6}$$

Thus, equation (3.5a) defines recursively the coefficients $\boldsymbol{\alpha}_{\boldsymbol{\nu}}^{(k)}$ for all $k\geq 1$ and all $\boldsymbol{\nu}\neq\mathbf{0}$, provided equation (3.5b) is satisfied for all $k\geq 1$.

The compatibility condition (3.5b) is automatically satisfied for $k=1$, because $\left[\partial_{\boldsymbol{\alpha}}f(\boldsymbol{\alpha})\right]_{\boldsymbol{\nu}}^{(0)}=\mathrm{i}\boldsymbol{\nu}f_{\boldsymbol{\nu}}$, which vanishes for $\boldsymbol{\nu}=\mathbf{0}$. It is a remarkable cancellation that the condition holds for all $k\geq 1$ – see Section 4.4 –, so implying that the perturbation series (3.4) is well-defined to all orders. The coefficients $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)}$, $k\geq 1$, can be arbitrarily fixed; for instance one can set $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)}=\mathbf{0}$ for all $k\geq 1$ – see Section 4.4. We shall see in Section 7 that for $\varepsilon$ small enough the series converges to a function analytic in $\boldsymbol{\psi}=\boldsymbol{\omega}t$. As a consequence, there exists a quasi-periodic solution of the form (3.4), analytic both in $\varepsilon$ and $\boldsymbol{\psi}$, and parameterised by $\boldsymbol{\alpha}_0\in\mathbb{T}^n$: hence such a solution describes an $n$-dimensional invariant torus (*maximal KAM torus*).

In this paper we confine ourselves to the Hamiltonian (3.1). Note that in the case (3.2) the unperturbed solution is $u_0=(\boldsymbol{\alpha}_0+\boldsymbol{\omega}(\boldsymbol{A}_0)t,\boldsymbol{A}_0)$, with $\boldsymbol{\omega}(\boldsymbol{A})=\partial_{\boldsymbol{A}}\mathcal{H}_0(\boldsymbol{A})$, hence it is still of the form $\boldsymbol{c}_0+\boldsymbol{\omega}t$ with $n=2p$ and $\omega_i=0$ for $i\geq n+1$. However, strictly speaking the Hamilton equation are not of the form (1.1), so the analysis should be suitably adapted – see [47].

We could also consider, instead of (3.3), the more general equation

$$\ddot{\boldsymbol{\alpha}}=-\varepsilon\partial_{\boldsymbol{\alpha}}f(\boldsymbol{\alpha},\underline{\omega}t),\tag{3.7}$$

which reduces to (3.3) for $m=0$. This is still a Hamiltonian system, with Hamiltonian

$$\mathcal{H}(\boldsymbol{\alpha},\underline{\alpha},\boldsymbol{A},\underline{A})=\frac{1}{2}\boldsymbol{A}\cdot\boldsymbol{A}+\underline{\omega}\cdot\underline{A}+\varepsilon f(\boldsymbol{\alpha},\underline{\alpha}).\tag{3.8}$$

The unperturbed solution to (3.7) is of the same form as before, but in that case we look for a quasi-periodic solution with rotation vector $\omega=(\boldsymbol{\omega},\underline{\omega})$.

Then the angles $\underline{\alpha}$ evolve trivially as $\underline{\alpha}(t)=\underline{\omega}t$, whereas equation (3.5a) has to be replaced with

$$\boldsymbol{\alpha}_{\nu}^{(k)}=\frac{1}{(\omega\cdot\nu)^2}\sum_{s=0}^{\infty}\frac{1}{s!}\sum_{\substack{\nu_0,\nu_1,\ldots,\nu_s\in\mathbb{Z}^d\\ \nu_0+\nu_1+\ldots+\nu_s=\nu}}(\mathrm{i}\boldsymbol{\nu}_0)^{s+1}f_{\nu_0}\sum_{\substack{k_1,\ldots,k_s\geq 1\\ k_1+\ldots+k_s=k-1}}\boldsymbol{\alpha}_{\nu_1}^{(k_1)}\ldots\boldsymbol{\alpha}_{\nu_s}^{(k_s)},\tag{3.9}$$

while the compatibility condition for $\nu=0$ reads

$$\mathbf{0}=\sum_{s=0}^{\infty}\frac{1}{s!}\sum_{\substack{\nu_0,\nu_1,\ldots,\nu_s\in\mathbb{Z}^d\\ \nu_0+\nu_1+\ldots+\nu_s=0}}(\mathrm{i}\boldsymbol{\nu}_0)^{s+1}f_{\nu_0}\sum_{\substack{k_1,\ldots,k_s\geq 1\\ k_1+\ldots+k_s=k-1}}\boldsymbol{\alpha}_{\nu_1}^{(k_1)}\ldots\boldsymbol{\alpha}_{\nu_s}^{(k_s)}.\tag{3.10}$$

However, the differences with respect to the previous case are just minor ones, as one can easily work out by himself.

## 3.2 Discrete systems: the standard map

Beside continuous dynamical systems we can consider discrete dynamical systems (maps), such as the *standard map* [21, 64] – also known as Chirikov-Greene-Taylor map.

The standard map is defined by the symplectic map from the cylinder to itself

$$\begin{cases} x' = x + y + \varepsilon \sin x, \\ y' = y + \varepsilon \sin x, \end{cases} \tag{3.11}$$

where $(x, y) \in \mathbb{T} \times \mathbb{R}$.

For $\varepsilon = 0$ the motion is trivial: one has a simple rotation $x' = x + 2\pi\omega$, while $y$ is fixed to $y = 2\pi\omega$, $\omega \in \mathbb{R}$. For $\varepsilon \neq 0$ one can look for solutions which are conjugate to a trivial rotation of some other variable (KAM invariant curves), i.e. one can look for solutions of the form

$$x = \alpha + u(\alpha, \varepsilon), \qquad y = 2\pi\omega + v(\alpha, \varepsilon), \tag{3.12}$$

with $\alpha \to \alpha' = \alpha + 2\pi\omega$ and the functions $u, v$ depending analytically on their arguments. The number $\omega$ will be called the *rotation number*.

The functions $u, v$ are not independent from each other. One has $v(\alpha, \varepsilon) = u(\alpha, \varepsilon) - u(\alpha - 2\pi\omega, \varepsilon)$, as it is straightforward to check: simply note that $x' = x + y'$ by (3.11) and express $x'$ and $y'$ in terms of $\alpha$ through (3.12), using that $\alpha' = \alpha + 2\pi\omega$. So one obtains a closed equation for the *conjugating function* $u$,

$$Du(\alpha, \varepsilon) \equiv u(\alpha + 2\pi\omega, \varepsilon) + u(\alpha - 2\pi\omega, \varepsilon) - 2u(\alpha, \varepsilon) = \varepsilon \sin(\alpha + u(\alpha, \varepsilon)). \tag{3.13}$$

If such a solution exists and is analytic in $\varepsilon$, then it has to be possible to expand the function $u$ as Taylor series in $\varepsilon$ and as Fourier series in $\alpha$. So we are led to write, at least formally,

$$u(\alpha, \varepsilon) = \sum_{\nu \in \mathbb{Z}} \sum_{k=1}^{\infty} \mathrm{e}^{\mathrm{i}\nu\alpha} \varepsilon^k u_\nu^{(k)}. \tag{3.14}$$

so implying, by calling $[F(\alpha, \varepsilon)]_\nu^{(k)}$ the coefficient of the function $F(\alpha, \varepsilon)$ with Fourier label $\nu$ and Taylor label $k$, according to (2.3),

$$\delta_0(\omega\nu) u_\nu^{(k)} = \left[\sin(\alpha + u(\alpha, \varepsilon))\right]_\nu^{(k-1)}, \qquad \delta_0(\omega\nu) = 2\left[\cos(2\pi\omega\nu) - 1\right]. \tag{3.15}$$

Note that, for $\omega\nu$ small (mod. 1), $\delta_0(\omega\nu) \sim \|\omega\nu\|^2$, if $\|x\| = \min_{p \in \mathbb{Z}} |x - p|$ denotes the distance of $x$ from the nearest integer. By explicitly writing $\sin\alpha = \sum_{\nu_0 = \pm 1} (2\mathrm{i})^{-1} \nu_0 \mathrm{e}^{\mathrm{i}\nu_0\alpha}$ and Taylor expanding $\sin(\alpha + u)$ in $u$ around $u = 0$, one finds, from (3.13),

$$u_\nu^{(1)} = -\frac{\mathrm{i}\nu}{2\delta_0(\omega\nu)}, \tag{3.16a}$$

$$u_\nu^{(k)} = \frac{1}{\delta_0(\omega\nu)} \sum_{s=0}^{\infty} \sum_{\substack{\nu_0 + \nu_1 + \ldots + \nu_s = \nu \\ k_1 + \ldots + k_s = k-1}} \frac{-(\mathrm{i}\nu_0)^{s+1}}{s!2} u_{\nu_1}^{(k_1)} \ldots u_{\nu_s}^{(k_s)}, \qquad k > 1, \tag{3.16b}$$

for $\nu \neq 0$. It is easy to check that for all $k \geq 1$ one has $u_\nu^{(k)} = 0$ if $|\nu| > k$, and one can choose $u_0^{(k)} = 0$. Then (3.16) can be iterate by taking into account that $k_j < k$ for any $j = 1, \ldots, s$.

When working in Fourier space, the recursive equations (3.16) look very similar to the equations (3.5a) for continuous systems, with $n = 2$. Then, one could be tempted to write the standard map as the stroboscopic map of a continuous system. However, it turns out that, formally, one should consider the singular system (known in physics as the *kicked rotator*) with Hamiltonian

$$\mathcal{H}(\boldsymbol{\alpha}, \boldsymbol{A}) = 2\pi A_1 + \frac{A_2^2}{2} + 2\pi\varepsilon \sum_{n \in \mathbb{Z}} \delta(\alpha_1 - 2\pi n) \left(\cos\alpha_2 - 1\right), \tag{3.17}$$

where $\delta$ is the delta function, and set $x = \alpha_2$ and $y = A_2 - (\varepsilon/2)\sin\alpha_2$ [34]. Therefore the corresponding Hamilton equations cannot be in the class (1.1), where smoothness was required. Nonetheless, as we have seen, in Fourier space the analysis is essentially the same.

### 3.3 Quasi-integrable Hamiltonian systems: lower dimensional tori

Consider the Hamiltonian

$$\mathcal{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{A},\boldsymbol{B}) = \frac{1}{2}\boldsymbol{A}^2 + \frac{1}{2}\boldsymbol{B}^2 + \varepsilon\, f(\boldsymbol{\alpha},\boldsymbol{\beta}), \tag{3.18}$$

where $(\boldsymbol{\alpha},\boldsymbol{A}) \in \mathbb{T}^r \times \mathbb{R}^r$ and $(\boldsymbol{\beta},\boldsymbol{B}) \in \mathbb{T}^s \times \mathbb{R}^s$ are angle-action coordinates, with $r+s = n$, and $f\colon \mathbb{T}^n \to \mathbb{R}$ is a real analytic function.

We can also consider the same Hamiltonian as (3.1), but assume that $\boldsymbol{\omega}$ is a resonant vector, that is that there exist $s$ integer vectors $\boldsymbol{\nu}_1,\dots,\boldsymbol{\nu}_s$ such that $\boldsymbol{\omega}\cdot\boldsymbol{\nu}_1 = \ldots = \boldsymbol{\omega}\cdot\boldsymbol{\nu}_s = 0$. If this happens, it is possible to perform a linear change of coordinates such that in the new coordinates $\boldsymbol{\omega} = (\omega_1,\dots,\omega_r,0,\dots,0)$, with $(\omega_1,\dots,\omega_r)$ satisfying a Diophantine condition in $\mathbb{R}^r$. Of course the corresponding Hamiltonian would be slightly more complicated than (3.18). For simplicity's sake we shall confine ourselves to (3.18), and study the problem of existence of a quasi-periodic solution with rotation vector $\boldsymbol{\omega} = (\omega_1,\dots,\omega_r)$ which for $\varepsilon = 0$ reduces to $u_0 = (\boldsymbol{\alpha}_0 + \boldsymbol{\omega}t, \boldsymbol{\beta}_0, \boldsymbol{A}_0, \boldsymbol{0})$, with $\boldsymbol{A}_0 = \boldsymbol{\omega}$.

In terms of the angles $(\boldsymbol{\alpha},\boldsymbol{\beta})$ the Hamilton equations become

$$\begin{cases} \ddot{\boldsymbol{\alpha}} = -\varepsilon\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha},\boldsymbol{\beta}), \\ \ddot{\boldsymbol{\beta}} = -\varepsilon\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha},\boldsymbol{\beta}), \end{cases} \tag{3.19}$$

so that the unperturbed solution is $(\boldsymbol{\alpha}_0 + \boldsymbol{\omega}t, \boldsymbol{\beta}_0)$, which is of the form $u_0 = c_0 + \Omega t$, with $\Omega = (\omega_1,\dots,\omega_r,0,\dots,0) = (\boldsymbol{\omega},\boldsymbol{0})$. Hence we look for solutions of the form (1.4) with $d = p = r$.

Since $\boldsymbol{\beta}$ is expected to remain close to $\boldsymbol{\beta}_0$, we Fourier expand $f$ only in the angle $\boldsymbol{\alpha}$, so writing

$$f(\boldsymbol{\alpha},\boldsymbol{\beta}) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}\boldsymbol{\nu}\cdot\boldsymbol{\alpha}} f_{\boldsymbol{\nu}}(\boldsymbol{\beta}). \tag{3.20}$$

Therefore, (3.19) gives, in Fourier space,

$$\begin{cases} (\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2\, \boldsymbol{\alpha}_{\boldsymbol{\nu}}^{(k)} = [\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha},\boldsymbol{\beta})]_{\boldsymbol{\nu}}^{(k-1)}, \\ (\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2\, \boldsymbol{\beta}_{\boldsymbol{\nu}}^{(k)} = [\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha},\boldsymbol{\beta})]_{\boldsymbol{\nu}}^{(k-1)}, \end{cases} \tag{3.21}$$

with

$$[\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha},\boldsymbol{\beta})]_{\boldsymbol{\nu}}^{(k)} = \sum_{p,q=0}^{\infty} \frac{1}{p!q!} \sum_{\substack{\boldsymbol{\nu}_0,\boldsymbol{\nu}_1,\dots,\boldsymbol{\nu}_{p+q}\in\mathbb{Z}^n \\ \boldsymbol{\nu}_0+\boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_{p+q}=\boldsymbol{\nu}}} (\mathrm{i}\boldsymbol{\nu}_0)^{p+1}\, \partial_{\boldsymbol{\beta}}^q f_{\boldsymbol{\nu}_0}(\boldsymbol{\beta}_0) \sum_{\substack{k_1,\dots,k_{p+q}\geq 1 \\ k_1+\ldots+k_{p+q}=k}} \boldsymbol{\alpha}_{\boldsymbol{\nu}_1}^{(k_1)} \dots \boldsymbol{\alpha}_{\boldsymbol{\nu}_p}^{(k_p)} \boldsymbol{\beta}_{\boldsymbol{\nu}_{p+1}}^{(k_{p+1})} \dots \boldsymbol{\beta}_{\boldsymbol{\nu}_{p+q}}^{(k_{p+q})}, \tag{3.22}$$

and an analogous expression holding for $[\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha},\boldsymbol{\beta})]_{\boldsymbol{\nu}}^{(k)}$ – with $(\mathrm{i}\boldsymbol{\nu}_0)^p\, \partial_{\boldsymbol{\beta}}^{q+1}$ instead of $(\mathrm{i}\boldsymbol{\nu}_0)^{p+1}\, \partial_{\boldsymbol{\beta}}^q$.

Again for $\boldsymbol{\nu} = \boldsymbol{0}$ we require both

$$[\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha},\boldsymbol{\beta})]_{\boldsymbol{0}}^{(k)} = \boldsymbol{0}, \qquad [\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha},\boldsymbol{\beta})]_{\boldsymbol{0}}^{(k)} = \boldsymbol{0}, \tag{3.23}$$

for all $k \geq 0$. We shall see that the first compatibility condition is automatically satisfied for all values of $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_{\boldsymbol{0}}^{(k)}$, $k \geq 1$, while the second one requires $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_{\boldsymbol{0}}^{(k)}$, $k \geq 1$, to be suitably fixed. This is clear already to first order, where we obtain

$$\begin{cases} (\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2\, \boldsymbol{\alpha}_{\boldsymbol{\nu}}^{(1)} = \mathrm{i}\boldsymbol{\nu} f_{\boldsymbol{\nu}}(\boldsymbol{\beta}_0), \\ (\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2\, \boldsymbol{\beta}_{\boldsymbol{\nu}}^{(1)} = \partial_{\boldsymbol{\beta}} f_{\boldsymbol{\nu}}(\boldsymbol{\beta}_0), \end{cases} \tag{3.24}$$

so that for $\boldsymbol{\nu} = \mathbf{0}$ the first equation trivially holds, whereas the second one fixes $\boldsymbol{\beta}_0$ to be such that $\partial_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0) = \mathbf{0}$, i.e. $\boldsymbol{\beta}_0$ must be a stationary point for the function $f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ (such a point always exists).

Moreover to higher orders one has

$$[\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha}, \boldsymbol{\beta})]_{\mathbf{0}}^{(k)} = \partial_{\boldsymbol{\beta}}^2 f_{\mathbf{0}}(\boldsymbol{\beta}_0)\, \boldsymbol{\beta}_{\mathbf{0}}^{(k)} + \boldsymbol{\Phi}_k, \tag{3.25}$$

for a suitable function $\boldsymbol{\Phi}_k$ depending only on the coefficients of order strictly less than $k$ – see Section 4.4 for details. Thus, if we further assume that the matrix $\partial_{\boldsymbol{\beta}}^2 f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ be nonsingular (*nondegeneracy condition*), then we can impose the compatibility conditions $[\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha}, \boldsymbol{\beta})]_{\mathbf{0}}^{(k)} = \mathbf{0}$ by suitably fixing the corrections $\boldsymbol{\beta}_{\mathbf{0}}^{(k)}$, $k \geq 1$, to the constant part of the $\boldsymbol{\beta}$ angles.

In Section 4 we shall see that, at least formally, a quasi-periodic solution parameterised by $\boldsymbol{\alpha}_0 \in \mathbb{T}^r$ exists for suitable values of $\boldsymbol{\beta}_0$. We shall see in Section 8 that, even if the formal series is divergent, however it can be suitably resummed for $\varepsilon$ small enough so as to be given a meaning as a well-defined function analytic in $\boldsymbol{\psi} = \boldsymbol{\omega} t$: hence the latter describes a *lower-dimensional torus*. If the matrix $\partial_{\boldsymbol{\beta}}^2 f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ is positive definite we shall say that the lower-dimensional torus is *hyperbolic* if $\varepsilon < 0$ and *elliptic* if $\varepsilon > 0$; in the latter case we shall see that the torus exists only for some values of $\varepsilon$, more precisely for $\varepsilon$ defined in a Cantor set with Lebesgue density point at the origin (*Cantorisation*) – see Section 10.

## 3.4 Strongly dissipative quasi-periodically forced systems

Consider a one-dimensional system subject to a mechanical force $g$, in the presence of dissipation and of a quasi-periodically forcing. The equation describing the system is the ordinary differential equation

$$\ddot{x} + \gamma \dot{x} + g(x) = f(\underline{\omega} t), \tag{3.26}$$

where $x \in \mathbb{R}$, $\gamma > 0$ is the *dissipation coefficient* and $\underline{\omega} \in \mathbb{R}^m$ is the frequency vector of the forcing. We assume that both $g \colon \mathcal{A} \to \mathbb{R}$ and $f \colon \mathbb{T}^m \to \mathbb{R}$ are real analytic functions, with $\mathcal{A} \subset \mathbb{R}$ an open set.

If the dissipation is large enough, it is natural to rewrite (3.26) in terms of the small parameter $\varepsilon = 1/\gamma$, so as to obtain the equation

$$\dot{x} + \varepsilon \ddot{x} + \varepsilon g(x) = \varepsilon f(\underline{\omega} t), \tag{3.27}$$

which is of the form (1.1) with $n = 1$, $u = x$, $D_\varepsilon = \partial_t + \varepsilon \partial_t^2$, and $F(u, \underline{\omega} t) = -g(u) + f(\underline{\omega} t)$. In particular one has $\delta_0(\omega \cdot \nu) = \mathrm{i}\omega \cdot \nu$, $\delta_1(\omega \cdot \nu) = (\mathrm{i}\omega \cdot \nu)^2$, $a_1 = 1$, and $\delta_k(\omega \cdot \nu) = a_k = 0$ for all $k \geq 2$.

We look for quasi-periodic solutions $x(\underline{\omega} t, \varepsilon)$ which are analytic in $\underline{\psi} = \underline{\omega} t$ and continue the unperturbed solutions $x_0 = c_0$, with $c_0$ constant. Such solutions (if any) are called *response solutions*, as they have the same frequency vector as the forcing. Thus, $\omega = \underline{\omega}$ and $d = m$, so that, if we write

$$x(\underline{\omega} t, \varepsilon) = c_0 + \sum_{k=1}^{\infty} \varepsilon^k \sum_{\underline{\nu} \in \mathbb{Z}^m} \mathrm{e}^{\mathrm{i}\underline{\nu} \cdot \underline{\omega} t} x_{\underline{\nu}}^{(k)}, \tag{3.28}$$

then we obtain

$$\mathrm{i}\omega \cdot \underline{\nu}\, x_{\underline{\nu}}^{(k)} + (\mathrm{i}\underline{\omega} \cdot \underline{\nu})^2\, x_{\underline{\nu}}^{(k-1)} + [g(x)]_{\underline{\nu}}^{(k-1)} = f_{\underline{\nu}} \delta_{k,1} \tag{3.29}$$

for $k \geq 1$, while $x_{\underline{0}}^{(0)} = c_0$ and $x_{\underline{\nu}}^{(0)} = 0$ for $\underline{\nu} \neq \underline{0}$.

The first order equation gives

$$\begin{cases} \mathrm{i}\underline{\omega} \cdot \underline{\nu}\, x_{\underline{\nu}}^{(1)} = f_{\underline{\nu}}, & \underline{\nu} \neq \underline{0} \\ g(c_0) = f_{\underline{0}}, \end{cases} \tag{3.30}$$

which fixes the value of $c_0$ (of course $f_{\underline{0}}$ must belong to the range of $g$). The unperturbed solution can be seen as a quasi-periodic solution in the extended phase space $(x, \underline{\psi}) \in \mathbb{R} \times \mathbb{T}^m$, where it looks like $(c_0, \underline{\omega} t)$, so that the full solution is of the form $(c_0 + X(\underline{\omega} t, c_0, \varepsilon), \underline{\omega} t)$.

If $\partial_x g(c_0) \neq 0$, then to higher orders one has the compatibility conditions

$$[g(x)]_{\underline{0}}^{(k)} = \partial_x g(c_0)\, x_{\underline{0}}^{(k)} + G_k(c_0, x^{(1)}, \dots, x^{(k-1)}) = 0, \qquad k \geq 1, \tag{3.31}$$

for a suitable function $G_k$ depending only on the coefficients of orders $k' < k$. An explicit calculation gives

$$G_k(c_0, x^{(1)}, \dots, x^{(k-1)}) = \sum_{s=2}^{\infty} \frac{1}{s!} \partial_x^s g(c_0) \sum_{\substack{\underline{\nu}_1, \dots, \underline{\nu}_s \in \mathbb{Z}^m \\ \underline{\nu}_1 + \dots + \underline{\nu}_s = \underline{0}}} \sum_{\substack{k_1, \dots, k_s \geq 1 \\ k_1 + \dots + k_s = k}} x_{\underline{\nu}_1}^{(k_1)} \dots x_{\underline{\nu}_s}^{(k_s)}. \tag{3.32}$$

Thus, (3.31) can be used to fix the corrections $x_{\underline{0}}^{(k)}$, $k \geq 1$, to the constant part of the solution $x(\underline{\omega} t, \varepsilon)$.

We shall see in Section 9 that, under the *nondegeneracy condition* $\partial_x g(c_0) \neq 0$, the series (3.28) can be resummed for $\varepsilon$ small enough into a function which depends analytically on $\underline{\psi} = \underline{\omega} t$.

# 4 Diagrammatic representation and tree formalism

We have to study the recursive equations (2.11) and (2.12), with $[F(u, \underline{\omega} t)]_{\nu}^{(k-1)}$ given by (2.10). This will be done through a diagrammatic formalism, known as the *tree formalism*.

Let us assume that one can decompose $u = (\widetilde{u}, \widehat{u})$, with $\widetilde{u} \in \mathbb{R}^{\widetilde{n}}$ and $\widehat{u} \in \mathbb{R}^{\widehat{n}}$, $\widetilde{n} + \widehat{n} = n$, and, accordingly, $F = (\widetilde{F}, \widehat{F})$, so that for $\nu = 0$ one has

$$[\widetilde{F}(u, \underline{\omega} t)]_0^{(k)} = \widetilde{0}, \tag{4.1a}$$

$$[\widehat{F}(u, \underline{\omega} t)]_0^{(k)} = A\, \widehat{u}_0^{(k)} + \Phi_k, \tag{4.1b}$$

for a suitable nonsingular matrix $A$ and a suitable vector $\Phi_k$ ($\widetilde{0}$ is the null vector in $\mathbb{R}^{\widetilde{n}}$). We shall see that in all cases considered in Section 3 this holds true – see Section 4.4 below.

We first introduce the trees (see also [11, 56, 57]) as the main combinatorial and graphical objects that we shall use in the forthcoming analysis. Then we shall provide some rules how to associate numerical values to the trees, so as to represent the coefficients $u_{\nu}^{(k)}$ 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 by $N(\mathcal{G})$ 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 $\ell_0$ 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 *rooted tree* $\theta$ defined by $N(\theta) = N(\mathcal{G})$ and $L(\theta) = L(\mathcal{G}) \cup \ell_0$. A *labelled tree* is a rooted tree $\theta$ 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 $\ell$ 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 nodes* the nodes which have no entering line. We denote by $V(\theta)$ and $E(\theta)$ the set of internal nodes and end nodes, respectively. Of course $N(\theta) = V(\theta) \cup E(\theta)$.

The number of unlabelled trees (i.e. of rooted trees with no labels) with $N$ nodes – and hence with $N$ lines – is bounded by $2^{2N}$, which is a bound on the number of random walks with $2N$ steps [48]. An example of unlabelled tree is represented in Figure 1.

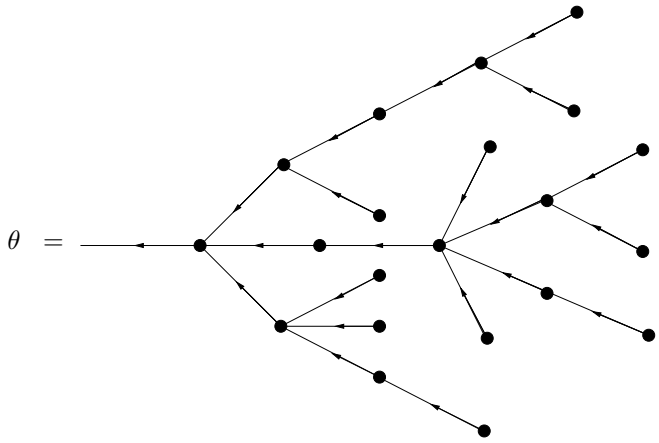


Figure 1: An example of unlabelled tree.

For each node $v$ denote by $S(v)$ the set of the lines entering $v$ and set $s_v = |S(v)|$; here and henceforth, given a set $A$, with denote by $|A|$ its cardinality. Hence $s_v = 0$ if $v$ is an end node, and $s_v \geq 1$ if $v$ is an internal node. One has

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

this can be easily checked by induction on the number of nodes of the tree.

## 4.2 Labels

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 constraints that $\nu_v \neq 0$ if $v \in E(\theta)$ and

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

which represents a *conservation rule* for each node.

We also associate with each node $v \in N(\theta)$ an *order* label $k_v \in \{0, 1, \dots, k_0\}$, such that $k_v = 0$ if

$\nu_{\ell_v} = 0$ and $k_v \geq 1$ if $\nu_{\ell_v} \neq 0$. We set

$$k(\theta) = \sum_{v \in N(\theta)} k_v, \qquad \nu(\theta) = \sum_{v \in N(\theta)} \nu_v, \tag{4.4}$$

which are called the *order* and the *momentum* of $\theta$, respectively; note that $\nu(\theta)$ is the momentum of the root line of $\theta$.

Finally we associate with each node $v \in V(\theta)$ a *badge* label $\rho_v \in \{0, 1\}$, such that $k_v \in \{0, 1\}$ if $\rho_v = 1$, while $\nu_{\ell_v} \neq 0$, $\nu_v = 0$, and $s_v = 1$ if $\rho_v = 0$.

Call $\mathcal{T}_{k,\nu}$ the set of all trees $\theta$ with order $k$ and momentum $\nu$, with the constraint that if a line $\ell \in L(\theta)$ has $\nu_\ell = 0$ and exits a node $v$ with $\nu_v = 0$ then $s_v \geq 2$. It is easy to check that there exists a positive constant $\kappa$ such that $k(\theta) \leq \kappa |N(\theta)|$; simply use that $s_v \neq 1$ when $k_v = 0$.

## 4.3 Diagrammatic rules

We want to show that trees naturally arise when studying the equations (2.11). Let $u_\nu^{(k)}$ be represented with the graph element in Figure 2 as a line with label $\nu$ exiting from a ball with label $(k)$.

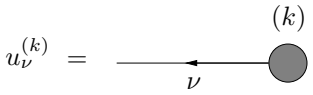


Figure 2: Graph element.

Then we can represent (2.11) graphically as depicted in Figure 3. Simply expand $[F(u, \underline{\omega} t)]_\nu^{(k)}$ as in (2.10) and represent each factor $u_{\nu_i}^{(k_i)}$ 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 $\nu$ of the root line is the sum of the mode label $\nu_0$ of the node $v_0$ plus the sum of the momenta of the lines entering $v_0$. Note that $k_{v_0} \geq 1$ as $\nu \neq 0$ in (2.11).

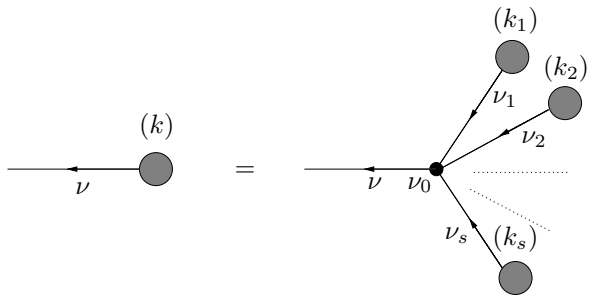


Figure 3: Graphical representation of the recursive equations.

We represent also (2.12) as in Figure 3, with the only difference that now $k_{v_0} = 0$ and hence $s_{v_0} \geq 2$ (recall the definition of $\mathcal{T}_{k,\nu}$ at the end of Section 4.2).

Given any tree $\theta \in \mathcal{T}_{k,\nu}$ we associate with each node $v \in N(\theta)$ a *node factor* $\mathfrak{F}_v$ and with each line $\ell \in L(\theta)$ a *propagator* $\mathfrak{G}_\ell$, by setting

$$\mathfrak{F}_v := \begin{cases} (s_v!)^{-1} \mathcal{F}_{s_v, \nu_v}, & \rho_v = 1, \\ -\delta_{k_v} (\omega \cdot \nu_{\ell_v}) \, \mathbb{1}, & \rho_v = 0, \end{cases} \qquad \mathfrak{G}_\ell := \begin{cases} \delta_0^{-1} (\omega \cdot \nu_\ell) \, \mathbb{1}, & \nu_\ell \neq 0, \\ G, & \nu_\ell = 0, \end{cases} \tag{4.5}$$

where $\mathbb{1}$ is the $n \times n$ identity, and $G$ is the $n \times n$ matrix of the form

$$G = \begin{pmatrix} 0 & 0 \\ 0 & -A^{-1} \end{pmatrix}, \tag{4.6}$$

where the null matrices 0 are $\widetilde{n} \times \widetilde{n}$, $\widetilde{n} \times \widehat{n}$, and $\widehat{n} \times \widetilde{n}$, respectively, while $A$ is the invertible matrix appearing in (4.1b). Define the *value* of the tree $\theta$ as

$$\mathrm{Val}(\theta) := \Big( \prod_{v \in N(\theta)} \mathfrak{F}_v \Big) \Big( \prod_{\ell \in L(\theta)} \mathfrak{G}_\ell \Big). \tag{4.7}$$

The propagators $\mathfrak{G}_\ell$ are matrices, whereas each $\mathfrak{F}_v$ is a tensor with $s_v + 1$ indices, which can be associated with the $s_v + 1$ lines entering or exiting $v$. In (4.7) the indices of the tensors $\mathfrak{F}_v$ must be contracted: this means that if a node $v$ is connected to a node $v'$ by a line $\ell$ then the indices of $\mathfrak{F}_v$ and $\mathfrak{F}_{v'}$ associated with $\ell$ are equal to each other, and eventually one has to sum over all the indices except that associated with the root line.

The node factors in (4.5) are bounded as $\max_{j_1,\dots,j_{s_v+1}} |(\mathfrak{F}_v)_{j_1 \dots j_{s_v+1}}| \le \Xi_0 \Xi_1^{s_v} \mathrm{e}^{-\xi|\nu_v|}$ if $\rho_v = 1$, while one has $s_v = 1$ and $|\delta_0^{-1}(\omega \cdot \nu_{\ell_v})| \max_{j_1,j_2} |(\mathfrak{F}_v)_{j_1,j_2}| \le |\omega \cdot \nu_{\ell_v}|^{(i_{k_v} - i_0)\tau}$ if $\rho_v = 0$. As to the propagators one has $\|\mathfrak{G}_\ell\| \le \gamma_0^{-1} |\nu_\ell|^{\tau_0}$ for $\nu_\ell \neq 0$ and $\|\mathfrak{G}_\ell\| \le \|A^{-1}\|$ for $\nu_\ell = 0$, where $\|\cdot\|$ denotes – say – the uniform norm.

By iterating the graphical representation in Figure 3 until only graph elements with $k = 1$ appear, one finds

$$u_\nu^{(k)} = \sum_{\theta \in \mathcal{T}_{k,\nu}} \mathrm{Val}(\theta), \qquad k \ge 1. \tag{4.8}$$

The *tree expansion* (4.8) makes sense since all node factors and propagators are finite quantities, and the sum over the labels can be performed. Except the mode labels, the last assertion is trivial for all other labels (as they can assume only a finite number of values). Finally, the sum over the mode labels is controlled by the exponential decay $\mathrm{e}^{-\xi|\nu_v|}$ of the Fourier coefficients $\mathcal{F}_{\nu_v}$ of the node factors.

The study of the convergence of the perturbation series is made difficult by the product of propagators in (4.7). Indeed, the denominators $\delta_0(\omega \cdot \nu)$ can be arbitrarily close to zero for $\nu$ large enough. This problem is usually referred to as the *small divisor problem.*

## 4.4 Compatibility conditions

Now, we show that (4.1) holds for all the models considered in Section 3. Note that to prove (4.1) is not a purely technical problem: for the very models of Section 3, quasi-periodic solutions in the form of formal power series or even quasi-periodic solution *tout court* can fail to exist, if one weaken too much the assumptions – see also Section 12.2.

Let us start from the model in Section 3.1; recall that $d = p = n$ in such a case. We have already checked that (3.10) trivially holds for $k = 1$. Then we can prove by induction that for all $k \ge 2$ the compatibility condition (3.10) holds and one can set $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)} = \mathbf{0}$. The proof proceeds as follows. By using (4.8) we express $[F(u, \underline{\omega} t)]_0^{(k)}$ according to (2.6) as sum of trees in which all lines except the root line have nonzero momenta, by the inductive hypothesis. Given a tree $\theta$ consider together all trees which can be obtained from $\theta$ by detaching the root line and attaching it to any other node; see Figure 4.

In that way, we obtain as many trees as nodes of $\theta$; call $\mathcal{F}(\theta)$ the set of all such trees. Of course, all arrows must point toward the root, so that the trees $\theta' \in \mathcal{F}(\theta)$ have all the same mode labels (by construction), but they can have different momenta. On the other hand, since $\nu = 0$, a line $\ell \in L(\theta')$ either has the same momentum $\nu_\ell$ as in $\theta$ (if the arrow has not been reverted) or has momentum $-\nu_\ell$ (if the arrow has been reverted). Since $\delta_0(\nu_\ell) = \delta_0(-\nu_\ell)$, this means that the corresponding propagator $\mathfrak{G}_\ell$

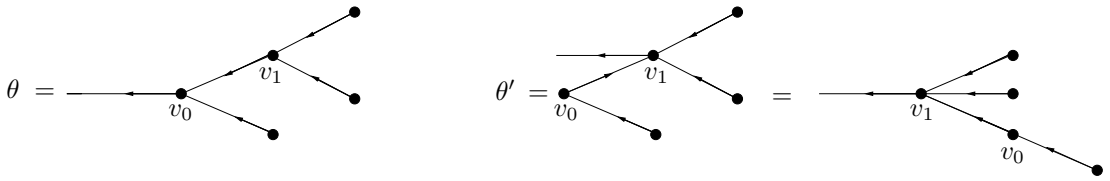


Figure 4: An example of tree $\theta'$ obtained from $\theta$ by detaching the root line from the node $v_0$ and reattaching it to the node $v_1$. The arrow of the line connecting the nodes $v_0$ and $v_1$ is reverted so as to point toward the new location of the root. One can always stretch the lines so as to make all arrows go right to left, as in the last graph.

does not change. The combinatorial factors of the trees $\theta' \in \mathcal{F}(\theta)$ are in general different from those of $\theta$ (because the values of $s_v$, $v \in N(\theta')$, can change), but if we sum together all nonequivalent trees we realise (with a little effort: one must perform the computation to convince himself that the assertion is true!) that we obtain a common value times a factor $\mathrm{i}\boldsymbol{\nu}_{v_0}$, if $v_0$ is the node which the root line is attached to. Therefore, since

$$\nu = \sum_{v_0 \in N(\theta)} \nu_{v_0} = 0, \tag{4.9}$$

the sum of all the tree values gives zero. This implies (4.1) with $\widetilde{n} = n$ and $\widehat{n} = 0$, and the coefficient $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)}$ is left arbitrary, and it can be chosen to be zero.

In fact, the argument above does not depend on the value of the coefficients $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)}$, which therefore can be arbitrarily chosen; in particular they can and will arbitrarily fixed to be zero. This was expected: changing $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)}$ means changing the constant $\boldsymbol{\alpha}_0$ in (3.4), which is arbitrary since it is the vector parameterising the torus.

The case of the standard map – see Section 3.2 – can be discussed in the same way. We omit the details.

In the case of the model in Section 3.3, the identity $[\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}, \boldsymbol{\beta})]_{\mathbf{0}}^{(k)} = \mathbf{0}$ for $k \geq 2$ can be proved as above, by relying on the same cancellation mechanism. The compatibility condition $[\partial_{\boldsymbol{\beta}} f(\boldsymbol{\alpha}, \boldsymbol{\beta})]_{\mathbf{0}}^{(k)} = \mathbf{0}$ for $k \geq 2$ can be imposed by using (3.25) and fixing $\boldsymbol{\beta}_{\mathbf{0}}^{(k)} = -(\partial_{\boldsymbol{\beta}}^2 f_{\mathbf{0}}(\boldsymbol{\beta}_0))^{-1} \boldsymbol{\Phi}_k$. This implies once more (4.1) with $\widetilde{n} = r$ and $\widehat{n} = s$. Again the coefficients $\boldsymbol{\alpha}_{\mathbf{0}}^{(k)}$ can be arbitrarily set to be zero.

Finally for the model in Section 3.4 one can use (3.31) to obtain (4.1) with $\widetilde{n} = 0$ and $\widehat{n} = n = 1$.

# 5 Multiscale analysis

To be able to bound the tree value (4.7), we need to control the product of propagators. This will be done through a *multiscale analysis*. To this aim, for each tree line we introduce a new label characterising the size of the corresponding propagator, that we call the *scale* label.

Essentially, we say that $\nu \in \mathbb{Z}^d \setminus \{0\}$ is on scale

$$\begin{cases} n \geq 1, & \text{if } 2^{-n}\gamma \leq |\omega \cdot \nu| < 2^{-(n-1)}\gamma, \\ n = 0, & \text{if } \gamma \leq |\omega \cdot \nu|, \end{cases} \tag{5.1}$$

where $\gamma$ is the constant appearing in (1.5), and we say that a line $\ell$ has a scale label $n_\ell = n$ if $\nu_\ell$ is on scale $n$.

As a matter of fact, in practice the sharp multiscale decomposition in (5.1) is a little annoying because, as we shall see, we have to consider derivatives. Thus, it is actually more convenient to replace it with a

smooth decomposition through $C^\infty$ compact support functions. Let $\psi$ be a nondecreasing $C^\infty$ function defined in $\mathbb{R}_+$, such that

$$\psi(x) = \begin{cases} 1, & \text{for } x \geq \gamma, \\ 0, & \text{for } x \leq \gamma/2, \end{cases} \tag{5.2}$$

and set $\chi(x) := 1 - \psi(x)$. For all $n \in \mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ define $\chi_n(x) := \chi(2^n x)$ and $\psi_n(x) := \psi(2^n x)$, and set

$$\Xi_n(x) = \chi_0(|x|) \dots \chi_{n-1}(|x|)\chi_n(|x|), \qquad \Psi_n(x) = \chi_0(|x|) \dots \chi_{n-1}(|x|)\psi_n(|x|), \tag{5.3}$$

where $\Psi_0(x)$ is meant as $\Psi_0(x) = \psi_0(|x|)$.

Then we change the definition of the propagator to be associated with each $\ell$ with $\nu_\ell \neq 0$, by associating with each such line $\ell$ a scale label $n_\ell \in \mathbb{Z}_+$ and a propagator

$$\mathfrak{G}_\ell = G^{[n_\ell]}(\omega \cdot \nu_\ell), \qquad G^{[n]}(\omega \cdot \nu) := \Psi_n(\omega \cdot \nu)\, \delta_0^{-1}(\omega \cdot \nu)\, \mathbb{1}, \tag{5.4}$$

which replaces the previous definition in (4.5). If $\Psi_n(x) \neq 0$ then $2^{-n-1}\gamma \leq |x| \leq 2^{-n+1}\gamma$, so that for any each $\ell \in L(\theta)$ one has either $\mathfrak{G}_\ell = 0$ or $\|\mathfrak{G}_\ell\| \leq \gamma_0^{-1} 2^{(n_\ell+1)i_0}$. For completeness we also associate a scale label $n_\ell = -1$ with each line $\ell$ with momentum $\nu_\ell = 0$. Note that, while with the sharp decomposition (5.1) a momentum $\nu$ identifies uniquely the scale $n$, on the contrary by using the smooth decomposition for each momentum $\nu$ there are two possible (adjacent) values $n$ such that $G^{[n]}(\omega \cdot \nu) \neq 0$.

A tree expansion like (4.8) still holds, with the difference that now the trees $\theta \in \mathcal{T}_{k,\nu}$ carry also the scale labels, and we have to sum also on these labels. The equality between the two expansions follows immediately from the observation that $\sum_{n=0}^\infty \Psi_n(x) = 1$ for all $x \in \mathbb{R} \setminus \{0\}$.

If $\mathfrak{N}_n(\theta)$ denotes the number of lines $\ell \in L(\theta)$ with scale $n_\ell = n$, then we can bound in (4.7)

$$\prod_{\ell \in L(\theta)} \|\mathfrak{G}_\ell\| \leq \gamma_0^{-k} 2^{k i_0} \prod_{n=0}^\infty 2^{n i_0 \mathfrak{N}_n(\theta)}, \tag{5.5}$$

with $i_0$ defined in (1.2), 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 $\theta$ 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 $n$. Given a cluster $T$ call $L(T)$ the set of lines of $\theta$ contained in $T$, and denote by $N(T)$ the set of nodes connected by such lines. We define $k(T) = \sum_{v \in N(T)} k_v$ 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; the terminology is borrowed from quantum field theory. This means that if $T$ is a self-energy cluster and $\ell_1$ and $\ell_2$ are the lines entering and exiting $T$, respectively, then $\nu_{\ell_1} = \nu_{\ell_2}$, so that

$$\sum_{v \in N(T)} \nu_v = 0. \tag{5.6}$$

By construction the scales of the lines $\ell_1$ and $\ell_2$ can differ at most by 1, and setting $n_T = \min\{n_{\ell_1}, n_{\ell_2}\}$, by definition of cluster one has $n_\ell < n_T$ for all $\ell \in L(T)$.

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

$$\mathcal{V}_T(\omega \cdot \nu) := \Big( \prod_{v \in N(T)} \mathfrak{F}_v \Big) \Big( \prod_{\ell \in L(T)} \mathfrak{G}_\ell \Big), \tag{5.7}$$

where all the indices of the node factors must be contracted except those associated with the line $\ell_1$ entering $T$ and with the line $\ell_2$ exiting $T$.

We can extend the notion of self-energy cluster 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 (5.7) has to be interpreted as $\mathcal{V}_T(\omega \cdot \nu) = \mathfrak{F}_v$: in particular it is independent of $\omega \cdot \nu$ if $\rho_v = 1$. If $T$ consists of only one node (and hence contains no line) we say that $T$ is a cluster on scale $-1$.

The simplest self-energy cluster one can think of consists of only one node $v$, but then (5.6) implies $\nu_v = 0$. For the model in Section 3.1, the corresponding value is zero, and hence the simplest nontrivial self-energy clusters contain at least two nodes. On the contrary for the models in Sections 3.3 and 3.4 one can also have clusters with only one node.

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

$$\mathfrak{N}^*_n(\theta) \le c\, 2^{-n/\tau} K(\theta), \qquad K(\theta) := \sum_{v \in N(\theta)} |\nu_v|, \tag{5.8}$$

for some constant $c$, where $\tau$ is the Diophantine exponent in (1.5). See Appendix A for a proof.

If no self-energy clusters could occur (so that $\mathfrak{R}_n(\theta) = 0$) the Siegel-Bryuno lemma would allow us to bound in (5.5)

$$\prod_{n=0}^{\infty} 2^{ni_0 \mathfrak{N}_n(\theta)} = \prod_{n=0}^{\infty} 2^{ni_0 \mathfrak{N}^*_n(\theta)} \le 2^{n_0 i_0 k} \prod_{n=n_0+1}^{\infty} 2^{ni_0 \mathfrak{N}^*_n(\theta)} \le C_1^k \exp\Big(\xi(n_0) K(\theta)\Big), \tag{5.9}$$

with $C_1 = 2^{n_0 i_0}$ and

$$\xi(n_0) := i_0 c \sum_{n=n_0+1}^{\infty} n 2^{-n/\tau}. \tag{5.10}$$

Since $\xi(n_0) \to 0$ as $n_0 \to \infty$ one can fix $n_0$ in such a way that $\xi(n_0) \le \xi/4$ (the constant $\xi$ being defined after (2.6)). Then, by extracting a factor $\mathrm{e}^{-\xi|\nu_v|/2}$ from each node factor $\mathfrak{F}_v$, one could easily perform the sum over the Fourier labels, so as to obtain an overall bound $C_2^k \mathrm{e}^{-\xi|\nu|/2}$ on $u^{(k)}_\nu$ for a suitable constant $C_2$. This would imply the convergence of the perturbation series (2.1) for $\varepsilon$ small enough, say for $|\varepsilon| < \varepsilon_0$ for some $\varepsilon_0 > 0$. However, there are self-energy clusters and they produce factorials, as the example in Appendix A shows, so that we have to deal with them.

# 6 Resummation of the series

Let us come back to the equation (2.11). To simplify the analysis, let us initially assume that we are using the sharp multiscale decomposition (5.1), so that each momentum fixes uniquely the corresponding scale. If we take the tree expansion of the right hand side of (2.11), 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

$$\sum_{T : k(T) < k} \mathcal{V}_T(\omega \cdot \nu)\, u^{(k-k(T))}_\nu, \tag{6.1}$$

and all the other contributions, that we denote by $[F(u, \underline{\omega} t)]^{(k-1)*}_\nu$. In (6.1) both the entering and exiting lines of $T$ have the same scale $n_T$, and the sum is over all clusters $T$ on scale $< n_T$.

By writing (2.11) as

$$\delta_0(\omega \cdot \nu)\, u^{(k)}_\nu = - \sum_{p=1}^{\min\{k, k_0\}} \delta_p(\omega \cdot \nu) u^{(k-p)}_\nu + [F(u, \underline{\omega} t)]^{(k-1)}_\nu \tag{6.2}$$

we can shift the contributions (6.1) to the left hand side of (6.2), so as to obtain

$$\delta_0(\omega\cdot\nu)\,u_\nu^{(k)} - \sum_{T:k(T)<k} \mathcal{V}_T(\omega\cdot\nu)\,u_\nu^{(k-k(T))} = [F(u,\underline{\omega}t)]_\nu^{(k-1)*}\,. \tag{6.3}$$

By summing over $k$ and setting

$$M(\omega\cdot\nu,\varepsilon) := \sum_{k=1}^{\infty}\varepsilon^k \sum_{T:k(T)=k} \mathcal{V}_T(\omega\cdot\nu), \tag{6.4}$$

then (6.3) gives, formally,

$$\mathcal{D}(\omega\cdot\nu,\varepsilon)\,u_\nu = [F(u,\underline{\omega}t)]_\nu^*\,. \qquad \mathcal{D}(\omega\cdot\nu,\varepsilon) := \delta_0(\omega\cdot\nu)\,\mathbb{1} - M(\omega\cdot\nu,\varepsilon). \tag{6.5}$$

The motivation for proceeding in this way is that, at the price of changing $\delta_0(\omega\cdot\nu)$ into $\mathcal{D}(\omega\cdot\nu,\varepsilon)$, 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 $\nu$ which are on scale $n=0$. Since the only self-energy clusters $T$ with $n_T=0$ are those (on scale $-1$) containing only one node, for such $\nu$ the matrix $M(\omega\cdot\nu,\varepsilon)$ is just a constant. Next, we pass to the momenta $\nu$ which are on scale $n=1$, and we consider (6.5) for such $\nu$; now all self-energy clusters $T$ whose values contribute to $M(\omega\cdot\nu,\varepsilon)$ cannot contain any self-energy clusters, because the lines $\ell\in L(T)$ are on scale $n_\ell=0$. Then, we consider the momenta $\nu$ 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 $u(\omega t,\varepsilon)$, that we call a *resummed series*,

$$u(\omega t,\varepsilon) = \sum_{\nu\in\mathbb{Z}^d} \mathrm{e}^{\mathrm{i}\omega\cdot\nu t}u_\nu, \qquad u_\nu = \sum_{k=1}^{\infty}\varepsilon^k u_\nu^{[k]}(\varepsilon), \tag{6.6}$$

where the self-energy clusters do not appear any more in the tree expansion and the propagators must be defined recursively, as follows. The propagator $\mathfrak{G}_\ell$ of a line $\ell$ on scale $n_\ell=n$ and momentum $\nu_\ell=\nu$ is the matrix

$$\mathfrak{G}_\ell := G^{[n]}(\omega\cdot\nu,\varepsilon) = \Big(\delta_0(\omega\cdot\nu)\,\mathbb{1} - \mathcal{M}^{[n-1]}(\omega\cdot\nu,\varepsilon)\Big)^{-1}, \tag{6.7}$$

with

$$\mathcal{M}^{[n]}(\omega\cdot\nu,\varepsilon) := \sum_{T\,\text{on scale}\,\le n} \varepsilon^{k(T)}\,\mathcal{V}_T(\omega\cdot\nu), \tag{6.8}$$

where the value $\mathcal{V}_T(\omega\cdot\nu)$ is written in accord with (5.7), with all the lines $\ell'\in L(T)$ on scales $n_{\ell'}<n$ and the corresponding propagators $\mathfrak{G}_{\ell'}$ expressed in terms of matrices $\mathcal{M}^{[n_{\ell'}]}(\omega\cdot\nu_{\ell'},\varepsilon)$ as in (6.7).

By construction, the new propagators depend on $\varepsilon$, so that the coefficients $u_\nu^{[k]}(\varepsilon)$ depend explicitly on $\varepsilon$: hence (6.6) is not a power series expansion.

If we use the smooth multiscale decomposition, then the algorithm above must be suitably modified. We define recursively the propagators $\mathfrak{G}_\ell = G^{[n_\ell]}(\omega \cdot \nu_\ell, \varepsilon)$ by setting for $n \geq 0$

$$G^{[n]}(\omega \cdot \nu, \varepsilon) = \Psi_n(\omega \cdot \nu) \Big( \delta_0(\omega \cdot \nu) \mathbb{1} - \mathcal{M}^{[n-1]}(\omega \cdot \nu, \varepsilon) \Big)^{-1} \tag{6.9a}$$

$$\mathcal{M}^{[n]}(x, \varepsilon) = \mathcal{M}^{[n-1]}(x, \varepsilon) + \Xi_n(x)\, M^{[n]}(x, \varepsilon), \qquad M^{[n]}(x, \varepsilon) = \sum_{T \in \mathcal{R}_n} \varepsilon^{k(T)}\, \mathcal{V}_T(x), \tag{6.9b}$$

where $\mathcal{R}_n$ is the set of self-energy clusters on scale $n$ which do not contain any other self-energy clusters.

The matrices $\mathcal{M}^{[n]}(\omega \cdot \nu, \varepsilon)$ are called the *self-energies*. The new propagators (6.9a) are called, by exploiting once more the analogy with quantum field theory, the *dressed propagators*.

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

$$u_\nu^{[k]}(\varepsilon) = \sum_{\theta \in \mathcal{T}^{\mathcal{R}}_{k,\nu}} \mathrm{Val}(\theta), \qquad \mathrm{Val}(\theta) := \Big( \prod_{v \in N(\theta)} \mathfrak{F}_v \Big) \Big( \prod_{\ell \in L(\theta)} \mathfrak{G}_\ell \Big) \qquad \nu \neq 0, \quad k \geq 1, \tag{6.10}$$

which replaces (4.8). In particular $\mathcal{T}^{\mathcal{R}}_{k,\nu}$ is defined as the set of *renormalised trees* of order $k$ and momentum $\nu$, where "renormalised" means that the trees do not contain any self-energy clusters.

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 (5.5) and (5.9), provided the propagators on scale $n$ can still be bounded proportionally to $2^{ni_0}$. In general there is no reason why this should occur, because of the extra term $\mathcal{M}^{[n-1]}(\omega \cdot \nu, \varepsilon)$ appearing in (6.9a).

The discussion of such an issue depends on the particular model one is studying. We shall see in the next sections what happens for the models considered in Section 3. We shall first consider cases in which the dressed propagators can be essentially bounded as the old ones, and then cases in which this is no longer true. By modifying further the resummation procedure described above, we shall see that something can still be achieved also in these cases.

# 7 Cancellations and convergence of the series – maximal tori

Let us consider the matrix $\mathcal{M}^{[n]}(x, \varepsilon)$ introduced in (6.9b), and let us study its dependence on the first argument $x = \omega \cdot \nu$ for the models of Section 3. As usual, let us consider first the model in Section 3.1.

It is a remarkable cancellation that $\mathcal{M}^{[n]}(x, \varepsilon)$ vanishes in $x$ up to second order, that is $\mathcal{M}^{[n]}(x, \varepsilon) = O(x^2)$. The symmetry properties

$$\mathcal{M}^{[n]}(x, \varepsilon) = (\mathcal{M}^{[n]}(-x, \varepsilon))^T = (\mathcal{M}^{[n]}(x, \varepsilon))^\dagger, \tag{7.1}$$

with $T$ and $\dagger$ denoting transposition and adjointness, are essential for the proof (such properties are trivially satisfied for $n = 0$, and can be proved by induction on $n$ – see [47, 34] for more details). Indeed, by using (7.1), the cancellation $\mathcal{M}^{[n]}(0, \varepsilon) = 0$ can be proved as the cancellation $[F(u, \underline{\omega} t)]_0^{(k)} = [\partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha})]_{\mathbf{0}}^{(k)} = 0$ discussed in Section 4.4, with the exiting line of the self-energy clusters playing the role of the root line. The first order cancellation requires $\partial_x \mathcal{M}^{[n]}(0, \varepsilon) = 0$, and this can be proved through a similar cancellation mechanism: besides the exiting line one has to detach also the entering line and reattach it to all the other nodes inside the self-energy clusters. If the function $f$ in (3.1) is even in $\boldsymbol{\alpha}$, then the first order cancellation follows also from parity properties [32].

Both the cancellations and the symmetry properties are only formal as far as we have not proved that the self-energies are well-defined quantities. To this aim we need to control the product of propagators in (5.7), with the propagators defined according to (6.9). An important ingredient of the analysis is that

also for all self-energy clusters one can prove a bound like (5.8). More precisely, if we denote by $\mathfrak{N}_n(T)$ the number of lines $\ell \in L(T)$ on scale $n$, for all $n \geq 0$ and all $T \in \mathcal{R}_n$ one has

$$\mathfrak{N}^*_{n'}(T) \leq c' \, 2^{-n'/\tau} K(T), \qquad K(T) := \sum_{v \in N(T)} |\nu_v|, \qquad n' \leq n, \tag{7.2}$$

for some constant $c'$. To prove (7.2) one first show that, for all $n \geq 0$ and all $T \in \mathcal{R}_n$, one has

$$\sum_{v \in N(T)} |\nu_v| > c'' 2^{n/\tau}, \tag{7.3}$$

for some constant $c''$, then one proceeds by induction on the order $k(T)$; see Appendix B for details.

Therefore, if we were able to prove for the dressed propagators an estimate like $\|G^{[n]}(x,\varepsilon)\| \leq 2/x^2$, then we could use (7.2) and (7.3) to bound

$$\prod_{\ell \in L(T)} \|\mathfrak{G}_\ell\| \leq \gamma_0^{-k} 2^{3k} \prod_{n'=0}^{n} 2^{2n' \mathfrak{N}^*_{n'}(T)}, \tag{7.4}$$

in such a way to obtain

$$|\mathcal{V}_T(\omega \cdot \nu)| \leq C_2^{k(T)} \mathrm{e}^{-\xi K(T)/2}, \tag{7.5}$$

for a suitable constant $C_2$, independent of $T$. In particular, this would ensure the well-definedness of the self-energies. At this point, the cancellations would allow us to write, for some constant $C$ and for all $n \geq 0$,

$$\mathcal{M}^{[n]}(x,\varepsilon) = \varepsilon^2 x^2 \overline{\mathcal{M}}^{[n]}(x,\varepsilon), \qquad \left\| \overline{\mathcal{M}}^{[n]}(x,\varepsilon) \right\| \leq C, \tag{7.6}$$

where we have taken into account also that $\mathcal{V}_T(x) \neq 0$ requires $k(T) \geq 2$ (cf. Section 5). In turn (7.6) would implies that, for $\varepsilon$ small enough,

$$\left\| G^{[n]}(x,\varepsilon) \right\| = \left\| \left( x^2 - \varepsilon^2 x^2 \overline{\mathcal{M}}^{[n-1]}(x,\varepsilon) \right)^{-1} \right\| \leq \frac{2}{x^2}. \tag{7.7}$$

The bounds (7.7) on the propagators, the symmetry properties, and the cancellations are proved all together, as follows. First note that, if the second order cancellation holds, one can write

$$\varepsilon^2 \overline{\mathcal{M}}^{[n]}(x,\varepsilon) = \int_0^1 \mathrm{d}t \; (1-t) \, \partial_x^2 \mathcal{M}^{[n]}(tx,\varepsilon), \tag{7.8}$$

so that the bound in (7.6) is essentially a bound on the second derivative of the self-energies. The case $n = 0$ is easily checked. Then the proof proceeds by induction, by relying on the recursive definition of $\mathcal{M}^{[n]}(x,\varepsilon)$ – see (6.9) and (5.7) –, and taking advantage of the smooth multiscale decomposition to perform the derivatives. More precisely, we assume that both the cancellations – and hence the bounds (7.6) – and the symmetry properties (7.1) hold for all $n' < n$. This means that all the dressed propagators of the lines on scales $\leq n$ are bounded according to (7.7), so that we can use the bound (7.5) to prove that also $\mathcal{M}^{[n]}(x,\varepsilon)$ is well-defined. Then the cancellation mechanism described at the beginning of the section shows that also at the step $n$ the symmetry properties (7.1) and the cancellations leading to (7.6) are satisfied; in particular also the propagators $G^{[n+1]}(x,\varepsilon)$ are bounded proportionally to $|x|^2$ according to (7.7).

The conclusion is that the series in (6.6) for $u_\nu = \boldsymbol{\alpha}_{\boldsymbol{\nu}}$ converges for $\varepsilon$ small enough. Therefore, the function $u(\omega t, \varepsilon) = \boldsymbol{\alpha}(\boldsymbol{\omega} t, \varepsilon)$ is analytic in $\varepsilon$ (notwithstanding that the expansion in $\varepsilon$ is not a power expansion), so that we can say *a posteriori* that the original power series (3.14) also converges. It is straightforward to see that $\boldsymbol{\alpha}_{\boldsymbol{\nu}}$ decays exponentially in $\boldsymbol{\nu}$, which implies that the function $\boldsymbol{\alpha}(\boldsymbol{\psi}, \varepsilon)$ is also analytic in $\boldsymbol{\psi}$.

The case of the standard map – see Section 3.2 – can be discussed in the same way. We do not repeat the analysis and refer to [37, 10] for details.

# 8 Summation of the divergent series – hyperbolic tori

Now we consider the model introduced in Section 3.3. In that case, one has

$$\mathcal{M}^{[n]}(x,\varepsilon) = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0) \end{pmatrix} + O(\varepsilon^2), \tag{8.1}$$

so that, already only keeping the first order terms, one realises that a cancellation like (7.6) cannot expected to hold. Indeed, in order to study the convergence of the series (6.6), we need at least the perturbation series (2.1) to be formally well-defined to all orders; in turn this requires matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ to be nonsingular – see Sections 3.4 and 4.4 – and hence different from 0.

Let us assume first that the matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ is positive definite, that is that its eigenvalues $a_1, \dots, a_s$ are positive, i.e $a_i > 0$ for $i = 1, \dots, s$ (in particular this means that $\boldsymbol{\beta}_0$ is a maximum point for the function $f_{\mathbf{0}}(\boldsymbol{\beta})$).

We write both the self-energies and the propagators as block matrices,

$$\mathcal{M}^{[n]}(x,\varepsilon) = \begin{pmatrix} \mathcal{M}^{[n]}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(x,\varepsilon) & \mathcal{M}^{[n]}_{\boldsymbol{\alpha}\boldsymbol{\beta}}(x,\varepsilon) \\ \mathcal{M}^{[n]}_{\boldsymbol{\beta}\boldsymbol{\alpha}}(x,\varepsilon) & \mathcal{M}^{[n]}_{\boldsymbol{\beta}\boldsymbol{\beta}}(x,\varepsilon) \end{pmatrix}, \qquad G^{[n]}(x,\varepsilon) = \begin{pmatrix} G^{[n]}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(x,\varepsilon) & G^{[n]}_{\boldsymbol{\alpha}\boldsymbol{\beta}}(x,\varepsilon) \\ G^{[n]}_{\boldsymbol{\beta}\boldsymbol{\alpha}}(x,\varepsilon) & G^{[n]}_{\boldsymbol{\beta}\boldsymbol{\beta}}(x,\varepsilon) \end{pmatrix}, \tag{8.2}$$

where the four blocks are $r \times r$, $r \times s$, $s \times r$, and $s \times s$ matrices, respectively.

Then, one can prove that the parity properties (7.1) still hold, and moreover, formally, one has the cancellations

$$\mathcal{M}^{[n]}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(x,\varepsilon) = O(\varepsilon^2 x^2), \qquad \mathcal{M}^{[n]}_{\boldsymbol{\alpha}\boldsymbol{\beta}}(x,\varepsilon) = \mathcal{M}^{[n]}_{\boldsymbol{\beta}\boldsymbol{\alpha}}(x,\varepsilon) = O(\varepsilon^2 x). \tag{8.3}$$

The proof of such assertions can be performed by induction, and follows the same pattern as described in Section 7 – see [35, 34] for details. We have used the word "formally" because the cancellations hold as far as the dressed propagators (6.9a) can be bounded essentially as the old ones (5.4) – a property that we have not yet proved.

The main implication of (7.1) and (8.3) is that the eigenvalues $\lambda^{[n]}_i(x,\varepsilon)$ of the self-energies $\mathcal{M}^{[n]}(x,\varepsilon)$ are of the form

$$\lambda^{[n]}_i(x,\varepsilon) = \begin{cases} O(\varepsilon^2 x^2), & i = 1, \dots, r, \\ a_{i-r}\varepsilon + O(\varepsilon^2), & i = r+1, \dots, d. \end{cases} \tag{8.4}$$

In particular, if $\varepsilon < 0$, the eigenvalues $x^2 - \lambda^{[n]}_i(x,\varepsilon)$ of the matrices $\delta_0(\omega\cdot\nu)\,\mathbb{1} - \mathcal{M}^{[n]}(x,\varepsilon)$ are such that $x^2 - \lambda^{[n]}_i(x,\varepsilon) \geq x^2/2$ for $i = 1, \dots, r$ and $x^2 - \lambda^{[n]}_i(x,\varepsilon) \geq x^2 + |\varepsilon a_{i-r}|/2$ for $i = r+1, \dots, d$, provided $\varepsilon$ is small enough and the higher order corrections in (8.4) remain small. The last property is automatically satisfied if the block matrices in (8.3) are dominated by the first nontrivial orders.

All the properties described above become rigorous if the dressed propagators $G^{[n]}(x,\varepsilon)$ are bounded proportionally to $x^{-2}$ (more generally any power of $|x|$ would suit), say

$$\left\| G^{[n]}(x,\varepsilon) \right\| \leq \frac{2}{x^2}. \tag{8.5}$$

So, all we have to do is to prove together all the above properties (7.1) and (8.3), by induction. Indeed, for $n = 0$ the properties are trivially satisfied, and at any step $n$, by the inductive hypothesis, the bounds $\|G^{[n']}(x,\varepsilon)\| \leq 2/x^2$ are satisfied for all $n' \leq n$, so that both (7.1) and (8.3) hold for $n$, and in turn this implies (8.4) and hence the bound (8.5) for $n+1$.

Therefore, the series (6.6) converges for $\varepsilon$ small enough, even if analyticity is prevented because of the condition $\varepsilon < 0$; then we say that the perturbation series (2.1) is a *divergent series* (in point of fact, it is very likely that it does not converge, though there is no proof of that). The function is only $C^\infty$ in $\varepsilon$ at $\varepsilon = 0$ – such a result improves a previous one by Treshchëv [79], where $C^\infty$-smoothness in $\sqrt{\varepsilon}$ was

proved at $\varepsilon = 0$. In fact, one can say a little more about the dependence of the invariant torus on the perturbation parameter $\varepsilon$ in the complex $\varepsilon$-plane: the lower-dimensional torus turns out to be analytic in the heart-shaped domain of Figure 5 [35].

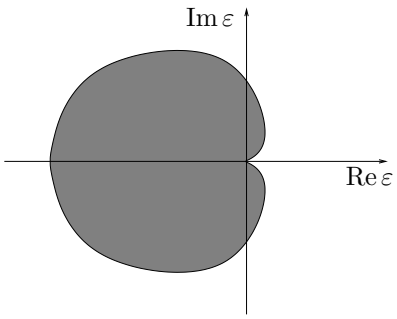


Figure 5: Analyticity domain for the hyperbolic invariant torus.

As said in Section 3.3 the torus will be said hyperbolic in that case. This is a somewhat improper terminology. Indeed, when studying lower-dimensional tori one usually considers Hamiltonians of the form [53, 67, 28, 73]

$$\mathcal{H}(\boldsymbol{\alpha}, \boldsymbol{q}, \boldsymbol{A}, \boldsymbol{p}) = \frac{1}{2}\boldsymbol{A}^2 + \frac{1}{2}\boldsymbol{p}^2 + \sum_{i=1}^{s} \lambda_i q_i^2 + \varepsilon f(\boldsymbol{\alpha}, \boldsymbol{q}), \tag{8.6}$$

(or generalisations of its), with $(\boldsymbol{\alpha}, \boldsymbol{A}) \in \mathbb{T}^r \times \mathbb{R}^r$ and $(\boldsymbol{q}, \boldsymbol{p}) \in \mathbb{R}^s \times \mathbb{R}^s$. Thus, for $\varepsilon = 0$ the coordinates $\boldsymbol{\alpha}$ freely rotates with some rotation vector $\boldsymbol{\omega} = (\omega_1, \dots, \omega_r)$, while the coordinates $(\boldsymbol{q}, \boldsymbol{p})$ moves around an equilibrium point, which is elliptic if $\lambda_i > 0$ for all $i = 1, \dots, s$ and hyperbolic if $\lambda_i < 0$ for all $i = 1, \dots, s$. The numbers $\omega_1, \dots, \omega_r$ are called the *proper frequencies*, while the numbers $\lambda_1, \dots, \lambda_s$ are called the *normal frequencies*.

Then one can study the problem of persistence of lower-dimensional tori under perturbation, that is for $\varepsilon \neq 0$. The Hamiltonian (3.18) can be interpreted as a Hamiltonian of the form (8.6) with $\lambda_i = O(\varepsilon)$. In that case one usually say that the lower-dimensional tori are *parabolic*. However, we can also interpret the persisting tori, in the case $a_i > 0$ and $\varepsilon < 0$, as *degenerate hyperbolic tori*, where "degenerate" refers to the fact that the normal frequencies vanish at $\varepsilon = 0$.

Of course if the matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ is negative definite, the same result of persistence of hyperbolic invariant tori holds for $\varepsilon > 0$. The case of indefinite (i.e. neither positive nor negative defined) matrices will be considered at the end of Section 10.

# 9 Summation of the divergent series – dissipative systems

The discussion of the model (3.26) introduced in Section 3.4 proceeds very closely to the case of the hyperbolic tori of Section 8. In that case $n = 1$, hence both the propagators and the self-energies are scalar. One has, formally,

$$\mathcal{M}^{[n]}(x, \varepsilon) = -(\mathrm{i}\varepsilon x)^2 + \widetilde{\mathcal{M}}^{[n]}(x, \varepsilon), \qquad \widetilde{\mathcal{M}}^{[n]}(0, \varepsilon) = a\varepsilon + O(\varepsilon^2), \qquad a := \partial_x g(c_0) \neq 0. \tag{9.1}$$

If for all $n' < n$ the dressed propagators $G^{[n']}(\omega \cdot \nu, \varepsilon)$ can still be bounded proportionally to $|x|^{-1}$ (as the undressed ones), the terms $O(\varepsilon^2)$ are defined by a convergent series, so that

$$\widetilde{\mathcal{M}}^{[n]}(x, \varepsilon) = \widetilde{\mathcal{M}}^{[n]}(0, \varepsilon) + x \int_0^1 \mathrm{d}t\, \partial_x \widetilde{\mathcal{M}}^{[n]}(tx, \varepsilon), \qquad \partial_x \widetilde{\mathcal{M}}^{[n]}(x, \varepsilon) = O(\varepsilon^2), \tag{9.2}$$

and hence

$$\delta_0(x) - \mathcal{M}^{[n-1]}(x,\varepsilon) = \mathrm{i}x - \mathcal{M}^{[n-1]}(x,\varepsilon) = \mathrm{i}x\,(1+\mathrm{i}\varepsilon x) - a\varepsilon - O(\varepsilon^2) - O(\varepsilon^2 x), \qquad x = \omega\cdot\nu, \tag{9.3}$$

and hence

$$\left|G^{[n]}(x,\varepsilon)\right| \le \left|\delta_0(x) - \mathcal{M}^{[n-1]}(x,\varepsilon)\right| \le \frac{2}{|x|}, \tag{9.4}$$

for $\varepsilon$ small enough. Actually one can prove that $\widetilde{\mathcal{M}}^{[n]}(0,\varepsilon)$ is real for real $\varepsilon$ [41], a property which becomes essential to deal with the case in which the nondegeneracy condition $a \neq 0$ is not satisfied – see Section 12.2 below.

Again the properties (9.1) and (9.4) are proved together, by induction on $n$: for $n = 0$ they trivially hold, and, by assuming that they are satisfied up to $n-1$, one sees that the series defining $\widetilde{\mathcal{M}}^{[n]}(x,\varepsilon)$ converge, and hence (9.1) can be proved for $n$; see [42, 40, 41] for details.

One can study the dependence of the response solution on $\varepsilon$ in the complex domain. Of course one expects an obstruction to analyticity along the imaginary axis (so as it happened along the positive real axis for the hyperbolic tori). In fact one can prove that the response solutions are analytic in two disks tangent to the imaginary axis at the origin – see Figure 6 –; of course only the disk to the right is physically relevant, as it corresponds to $\varepsilon > 0$ and hence to $\gamma > 0$.

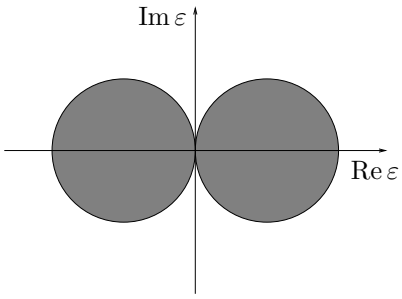


Figure 6: Analyticity domain for the response solution.

An interesting remark is that, while the periodic case ($d = p = 1$) is trivial in the case of lower-dimensional tori, this is no longer true for the model (3.26). Indeed, if one takes $r = 1$ in Section 3.3, then no small divisors appear, so that the perturbation series is easily proved to be convergent (in particular analyticity in $\varepsilon$ follows in that case). On the contrary if one takes $d = m = 1$ for the model (3.26), one can still have arbitrarily large powers of $|\nu|$ because $\delta_1(\omega\cdot\nu)/\delta_0(\omega\cdot\nu) = \mathrm{i}\omega\cdot\nu$: then, it is straightforward to see that one can construct trees whose value grows like a factorial – see Appendix C.

This means that also in the case of periodic forcing, the response solution to (3.26) is not analytic in $\varepsilon$. However, one can prove that such a solution is *Borel summable* [42, 43]; we recall in Appendix D the definition of Borel summability – see also [14, 58, 78]. Note that an equation like (3.26) with periodic forcing (and with $g(x) = x^\mu$ for $\mu \in [1.5, 2.5]$) naturally arises in electronic engineering, and is known as the *varactor equation* [7].

It is proved in [43] that also in the case of quasi-periodic forcing, the response solution turns out to be Borel summable if $d = 2$ and $\tau = 1$ – that is in the case of frequency vectors with components whose ratio is an irrational number of constant type [77]. A similar situation is encountered in the case of hyperbolic tori: also in that case the function $u(\boldsymbol{\omega}t,\varepsilon)$ is Borel summable if $d = 2$ and $\tau = 1$ [25].

# 10 Cantorisation – elliptic tori

Let us come back to the system of Section 3.3, still assuming that the eigenvalues $a_1, \dots, a_s$ of the matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ are all positive, but taking $\varepsilon > 0$. In that cases, already for $n = -1$ one has – see (6.8) and (8.1) –

$$\mathcal{M}^{[-1]}(x,\varepsilon) = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon \partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0) \end{pmatrix}, \tag{10.1}$$

so that the eigenvalues $x^2 - \lambda_i^{[-1]}(x,\varepsilon)$ of the matrix $\delta_0(x)\,\mathbb{1} - \mathcal{M}^{[-1]}(x,\varepsilon)$ are

$$x^2 - \lambda_i^{[-1]}(x,\varepsilon) = \begin{cases} x^2, & i = 1, \dots, r, \\ x^2 - \varepsilon a_{i-r}, & i = r+1, \dots, d. \end{cases} \tag{10.2}$$

Hence, for fixed $\varepsilon$, we have problems for all $\nu \in \mathbb{Z}^r$ such that $\omega \cdot \nu$ is too close to some value $\pm\sqrt{a_i \varepsilon}$. So, to give a meaning to $G^{[0]}(x,\varepsilon)$ we must require some further Diophantine conditions, say

$$|\,|\omega \cdot \nu| - \sqrt{\varepsilon a_i}\,| > \gamma |\nu|^{-\tau'} \quad \forall \nu \in \mathbb{Z}^r \setminus \{0\}, \tag{10.3}$$

for some Diophantine exponent $\tau' \geq d$. This can be achieved at the price of eliminating some values of $\varepsilon$. Fixed $\varepsilon_0 > 0$ small enough, the subset $\mathfrak{E}'_{-1}$ of values $\varepsilon \in [0, \varepsilon_0]$ for which all Diophantine conditions (10.3) are satisfied has large Lebesgue measure in $[0, \varepsilon_0]$, in the sense that

$$\lim_{\varepsilon \to 0^+} \frac{\mathrm{meas}(\mathfrak{E}'_{-1} \cap [0,\varepsilon])}{\varepsilon} = 1, \tag{10.4}$$

provided $\tau'$ is chosen larger enough than $\tau$, say $\tau' > \tau + r$; see Appendix E. The property (10.4) can be stated by saying that $\mathfrak{E}'_{-1}$ has a Lebesgue density point at $\varepsilon = 0$.

For all values $\varepsilon \in \mathfrak{E}'_{-1}$ the propagators $G^{[0]}(x,\varepsilon) = \Psi_0(x)\,(\delta_0(x)\,\mathbb{1} - \mathcal{M}^{[-1]}(x,\varepsilon))^{-1}$ can be formally defined. At this point, one could hope to iterate the procedure. The main obstacle is that now the dressed propagators are no longer bounded proportionally to the undressed one: indeed it may happen that $x^2 - \varepsilon a_{i-r}$ is much smaller than $x^2$. So we have to modify the algorithm.

For simplicity, let us first reason once more by taking a sharp decomposition as initially done in Section 5. Let us also assume, in the discussion below, the self-energies to be well-defined: we shall back later to such an issue.

We say that $\nu \neq 0$ is on scale 0 if $|\omega \cdot \nu| \geq \gamma$ and on scale $[\geq 1]$ otherwise: for $\nu$ on scale 0 we define $G^{[0]}(\omega \cdot \nu, \varepsilon)$ as in (6.7), with $n = 0$ and $\mathcal{M}^{[-1]}(x,\varepsilon)$ given in (10.1). Given $\nu$ on scale $[\geq 1]$ we say that $\nu$ is on scale 1 if $\min_{i=1,\dots,d} |(\omega \cdot \nu)^2 - \lambda_i^{[-1]}(\omega \cdot \nu, \varepsilon)| \geq (2^{-1}\gamma)^2$, and on scale $[\geq 2]$ if $\min_{i=1,\dots,d} |(\omega \cdot \nu)^2 - \lambda_i^{[-1]}(\omega \cdot \nu, \varepsilon)| < (2^{-1}\gamma)^2$. For $\nu$ on scale 1 we write $G^{[1]}(x,\varepsilon)$ as in (6.7), with $n = 1$ and $\mathcal{M}^{[0]}(x,\varepsilon)$ written according to (6.8). Call $\lambda_i^{[0]}(x,\varepsilon)$ the eigenvalues of $\mathcal{M}^{[0]}(x,\varepsilon)$: given $\nu$ on scale $[\geq 2]$ we say that $\nu$ is on scale 2 if $\min_{i=1,\dots,d} |(\omega \cdot \nu)^2 - \lambda_i^{[0]}(\omega \cdot \nu, \varepsilon)| \geq (2^{-2}\gamma)^2$, and on scale $[\geq 3]$ if $\min_{i=1,\dots,d} |(\omega \cdot \nu)^2 - \lambda_i^{[0]}(\omega \cdot \nu, \varepsilon)| < (2^{-2}\gamma)^2$. And so on: eventually we impose infinitely many Diophantine conditions, i.e.

$$\left| |\omega \cdot \nu| - \sqrt{|\lambda_i^{[n]}(\omega \cdot \nu, \varepsilon)|} \right| > 2^{-(n+1)/2} \gamma |\nu|^{-\tau'} \quad \forall \nu \in \mathbb{Z}^r \setminus \{0\} \tag{10.5}$$

for all $i = 1, \dots, d$ and $n \geq -1$.

Even if we are successful in imposing the conditions (10.5), the argument above is still incomplete, and it needs a further modification. In order to bound the tree values and the self-energies we need a bound on the number of lines of fixed scale, in the spirit of the Siegel-Bryuno lemma. This requires to

compare the propagators of the lines entering and exiting clusters which are not self-energy clusters – see the discussion at the end of Appendix F – and this leads to further Diophantine conditions,

$$\left|\,|\omega\cdot(\nu_1-\nu_2)|\pm\sqrt{|\lambda_i^{[n]}(\omega\cdot\nu_1,\varepsilon)|}\pm\sqrt{|\lambda_j^{[n]}(\omega\cdot\nu_2,\varepsilon)|}\,\right|>2^{-(n+1)/2}\gamma|\nu_1-\nu_2|^{-\tau'}\quad\forall\nu_1\neq\nu_2\in\mathbb{Z}^r\setminus\{0\},\ (10.6)$$

for all $i,j=1,\dots,d$ and $n\geq -1$. For any fixed $\nu$ this would mean to impose the conditions for all $\nu_1$ and $\nu_2$ such that $\nu_1-\nu_2=\nu$. Unfortunately, these conditions are infinitely many, and for all of them we would eliminate intervals of the same size: therefore we would left with a zero measure set. Of course, at the step $n=-1$, there would be no difficulty, since the eigenvalues $\lambda_i^{[-1]}(x,\varepsilon)$ are independent of $x$ – see (10.2) –, but already at the step $n=0$ problems would arise.

So, instead of (10.5) and (10.6), we can try to impose the Diophantine conditions

$$\left|\,|\omega\cdot\nu|-\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon)}\,\right|>2^{-(n+1)/2}\gamma|\nu|^{-\tau'}\quad\forall\nu\in\mathbb{Z}^r\setminus\{0\},\tag{10.7a}$$

$$\left|\,|\omega\cdot\nu|\pm\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon)}\pm\sqrt{\underline{\lambda}_j^{[n]}(\varepsilon)}\,\right|>2^{(n+1)/2}\gamma|\nu|^{-\tau'}\quad\forall\nu\in\mathbb{Z}^r\setminus\{0\},\tag{10.7b}$$

for all $i,j=1,\dots,d$ and $n\geq 0$, for suitable numbers $\underline{\lambda}_i^{[n]}(\varepsilon)$ independent of $\nu$. The advantage of (10.7) with respect to (10.6) is that for any $n$ we have to impose that the quantities $\omega\cdot\nu$ are far enough only from a finite number of values, that is $d$ values for (10.7a) and $\leq 4d^2$ values for (10.7b).

Thus, already for $n=-1$ we have to impose, besides the conditions (10.3), also the conditions (10.7b). This leaves a subset $\mathfrak{E}_{-1}\subset\mathfrak{E}'_{-1}$. To prove that the set $\mathfrak{E}_{-1}$ has still a Lebesgue density point at $\varepsilon=0$, we need a lower bound on all the derivatives $\mathrm{d}(\sqrt{\varepsilon a_i}\pm\sqrt{\varepsilon a_j})/\mathrm{d}\varepsilon$, with $i\neq j$ when the sign minus is considered. This is easily obtained if we assume that the eigenvalues $a_1,\dots,a_s$ are distinct, i.e. that there exists $a_0>0$ such that $|a_i-a_j|>a_0$ for all $1\leq i<j\leq s$; see Appendix E. Of course, this provides a further assumption on the function $f_{\mathbf{0}}(\boldsymbol{\beta})$.

To deal with the cases $n\geq 0$ we define, iteratively,

$$\underline{\lambda}_i^{[n]}(\varepsilon)=\begin{cases}0, & i=1,\dots,r,\\ \lambda_i^{[n]}(\sqrt{\underline{\lambda}_i^{[n-1]}(\varepsilon)},\varepsilon), & i=r+1,\dots,d.\end{cases}\tag{10.8}$$

In this way we obtain both that the eigenvalues $x^2-\lambda_i^{[n]}(x,\varepsilon)$ are bounded in terms of the quantities $x^2-\underline{\lambda}_i^{[n]}(\varepsilon)$ and that the sequences $\{\underline{\lambda}_i^{[n]}(\varepsilon)\}_{n=-1}^{\infty}$ converge exponentially fast for all $i=r+1,\dots,d$, that is $|\underline{\lambda}_i^{[n]}(\varepsilon)-\underline{\lambda}_i^{[n-1]}(\varepsilon)|\leq K_1\mathrm{e}^{-\kappa_1 2^{n/\tau'}}\varepsilon^2$ for suitable positive constants $K_1$ and $\kappa_1$; see Appendix F. Again, in order to impose the conditions (10.7b) we need a lower bound on the derivatives $\mathrm{d}(\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon)}\pm\sqrt{\underline{\lambda}_j^{[n]}(\varepsilon)})/\mathrm{d}\varepsilon$, but these can be discussed as in the case $n=-1$; see Appendix E.

The discussion above is correct as far as the self-energies are well-defined – which we have simply assumed to be for the moment. For instance, only if this is the case, when we write $\underline{\lambda}_i^{[n]}(\varepsilon)=a_{i-r}\varepsilon+O(\varepsilon^2)$ for $i=r+1,\dots,d$, we can really say that the high order terms are negligible with respect to the liner ones. As in the case of maximal and hyperbolic tori, we prove by induction that the self-energies are well defined. To this aim, we need bounds on the number of lines on scale $n$: one can prove that bounds of the form (5.8) and (7.2) still holds, but with $\tau'$ instead of $\tau$ – since the Diophantine conditions involve the Diophantine exponent $\tau'$; see Appendix F. Up to this difference, the strategy of the inductive proof is exactly as in the previous cases.

The Diophantine conditions (10.7a) and (10.7b) are known as the *first Melnikov conditions* and *second Melnikov conditions*, respectively. Each condition shrinks further the set of allowed values of $\varepsilon$: if $\mathfrak{E}_{n-1}$ is the set of allowed values found at the step $n-1$, then imposing the conditions (10.7) leaves a subset

$\mathfrak{E}_n \subset \mathfrak{E}_{n-1}$. At each step the set of values which are removed has measure proportional to a common value times $2^{-(n+1)/2}$: it was to obtain this exponential factor that a factor $2^{-(n+1)/2}$ was introduced in (10.7). Thus, eventually one is left with a set $\mathfrak{E}_\infty$, which is still of large measure. A closer inspection of $\mathfrak{E}_\infty$ reveals that $\mathfrak{E}_\infty$ is a Cantor set (that is a perfect, nowhere dense set).

To make the argument above really rigorous one should take a smooth decomposition, such as that considered in Section 5. Moreover, it turns out to be convenient to use functions $\underline{\lambda}_i^{[n]}(\varepsilon)$ which are differentiable (in the sense of Whitney [72]) in $\varepsilon$, so that, instead of the minimum of the eigenvalues, one should consider a smooth version of it – see Appendix F for details (see also [38, 45, 39]).

The assumptions on $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ can be weakened by requiring that the eigenvalues $a_1, \dots, a_s$ are such that $a_i \neq 0$ for all $i$ and $a_i - a_j \neq 0$ for all $i \neq j$. In this case *lower-dimensional tori of mixed type* can be proved to exist [45, 39].

# 11 Stability and uniqueness

The quasi-periodic solutions describing the maximal tori are linearly stable [3]. An interesting problem is that of uniqueness of solutions. In other words, one can wonder whether there are other quasi-periodic solutions with the same rotation vector $\boldsymbol{\omega}$ as the solution studied in the previous sections. Despite the apparent simplicity of the problem, a proof of uniqueness has been given only recently [31].

The case of lower-dimensional tori is more difficult. In principle there could be other quasi-periodic solutions with the same rotation vector, which either do not admit any perturbation expansion or admit a different expansion or, even admitting the same expansion, are different. For instance, the resummed expansion (6.6) *a priori* depends on the particular way the multiscale analysis is implemented, and by slightly changing the procedure one could obtain a different solution: this would imply infinitely many solutions which have the same formal perturbation series. In the case $d = 2$ and $\tau = 1$ all such functions coincide, because they are all Borel summable – see the last remark in Appendix D – but in general there is no reason why this should happen. At the present moment the problem of uniqueness is still open.

In the case of the dissipative systems of Section 3.4 one expects, on the ground of physical considerations, the response solution to be either attractive or repulsive (which means attractive for the time-reverted dynamics). More precisely, under the further assumption that $\partial_x g(c_0) > 0$ the response solution is expected to be asymptotically stable. Indeed, this is what happens.

The proof – very easy – proceeds as follows [4]. The analysis of Section 9 shows that there exists a response solution $x_0(t) = x(\underline{\omega} t, \varepsilon) = c_0 + O(\varepsilon)$. If we look for solutions of the form $x = x_0 + \xi$: then $x(t) \to x_0(t)$ as $t \to \infty$ (i.e. $x_0$ is attracting) if and only if $\xi(t) \to 0$ as $t \to \infty$. The function $\xi$ must solves the differential equation

$$\ddot{\xi} + \gamma\dot{\xi} + P(\xi, x_0(t)) = 0, \qquad P(\xi, x) = g(x + \xi) - g(x) = \partial_x g(x)\, \xi + O(\xi^2), \tag{11.1}$$

which can be rewritten as a system of first order equations

$$\begin{cases} \dot{\xi} = y, \\ \dot{y} = -\gamma y - P(\xi, x_0(t)). \end{cases} \tag{11.2}$$

If we define $R(\xi, t) = P(\xi, x_0(t))/P(\xi, c_0)$ we have $R(0, t) = 1 + O(\varepsilon)$, so that $1/2 \le R(\xi, t) \le 2$ uniformly in $t$ and $\xi$, for $\varepsilon$ and $\xi$ small enough. Then we can rescale time and variables by setting

$$\tau(t) = \int_0^t \mathrm{d}t' \, \sqrt{R(\xi(t'), t')}, \qquad \xi(t) = v(\tau(t)), \qquad y(t) = \sqrt{R(\xi(t), t)}\, w(\tau(t)), \tag{11.3}$$

which transforms the system (11.2) into

$$\begin{cases} v' = w, \\ w' = -\gamma(v,t)\, w - P(v, c_0), \end{cases} \tag{11.4}$$

with the prime denoting differentiation with respect to time $\tau$ and

$$\gamma(v,t) := \frac{1}{\sqrt{R}} \left( \gamma + \frac{R'}{2\sqrt{R}} \right). \tag{11.5}$$

If we neglect the friction term $\gamma(v,t)\, w$ in (11.4) we obtain an autonomous system with constant of motion

$$H(v,w) = \frac{1}{2} w^2 + \int_0^v \mathrm{d}v'\, P(v', c_0) = \frac{1}{2} w^2 + \frac{1}{2} \partial_x g(c_0)\, v^2 + O(v^3), \tag{11.6}$$

so that the origin is a stable equilibrium point. Moreover $\gamma(v,t) > 0$ for $\varepsilon$ small enough (recall that $\gamma = 1/\varepsilon$), in a neighbourhood $U$ of the origin. Hence we can apply Barbashin-Krasokvsky's theorem [62] (or Lasalle's invariance principle [55]) to conclude that the origin is asymptotically stable and $U$ is contained in its basin of attraction.

In some cases, for instance if $g(x) = x^{2p+1}$, $p \in \mathbb{N}$, and $f_{\mathbf{0}} \neq 0$ (so that $\partial_x g(c_0) > 0$), the response solution turns out to be a global attractor [4], but of course in general it is only locally attracting.

The problem of uniqueness, mentioned about the lower-dimensional tori, can be addressed also as to such response solutions. Of course, the local attractiveness of the solution implies local uniqueness. In other words, the response solution $x(\underline{\omega} t, \varepsilon)$ is the only quasi-periodic solution which reduces to $c_0$ as $\varepsilon \to 0$.

# 12 Generalisations

In this last section, we review some possible directions one can follow to generalise the results described in the previous sections. Some of the these generalisations are discussed in the literature, other are still open problems.

## 12.1 Weaker Diophantine conditions

Instead of the standard Diophantine condition (1.5) one can consider weaker conditions, such as the *Bryuno condition* [19]: a vector $\omega$ is said to satisfy the Bryuno condition if $\mathcal{B}(\omega) < \infty$, where

$$\mathcal{B}(\omega) = \sum_{n=1}^{\infty} \frac{1}{2^n} \log \frac{1}{\alpha_n(\omega)}, \qquad \alpha_n(\omega) = \inf\{|\omega \cdot \nu| : \nu \in \mathbb{Z}^d \text{ such that } 0 < |\nu| \leq 2^n\}. \tag{12.1}$$

All the results of the previous sections can be extended to rotation vectors satisfying the Bryuno condition: see [39] for maximal and lower-dimensional tori, [9] for the standard map, and [40] for dissipative systems.

For $d = 2$ one can write $\omega = (\omega_1, \omega_2) = (1, \alpha)\omega_1$, where $\alpha = \omega_2/\omega_1$ is the rotation number. One can define the *Bryuno function*

$$B(\alpha) = \sum_{n=1}^{\infty} \frac{1}{q_n} \log q_{n+1}, \tag{12.2}$$

where $\{q_n\}_{n \in \mathbb{Z}}$ are the denominators of the *best approximants* of $\alpha$ [77]. Then the function $\mathcal{B}(\omega)$ is equivalent to the Bryuno function $B(\alpha)$, in the sense that one has $C^{-1} < \mathcal{B}(\omega)/B(\alpha) < C$ for a universal constant $C$ [39].

An open problem is whether such a condition can be further weakened. For $d = 2$ the Bryuno condition is optimal: indeed, in the case of the standard map Davie [27] proved that if the rotation number $\omega \in \mathbb{R}$ does not satisfy the Bryuno condition then there is no invariant curve with that rotation number $\omega$. Such a result can be even strengthened by saying that the radius of convergence $\rho(\omega)$ of the conjugating function and the function $B(\omega)$ are such that

$$C_1 \mathrm{e}^{-2B(\omega)} < \rho(\omega) < C_2 \mathrm{e}^{-2B(\omega)}, \tag{12.3}$$

for suitable universal constants $C_1$ and $C_2$; see [27, 9] for a proof of the last statement. The proof of the lower bound in (12.3) relies on deeper cancellations than those discussed in Section 7; see [9] for details.

Another Diophantine condition considered in the literature is the so-called *Rüssmann condition* [75, 76, 73], which has a somewhat intricate definition if compared to (12.1). For $d = 2$ such a condition is equivalent to the Bryuno condition [76].

## 12.2 Degenerate perturbations

The dissipative systems introduced in Section 3.4 have been considered under the *nondegeneracy condition* that $\partial_x g(c_0) \neq 0$ if $g(c_0) = f_{\underline{0}}$. Such a condition can be removed, and the existence of a response solution can be proved under the only condition that there exists a zero $c_0$ of odd order to the equation $g(x) - f_{\underline{0}} = 0$; see [40], where it is also proved that there is no response solution reducing to $c_0$ as $\varepsilon \to 0$ if $c_0$ is a zero of even order.

Also in the case of lower-dimensional tori one can think of relaxing the nondegeneracy condition that the matrix $\partial^2_{\boldsymbol{\beta}} f_{\boldsymbol{0}}(\boldsymbol{\beta}_0)$ be nonsingular. In full generality, this case is very hard. The case $s = 1$ is already nontrivial. In that case it has be proved that at least one lower-dimensional torus always persists [20].

A first difference with respect to the nondegenerate case considered in Section 3.3 is that a formal power series in $\varepsilon$ does not exist any longer. In [36] it is shown that, at least in some cases, a fractional power series in $\varepsilon$ can be envisaged. The situation is somewhat reminiscent of what happens in *Melnikov theory* [66, 22, 54], when the *subharmonic Melnikov function* has degenerate zeroes of odd order: in that case subharmonic solutions exist and are analytic in a fractional power of $\varepsilon$ (*Puiseux series* [74, 18]) – cf. [44, 23]; see also [1, 70] for a similar situation in the case of limit cycles. In the case of lower-dimensional tori, the fractional power series in $\varepsilon$ do not converge, but they can be resummed in order to give well-defined functions – see [36] for a complete discussion.

## 12.3 More general systems

One can also consider ordinary differential equations more general than (1.1), say of the form

$$D_\varepsilon u = F_0(u) + \varepsilon F(u, \underline{\omega} t), \tag{12.4}$$

with $F_0$ real analytic. In that case one still assumes that the unperturbed equation $D_0(u) = F_0(u)$ admits a quasi-periodic solution $u_0(\boldsymbol{\omega} t)$.

The most general formulation of KAM theorem is within this class – see the Hamiltonian (3.2). The analysis performed in the previous sections for the simplified Hamiltonians (3.1) and (3.18) can be extended to deal with these Hamiltonians; we refer to [47, 39] for details. If the perturbation depends explicitly on the action variables, then the formal solubility of the Hamilton equations relies on a nondegeneracy condition of the unperturbed Hamiltonian, such as the invertibility of the matrix $\partial^2_{\boldsymbol{A}} \mathcal{H}_0(\boldsymbol{A})$ (*anisochronous condition*). However, the KAM theorem can be extended also to *isochronous systems* with $\mathcal{H}_0(\boldsymbol{A}) = \boldsymbol{\omega} \cdot \boldsymbol{A}$, by assuming a nondegeneracy condition on the perturbation, for instance that the matrix

$$f_{\boldsymbol{0}}(\boldsymbol{A}) = \int_{\mathbb{T}^d} \frac{\mathrm{d}\boldsymbol{\alpha}}{(2\pi)^d} f(\boldsymbol{\alpha}, \boldsymbol{A}) \tag{12.5}$$

is invertible [33]. A proof along the lines of the previous sections passes through the so-called *translated torus theorem* [69] (also knows as *theorem of the modifying terms* or *theorem of the counterterms*), which says that, for any analytic function $f\colon \mathbb{T}^n \times \mathbb{R}^n$ and any Diophantine vector $\boldsymbol{\omega} \in \mathbb{R}^n$, there exists a vector $\boldsymbol{\mu}(\varepsilon, \boldsymbol{\omega})$ analytic in $\varepsilon$, such that the equations

$$\begin{cases} \dot{\boldsymbol{\alpha}} = \boldsymbol{\omega} + \varepsilon \partial_{\boldsymbol{A}} f(\boldsymbol{\alpha}, \boldsymbol{A}) + \boldsymbol{\mu}(\varepsilon, \boldsymbol{\omega}), \\ \dot{\boldsymbol{A}} = -\varepsilon \partial_{\boldsymbol{\alpha}} f(\boldsymbol{\alpha}, \boldsymbol{A}), \end{cases} \tag{12.6}$$

admit a quasi-periodic solution with rotation vector $\boldsymbol{\omega}$ which is analytic in $\varepsilon$. A proof of such a theorem by using the tree formalism can be found in [8]. The theorem of the translated torus can be also formulated in Cartesian coordinates; in that case the cancellation mechanisms leading to the convergence of the series work in a rather different way [24].

In Section 3.1 we have considered only analytic Hamiltonians. A more general formulation of the KAM theorem requires only finite smoothness [68, 72]. In certain cases, the tree formalism can be extended to nonanalytic systems, such as some quasi-integrable systems of the form (3.3) with $f$ in a class of $C^p$ functions for some finite $p$ [12, 13]. However, up to exceptional cases, the method described here seems to be intrinsically suited in cases in which the vector fields are analytic – from a physical point of view this a quite reasonable assumption. 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, as shown in Sections 8 and 9, the perturbation series can be given a meaning also in cases where the solutions are not analytic in the perturbation parameter.

Equations of the form (12.4) also arise in problems of electronic engineering and theory of circuits, usually with periodic forcing. Such systems are resistive and hence intrinsically dissipative. As examples one can consider the *saturating inductor circuit*, described by the equation

$$G(\dot{x})\,\ddot{x} + \beta x + \varepsilon\gamma\dot{x} = \varepsilon f(\underline{\omega} t), \qquad G(v) = \frac{\alpha + v^2}{1 + v^2}, \qquad \alpha > 1, \quad \beta > 0, \quad \gamma > 0, \tag{12.7}$$

and the *resonant injection-locked frequency divider*, described by the equation

$$\begin{cases} x' = \alpha y + \beta x \left(1 - x^2\right) + \varepsilon x \left(1 - x^2\right) f(\underline{\omega} t), \qquad \alpha > \beta > 1, \\ y' = -x - y, \end{cases} \tag{12.8}$$

with $m = 1$ (i.e. periodic forcing) in both cases. For both equations the dynamics at $\varepsilon = 0$ is known: the first system (12.7) admits a constant of motion (although it is not Hamiltonian), while the second one has a globally attracting limit cycle. By following the same approach as described in Sections 2 and 4, one can study for $\varepsilon \neq 0$ the existence of periodic solutions whose period is rational with respect to the period of the forcing – with the major simplification with respect to the previous analysis that no small divisors appear. More precisely one can study the existence of subharmonic solutions for the equation (12.7), and the frequency locking phenomenon for the equation (12.8). We refer to the literature [5, 6] for details and results.

## 12.4 Partial differential equations

Finally, the analysis developed so far for ordinary differential equations, can be extended to partial differential equations, such as the *nonlinear wave equation*

$$\partial_{tt} u - \partial_{xx} u + \mu u = u^3, \qquad x \in [0, \pi], \qquad \mu \geq 0, \tag{12.9}$$

and the *nonlinear Schrödinger equation*

$$\mathrm{i}\partial_t u - \partial_{xx} u + \mu u = |u|^2 u, \qquad x \in [0, \pi], \qquad \mu \geq 0, \tag{12.10}$$

with periodic or Dirichlet boundary conditions. There exists a very wide literature about periodic and quasi-periodic small amplitude solutions to nonlinear one-dimensional partial differential equations such as (12.9) and (12.10), starting from the seminal work by Kuksin, Craig and Wayne [63, 80, 26]. Recently results have been obtained also in higher space dimension [15, 16, 30], that is for $x \in [0,\pi]^D$, $D > 1$, with periodic boundary conditions.

By using the tree formalism, small amplitude periodic solutions have been proved to exist, in dimension $D = 1$, both in the nonresonant case – $\mu$ in a suitable Cantor set [49] – and in the resonant case – $\mu = 0$ [50, 51]. Results have been obtained also in the higher space dimensional case $D > 1$ [52]. We refer to the original papers for a precise formulation of the results and the proofs.

# A Proof of the Siegel-Bryuno lemma

The bound (5.8) follows from the fact that if $\mathfrak{N}^*_n(\theta) \neq 0$ then $\mathfrak{N}^*_n(\theta) \le E(n,\theta) := c\,K(\theta)2^{-n/\tau} - 1$, with $c = 2^{2+1/\tau}$. The last bound can be proved by induction on the order $k(\theta)$ as follows. Given a tree $\theta$ let $\ell_0$ be its root line, let $\ell_1, \dots, \ell_s$, $s \ge 0$, be the lines on scales $\ge n$ which are the closest to $\ell_0$, and let $\theta_1, \dots, \theta_s$ the trees with root lines $\ell_1, \dots, \ell_s$, respectively – cf. Figure 7. By construction all lines $\ell$ in the subgraph $T$ have scales $n_\ell < n$, so that if $n_{\ell_0} \ge n$ then $T$ is necessarily a cluster. Moreover, all trees $\theta_1, \dots, \theta_s$ have orders strictly less than $k(\theta)$, so that, by the inductive hypothesis, for each $i = 1, \dots, s$ one has either $\mathfrak{N}^*_n(\theta_i) \le E(n,\theta_i)$ or $\mathfrak{N}^*_n(\theta_i) = 0$.

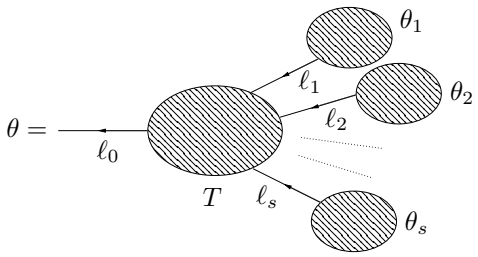


Figure 7: Construction for the proof of the Siegel-Bryuno lemma.

If either $\ell_0$ is not on scale $n$ or it is on scale $n$ but exits a self-energy cluster then $\mathfrak{N}^*_n(\theta) = \mathfrak{N}^*_n(\theta_1) + \dots + \mathfrak{N}^*_n(\theta_s)$ and the bound $\mathfrak{N}^*_n(\theta) \le E(n,\theta)$ follows by the inductive hypothesis. If $\ell_0$ does not exit a self-energy cluster and $n_{\ell_0} = n$ then $\mathfrak{N}^*_n(\theta) = 1 + \mathfrak{N}^*_n(\theta_1) + \dots + \mathfrak{N}^*_n(\theta_s)$, and the lines $\ell_1, \dots, \ell_s$ enter a cluster $T$ with $K(T) = K(\theta) - (K(\theta_1) + \dots + K(\theta_s))$. If $s \ge 2$ the bound $\mathfrak{N}^*_n(\theta) \le E(n,\theta)$ follows once more by the inductive hypothesis. If $s = 0$ then $\mathfrak{N}^*_n(\theta) = 1$; on the other hand for $\ell_0$ to be on scale $n_{\ell_0} = n$ one must have $|\omega \cdot \nu_{\ell_0}| < 2^{-n+1}\gamma$, which, by the Diophantine condition (1.8), implies $K(\theta) \ge |\nu_{\ell_0}| > 2^{(n-1)/\tau}$, hence $E(n,\theta) > 1$. If $s = 1$ call $\nu_1$ and $\nu_2$ the momenta of the lines $\ell_0$ and $\ell_1$, respectively (in particular $\nu_1 = \nu_{\ell_0}$). By construction $T$ cannot be a self-energy cluster, hence $\nu_1 \neq \nu_2$. Thus, by the Diophantine condition (1.5), one has

$$2^{-n+2}\gamma \ge |\omega \cdot \nu_1| + |\omega \cdot \nu_2| \ge |\omega \cdot (\nu_1 - \nu_2)| > \frac{\gamma}{|\nu_1 - \nu_2|^\tau}, \tag{A.1}$$

because $n_{\ell_0} = n$ and $n_{\ell_1} \ge n$, and hence

$$K(T) \ge \sum_{v \in N(T)} |\nu_v| \ge |\nu_1 - \nu_2| > 2^{(n-2)/\tau}, \tag{A.2}$$

hence $T$ must contain "many nodes". In particular, one finds also in this case $\mathfrak{N}^*_n(\theta) = 1 + \mathfrak{N}^*_n(\theta_1) \le 1 + E(n,\theta_1) \le 1 + E(n,\theta) - c\,K(T)2^{-n/\tau} \le E(n,\theta)$, where we have used that $c\,K(T)2^{-n/\tau} \ge 1$ by (A.2), provided $c = 2^{2+1/\tau}$.

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 8, where a tree $\theta$ 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 $\theta_1$ with root line $\ell$ is of order $k/3$. If the line $\ell$ entering the rightmost self-energy cluster $T_p$ has momentum $\nu$, also the lines exiting the $p$ self-energy clusters have the same momentum $\nu$. Suppose that $|\nu| \approx Nk/3$ and $|\omega \cdot \nu| \approx \gamma/|\nu|^\tau$ (this is certainly possible for some $\nu$). Then the value of the tree $\theta$ 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.

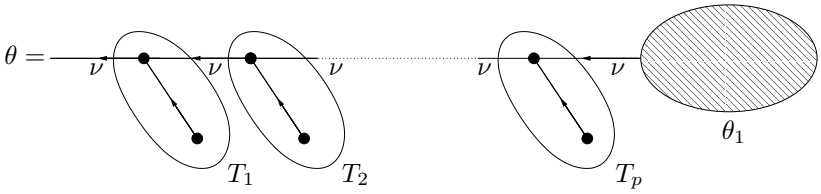


Figure 8: Example of tree with accumulation of small divisors.

# B Siegel-Bryuno lemma for self-energies

We first prove (7.3). Call $\mathcal{P}_T$ the path connecting the entering and exiting lines of $T$. If $T \in \mathcal{R}_n$ then $T$ contains at least a line on scale $\geq n$. If there is one line $\ell \notin \mathcal{P}_T$ on scale $\geq n$, one has $K(T) \geq |\nu_\ell|$ and $\gamma|\nu_\ell|^{-\tau} < |\omega \cdot \nu_\ell| \leq 2^{-n+1}\gamma$, so that $K(T) > 2^{(n-1)/\tau}$. Otherwise, let $\ell$ be the line $\ell \in \mathcal{P}_T$ on scale $\geq n$ which is closest to the line $\ell_1$ entering $T$. Call $\widetilde{T}$ the subgraph of $T$ consisting of all lines and nodes of $T$ preceding $\ell$. By construction, all lines in $\widetilde{T}$ have scale $< n$, and hence $\nu_\ell \neq \nu_{\ell_1}$, otherwise $\widetilde{T}$ would be a self-energy cluster. Therefore one has $\gamma|\nu_\ell - \nu_{\ell_1}|^{-\tau} < |\omega \cdot (\nu_\ell - \nu_{\ell_1})| < 2^{-n+2}\gamma$, which yields $K(T) \geq |\nu_\ell - \nu_{\ell_1}| > 2^{(n-2)/\tau}$.

To prove (7.2) one considers a more general class of graphs. We say that $T$ is a graph of class $\mathcal{S}(n, n')$ if it has one entering line and one exiting line both on scale $\geq n'$ and all the lines contained in $T$ are on scale $\leq n$. Define $K(T)$ and $\mathfrak{N}_n(T)$ as done in Section 7 for the self-energy clusters. We want to prove the bound $\mathfrak{N}^*_{n'}(T) \leq \max\{2K(T)2^{(3-n)/\tau} - 1, 0\}$. Let $\mathcal{P}_T$ the path connecting the entering and exiting lines of $T$, and let $N(\mathcal{P}_T)$ the set of nodes connected by lines of $\mathcal{P}_T$. If all the lines along $\mathcal{P}_T$ have scale $< n'$, then $\mathfrak{N}_{n'}(T) = \mathfrak{N}_{n'}(\theta_1) + \ldots + \mathfrak{N}_{n'}(\theta_m)$, where $\theta_1, \ldots, \theta_m$ are the trees contained in $T$ which have the root in a node $v \in N(\mathcal{P}_T)$. In that case the bound follows from (5.8). If there exists a line $\ell \in \mathcal{P}_T$ on scale $\geq n'$, call $T_1$ and $T_2$ the subgraphs of $T$ such that $L(T) = \{\ell\} \cup L(T_1) \cup L(T_2)$. Both $T_1$ and $T_2$ are of class $\mathcal{S}(n, n')$, so that, in the case in which both $T_1$ and $T_2$ contain lines on scale $\geq n'$, by the inductive hypothesis one finds $\mathfrak{N}_{n'}(T) \leq 1 + \mathfrak{N}_{n'}(T_1) + \mathfrak{N}_{n'}(T_2) \leq 2K(T)2^{(n-3)/\tau} - 1$. If $T_1$ contains no line on scale $\geq n'$ then one realises that one must have $K(T_1) > 2^{(n-2)/\tau}$, and the same holds for $T_2$, so that the bound follows also in these cases. Finally, (7.2) follows from the previous bound by noting that a self-energy cluster $T \in \mathcal{R}_n$ is a graph of class $\mathcal{S}(n, n')$ for all $n < n'$.

# C Accumulation of small divisors for dissipative systems

We want to construct for the model of Section 3.4 a tree $\theta$ whose value $\mathrm{Val}(\theta)$ grows like a factorial. Let $\theta$ be the tree with $k$ nodes $v_1, \ldots, v_k$, such that $s_{v_i} = k_{v_i} = 1$ and $\rho_{v_i} = 0$ for all $i = 1, \ldots, k-1$, while $s_{v_k} = 0$ and $\rho_{v_k} = k_{v_k} = 1$. Let $\nu_{v_k} = \nu$ such that $|\omega \cdot \nu| \approx \gamma/|\nu|^\tau$. The value of the labels $\rho_{v_i}$ for $i = 1, \ldots, k-1$ implies that $\nu_{v_i} = 0$ for $i = 1, \ldots, k-1$, and hence all the lines have the same

momentum $\nu$. Then one has $\mathrm{Val}(\theta) = (\mathrm{i}\omega\cdot\nu)^{2(k-1)} f_\nu (\mathrm{i}\omega\cdot\nu)^{-k} = (\mathrm{i}\omega\cdot\nu)^{k-2} f_\nu$, which can be bounded by $(k-2)!(2/\xi)^k \Xi_0 \mathrm{e}^{-\xi|\nu|/2}$ for large $k$.

# D Borel summability

Let $f(\varepsilon) = \sum_{n=1}^\infty a_n \varepsilon^n$ a formal power series (which means that the sequence $\{a_n\}_{n=1}^\infty$ is well-defined). We say that $f(\varepsilon)$ is *Borel summable* if

1. $B(p) := \sum_{n=1}^\infty a_n p^n/n!$ converges in some circle $|p| < \delta$,
2. $B(p)$ has an analytic continuation to a neighbourhood of the positive real axis, and
3. $g(\varepsilon) = \int_0^\infty \mathrm{e}^{-p/\varepsilon} B(p)\,\mathrm{d}p$ converges for some $\varepsilon > 0$.

Then the function $B(p)$ is called the *Borel transform* of $f(\varepsilon)$, and $g(\varepsilon)$ is the *Borel sum* of $f(\varepsilon)$. Moreover if the integral defining $g(\varepsilon)$ converges for some $\varepsilon_0 > 0$ then it converges in the circle $\mathrm{Re}\,\varepsilon^{-1} > \mathrm{Re}\,\varepsilon_0^{-1}$.

A function which admits the formal power series expansion $f(\varepsilon)$ is called Borel summable if $f(\varepsilon)$ is Borel summable; in that case the function equals the Borel sum $g(\varepsilon)$.

A remarkable property of Borel summable functions is that if two functions $f(\varepsilon)$ and $g(\varepsilon)$ are both Borel summable and admit the same power series expansion, then the two functions coincide.

# E Excluded values of the perturbation parameter

Set $a = \min\{a_1,\dots,a_s\}$ and $A = \max\{a_1,\dots,a_s\}$. In order to impose the Diophantine conditions (10.3) we have to exclude all values of $\varepsilon \in [0,\varepsilon_0]$ such that $\big||\omega\cdot\nu| - \sqrt{\varepsilon a_i}\big| \le \gamma|\nu|^{-\tau'}$ for some $i = 1,\dots,s$ and some $\nu \neq 0$. Of course, we can confine ourselves to the vectors $\nu \in \mathbb{Z}^r$ such that $|\nu| \ge m_0 := (\gamma/4\sqrt{\varepsilon_0 A})^{-1/\tau}$, because one has $|\omega\cdot\nu| > 4\sqrt{\varepsilon_0 A}$ and hence $\big||\omega\cdot\nu| - \sqrt{\varepsilon a_i}\big| > |\omega\cdot\nu|/2 > \gamma|\nu|^{-\tau}/2$ for $|\nu| < m_0$. For all $|\nu| \ge m_0$ we can introduce an interpolation parameter $t \in [-1,1]$ by setting $|\omega\cdot\nu| - \sqrt{\varepsilon(t,\nu)\,a_i} = t\,\gamma|\nu|^{-\tau'}$, so that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(t,\nu)\right| \le \frac{\gamma}{|\nu|^{\tau'}} \frac{2\sqrt{\varepsilon(t,\nu)}}{\sqrt{a_i}} \le \frac{\gamma}{|\nu|^{\tau'}} \frac{2\sqrt{\varepsilon_0}}{\sqrt{a}} \tag{E.1}$$

for all $\varepsilon(t,\nu) \in \mathfrak{E}'_{-1}$. Therefore we have to exclude a set $\mathfrak{E}'_{-1} \subset [0,\varepsilon_0]$ of measure

$$\mathrm{meas}(\mathfrak{E}'_{-1}) = \int_{\mathfrak{E}'_{-1}} \mathrm{d}\varepsilon \le \sum_{\substack{\nu\in\mathbb{Z}^r \\ |\nu|\ge m_0}} \int_{-1}^1 \mathrm{d}t \left|\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(t,\nu)\right| \le \sum_{\substack{\nu\in\mathbb{Z}^r \\ |\nu|\ge m_0}} \frac{\gamma}{|\nu|^{\tau'}} \frac{4\sqrt{\varepsilon_0}}{\sqrt{a}} \le C\gamma \left(\frac{\sqrt{\varepsilon_0 A}}{\gamma}\right)^{(\tau'-r)/\tau} \frac{\sqrt{\varepsilon_0}}{\sqrt{a}}, \tag{E.2}$$

for some universal constant $C$. Hence $\mathrm{meas}(\mathfrak{E}'_{-1})$ is much smaller than $\varepsilon_0$ if $\varepsilon_0$ is small and $\tau' > \tau + r$.

To impose the Diophantine conditions (10.7a) one can reason in the same way. One uses that $\underline{\lambda}_i^{[n]}(\varepsilon) = a_{i-r}\varepsilon + O(\varepsilon^2)$ for $i = r+1,\dots,d$ (see Appendix F), which yields $|\underline{\lambda}_i^{[n]}(\varepsilon)| \le 2a_{i-r}\varepsilon$ and $|\mathrm{d}\underline{\lambda}_i^{[n]}(\varepsilon)/\mathrm{d}\varepsilon| \ge a_{i-r}/2$; here and henceforth the derivative is in the sense of Whitney. Again we have to consider only the values $\nu \in \mathbb{Z}^r$ such that $|\nu| \ge m_0$. We define $\mathfrak{E}'_n$ as the set of values $\varepsilon \in [0,\varepsilon_0]$ which do not satisfy (10.7a) and for all $|\nu| \ge m_0$ and $t \in [-1,1]$ we write $|\omega\cdot\nu| - \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon(t,\nu))} = t\,2^{-(n+1)/2}\gamma|\nu|^{-\tau'}$. Then

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(t,\nu)\right| \le 2^{-(n+1)/2} \frac{\gamma}{|\nu|^{\tau'}} \frac{4\sqrt{2\varepsilon}}{\sqrt{a_{i-r}}} \le 2^{-(n+1)/2} \frac{\gamma}{|\nu|^{\tau'}} \frac{4\sqrt{2\varepsilon_0}}{\sqrt{a}} \tag{E.3}$$

for all $\varepsilon(t,\nu) \in \mathfrak{E}'_n$, and hence

$$\mathrm{meas}(\mathfrak{E}'_n) \le C 2^{-(n+1)/2} \gamma \left( \frac{\sqrt{\varepsilon_0 A}}{\gamma} \right)^{(\tau'-r)/\tau} \frac{\sqrt{\varepsilon_0}}{\sqrt{a}}, \tag{E.4}$$

for some universal constant $C$. Therefore one has $\mathrm{meas}(\mathfrak{E}'_n) = 2^{-(n+1)/2} o(\varepsilon_0)$, provided $\tau' > \tau + r$.

To impose the Diophantine conditions (10.7b), the only difference is that we need a lower bound on the derivatives $\mathrm{d}(\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon)} \pm \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon)})/\mathrm{d}\varepsilon$, $r+1 \le i, j \le d$, with $i \ne j$ when the sign minus is taken. One easily realises that the conditions with the sign plus do not present any further difficulty with respect to the first Melnikov conditions. Moreover, for $\varepsilon$ small enough and $r+1 \le i \ne j \le d$ one has

$$\left| \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \left( \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon)} - \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon)} \right) \right| \ge \frac{a_0}{8\sqrt{A\varepsilon}}, \qquad a_0 := \min_{1 \le k \ne h \le s} |a_k - a_h| \,. \tag{E.5}$$

To deduce (E.5) one uses that for $i = r+1, \dots, d.$ one has $\lambda_i^{[-1]}(\varepsilon) = a_{i-r}\varepsilon$ and $\lambda_i^{[n]}(\varepsilon) = a_{i-r}\varepsilon + O(\varepsilon^2)$, $n \ge 0$. By defining $\mathfrak{E}''_n$ as the set of values $\varepsilon \in [0, \varepsilon_0]$ which do not satisfy (10.7b), we obtain

$$\mathrm{meas}(\mathfrak{E}''_n) \le C 2^{-(n+1)/2} \gamma \left( \frac{\sqrt{\varepsilon_0 A}}{\gamma} \right)^{(\tau'-r)/\tau} \frac{\sqrt{A\varepsilon_0}}{a_0}, \tag{E.6}$$

for some universal constant $C$, so that once more $\mathrm{meas}(\mathfrak{E}''_n) = 2^{-(n+1)/2} o(\varepsilon_0)$ for $\tau' > \tau + r$. The sets $\mathfrak{E}_n$ in Section 10 are defined as $\mathfrak{E}_n = \mathfrak{E}'_n \cup \mathfrak{E}''_n$.

# F Multiscale analysis for elliptic tori

To extend the multiscale analysis to the case of elliptic tori, we slightly change the recursive definition of propagators and self-energies. We set $\Delta^{[-1]}(x,\varepsilon) = x^2$ and

$$\Delta^{[n]}(x,\varepsilon) = \left( \frac{1}{d} \sum_{i=1}^{d} \left( x^2 - \underline{\lambda}_i^{[n-1]}(\varepsilon) \right)^{-2} \right)^{-1/2}, \qquad n \ge 0, \tag{F.1}$$

and define

$$\Xi_n(x,\varepsilon) = \prod_{p=0}^{n} \chi_p(\Delta^{[p-1]}(x,\varepsilon)), \qquad \Psi_n(x,\varepsilon) = \psi_n(\Delta^{[n-1]}(x,\varepsilon)) \prod_{p=0}^{n-1} \chi_p(\Delta^{[p-1]}(x,\varepsilon)), \tag{F.2}$$

with the functions $\chi_n$ and $\psi_n$ defined as in Section 5, with the only difference that in (5.2) $\gamma$ and $\gamma/2$ are replaced with $\gamma^2$ and $\gamma^2/4$, respectively.

In terms of the quantities (F.1) and (F.2), the propagators $\mathfrak{G}_\ell = G^{[n_\ell]}(\omega \cdot \nu_\ell, \varepsilon)$ are defined iteratively as

$$G^{[n]}(x,\varepsilon) = \Psi_n(x,\varepsilon) \left( \delta_0(x) \mathbb{1} - \mathcal{M}^{[n-1]}(x,\varepsilon) \right)^{-1} \tag{F.3a}$$

$$\mathcal{M}^{[n]}(x,\varepsilon) = \mathcal{M}^{[n-1]}(x,\varepsilon) + \Xi_n(x,\varepsilon)\, M^{[n]}(x,\varepsilon), \qquad M^{[n]}(x,\varepsilon) = \sum_{T \in \mathcal{R}_n} \varepsilon^{k(T)}\, \mathcal{V}_T(x). \tag{F.3b}$$

Finally the numbers $\underline{\lambda}_i^{[n]}(\varepsilon)$ are defined according to (10.8), where $\lambda_i^{[n]}(x,\varepsilon)$ are the eigenvalues of the matrix $\mathcal{M}^{[n]}(x,\varepsilon)$.

If a line $\ell$ is on scale $n$ and $\mathfrak{G}_\ell \neq 0$, then one has

$$\min_{1\le i\le d} \left|(\omega\cdot\nu)^2 - \underline{\lambda}_i^{[p]}(\varepsilon)\right| \le 2^{-2p}\gamma^2, \qquad 0\le p\le n-2, \tag{F.4a}$$

$$\min_{1\le i\le d} \left|(\omega\cdot\nu)^2 - \underline{\lambda}_i^{[n-1]}(\varepsilon)\right| \ge \frac{1}{4\sqrt{d}} 2^{-2n}\gamma^2. \tag{F.4b}$$

Therefore, setting $\omega\cdot\nu = x$, if $x>0$ one has

$$\left|\lambda_i^{[n-1]}(x,\varepsilon) - \underline{\lambda}_i^{[n-1]}(\varepsilon)\right| \le \max_x \left|\partial_x \lambda_i^{[n-1]}(x,\varepsilon)\right| \left|\sqrt{\underline{\lambda}_i^{[n-2]}(\varepsilon)} - x\right|, \tag{F.5}$$

where $\partial_x \lambda_i^{[n-1]}(x,\varepsilon) = O(\varepsilon^2)$ and

$$\left|\sqrt{\underline{\lambda}_i^{[n-2]}(\varepsilon)} - x\right| \le \frac{|\underline{\lambda}_i^{[n-2]}(\varepsilon) - x^2|}{|\sqrt{\underline{\lambda}_i^{[n-2]}(\varepsilon)} + x|} \le \frac{2^{-2(n-2)}\gamma^2}{\sqrt{\varepsilon a}}, \tag{F.6}$$

and hence

$$\left|\lambda_i^{[n-1]}(x,\varepsilon) - \underline{\lambda}_i^{[n-1]}(\varepsilon)\right| \le C\,\varepsilon_0 2^{-2n}, \tag{F.7}$$

for some positive constant $C$. Therefore (F.4b) and (F.7) imply

$$\left|x^2 - \lambda_i^{[n-1]}(x,\varepsilon)\right| \ge \left|x^2 - \underline{\lambda}_i^{[n-1]}(\varepsilon)\right| - \left|\lambda_i^{[n-1]}(x,\varepsilon) - \underline{\lambda}_i^{[n-1]}(\varepsilon)\right| \ge \frac{1}{2}\left|x^2 - \underline{\lambda}_i^{[n-1]}(\varepsilon)\right|. \tag{F.8}$$

The case $x<0$ is discussed in the same way noting that $\lambda_i^{[n-1]}(-x,\varepsilon) = \lambda_i^{[n-1]}(x,\varepsilon)$, because of (7.1). Therefore the eigenvalues $x^2 - \lambda_i^{[n]}(x,\varepsilon)$ can be bounded from below by half the quantities $x^2 - \underline{\lambda}_i^{[n]}(\varepsilon)$.

The property $|\underline{\lambda}_i^{[n]}(\varepsilon) - \underline{\lambda}_i^{[n-1]}(\varepsilon)| \le C_1 \mathrm{e}^{-\kappa_1 2^{n/\tau'}}\varepsilon^2$ mentioned after (10.8) follows from the expression (F.3b) for $\mathcal{M}^{[n]}(x,\varepsilon) - \mathcal{M}^{[n-1]}(x,\varepsilon)$, the bound (7.5) for the values of the self-energy clusters, and the bound $\sum_{v\in N(T)} |\nu_v| > c'' 2^{n/\tau'}$ which holds for any $T\in\mathcal{R}_n$.

Finally we want to show that the bounds

$$\mathfrak{N}_n^*(\theta) \le c\, 2^{-n/\tau'} K(\theta), \qquad \mathfrak{N}_{n'}^*(T) \le c\, 2^{-n'/\tau'} K(T), \quad T\in\mathcal{R}_n, \quad n'\le n, \tag{F.9}$$

hold with the multiscale analysis described above. One proceed as in Appendices A and B, with the following changes. If the propagator of a line with momentum $\nu$ and scale $n$ is non-zero, then (F.4) imply $|\nu| \ge 2^{(n-2)/\tau'}$. When discussing the analogous of the case $s=1$ in Appendix A, then (A.1) must be replaced with

$$\begin{aligned} 2^{-n+3}\gamma &\ge |\,\omega\cdot\nu_1 + \sigma_1\sqrt{\underline{\lambda}_i^{[n-2]}(\varepsilon)}\,| + |\,\omega\cdot\nu_2 + \sigma_2\sqrt{\underline{\lambda}_j^{[n-2]}(\varepsilon)}\,| \\ &\ge |\,\omega\cdot(\nu_1-\nu_2) + \sigma_1\sqrt{\underline{\lambda}_i^{[n-2]}(\varepsilon)} - \sigma_2\sqrt{\underline{\lambda}_j^{[n-2]}(\varepsilon)}\,| \ge \frac{\gamma}{|\nu_1-\nu_2|^{\tau'}}, \end{aligned} \tag{F.10}$$

where the signs $\sigma_1,\sigma_2\in\{\pm\}$ and the labels $i,j\in\{1,\dots,d\}$ are such that the first inequality is satisfied. Analogously one discussed the case of the self-energy clusters.

In particular (F.10) explains why the second Melnikov conditions are necessary. Of course, if we had used directly the eigenvalues $\lambda_i^{[n]}(x,\varepsilon)$, instead of the quantities $\underline{\lambda}_i^{[n]}(\varepsilon)$, we would have require (10.6) instead of (10.7b). We have already seen in Section 10 why this was not allowed.

**Acknowledgments.** I thank L. Corsi for her very careful comments on the manuscript.

# References


[1] A.A. Andronov, E.A. Leontovich, I.I. Gordon, A.G. Maĭer, *Theory of bifurcations of dynamic systems on a plane*, Halsted Press, Israel Program for Scientific Translations, Jerusalem-London, 1973.

[2] V.I. Arnold, *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** (1963), no. 5 (113), 13–40.

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

[4] M.V. Bartuccelli, J.H.B. Deane, G. Gentile, *Globally and locally attractive solutions for quasi-periodically forced systems*, J. Math. Anal. Appl. **328** (2007), no. 1, 699-714.

[5] M.V. Bartuccelli, J.H.B. Deane, G. Gentile, *Bifurcation phenomena and attractive periodic solutions in the saturating inductor circuit*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **463** (2007), no. 2085, 2351-2369.

[6] M.V. Bartuccelli, J.H.B. Deane, G. Gentile, *Frequency locking in an injection-locked frequency divider equation*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **465** (2009), no. 2101, 283-306.

[7] M.V. Bartuccelli, J.H.B. Deane, G. Gentile, L. Marsh, *Invariant sets for the varactor equation*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **462** (2006), no. 2066, 439-457.

[8] M. Bartuccelli, G. Gentile, *Lindstedt series for perturbations of isochronous systems. A review of the general theory*, Rev. Math. Phys. **14** (2002), no. 2, 121-171.

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

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

[11] B. Bollobás, *Graph theory. An introductory course*, Graduate Texts in Mathematics Vol. 63, Springer-Verlag, New York-Berlin, 1979.

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

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

[14] E. Borel, *Leçons sur les séries divergentes*, Gauthier-Villars, Paris, 1928.

[15] J. Bourgain, *Construction of periodic solutions of nonlinear wave equations in higher dimension*, Geom. Funct. Anal. **5** (1995), no. 4, 629–639.

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

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

[18] E. Brieskorn, H. Knörrer, *Plane algebraic curves*, Birkhäuser, Basel, 1986.

[19] A.D. Bryuno, *Analytic form of differential equations. I, II*, Trudy Moskov. Mat. Obšč. **25** (1971), 119–262; ibid. **26** (1972), 199–239. English translation: *Trans. Moscow Math. Soc.* **25** (1971), 131–288 (1973); ibid. **26** (1972), 199–239 (1974).

[20] Ch.-Q. Cheng, *Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems*, Commun. Math. Phys. **177** (1996), 529–559.

[21] B.V. Chirikov, *A universal instability of many-dimensional oscillator systems*, Phys. Rep. **52** (1979), no. 5, 264–379.

[22] S.-N. Chow, J.K. Hale, *Methods of bifurcation theory*, Grundlehren der Mathematischen Wissenschaften Vol. 251, Springer-Verlag, New York-Berlin, 1982.

[23] L. Corsi, G. Gentile, *Melnikov theory to all orders and Puiseux series for subharmonic solutions*, J. Math. Phys. **49** (2008), no. 11, 112701, 29 pp.

[24] L. Corsi, G. Gentile, M. Procesi, *KAM theory in configuration space and cancellations in the Lindstedt series*, preprint.

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

[26] W. Craig, C.E. Wayne, *Newton's method and periodic solutions of nonlinear wave equations*, Comm. Pure Appl. Math. **46** (1993), 1409–1498.

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

[28] L.H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **15** (1988), no. 1, 115–147.

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

[30] L.H. Eliasson, S.B. Kuksin, *KAM for nonlinear Schrödinger equation*, Ann. Math., to appear.

[31] A. Fathi, A. Giuliani, A. Sorrentino, *Uniqueness of invariant Lagrangian graphs in a homology or a cohomology class*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., to appear.

[32] G. Gallavotti, *Perturbation theory for classical Hamiltonian systems*, Scaling and self-similarity in physics (Bures-sur-Yvette, 1981/1982), 359–426, Progress in Physics 7, Birkhäuser Boston, Boston, 1983.

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

[34] G. Gallavotti, F. Bonetto, G. Gentile, *Aspects of ergodic, qualitative and statistical theory of motion*, Texts and Monographs in Physics, Springer-Verlag, Berlin, 2004.

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

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

[37] G. Gentile, *Diagrammatic techniques in perturbations theory, and applications*, Symmetry and perturbation theory (Rome, 1998), 59–78, World Scientific Publishing, River Edge, NJ, 1999.

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

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

[40] G. Gentile, *Quasi-periodic motions in strongly dissipative forced systems*, Ergodic Theory Dynam. Systems, to appear.

[41] G. Gentile, *Construction of quasi-periodic response solutions in forced strongly dissipative systems*, preprint.

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

[43] G. Gentile, M. Bartuccelli, J. Deane, *Quasi-periodic attractors, Borel summability and the Bryuno condition for strongly dissipative systems*, J. Math. Phys. **47** (2006), no. 7, 072702, 10 pp.

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

[45] G. Gentile, G. Gallavotti, *Degenerate elliptic resonances*, Comm. Math. Phys. **257** (2005), no. 2, 319-362.

[46] G. Gentile, V. Mastropietro, *KAM theorem revisited*, Phys. D **90** (1996), no. 3, 225–234.

[47] 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*, Rev. Math. Phys. **8** (1996), no. 3, 393–444.

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

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

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

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

[52] G. Gentile, M. Procesi, *Periodic solutions for a class of nonlinear partial differential equations in higher dimension*, Comm. Math. Phys. **289** (2009), no. 3, 863-906.

[53] S.M. Graff, *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations **15** (1974), 1–69.

[54] J. Guckenheimer, Ph. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Applied Mathematical Sciences Vol. 42, Springer-Verlag, New York, 1990.

[55] J. Hale, M. Koçak, *Dynamics and bifurcations*, Texts in Applied Mathematics Vol. 3, Springer-Verlag, New York, 1991.

[56] F. Harary, *Graph theory*, Addison-Wesley, Reading, 1969.

[57] F. Harary, E.M. Palmer, *Graphical enumeration*, Academic Press, New York-London, 1973.

[58] G.H. Hardy, *Divergent series*, Oxford University Press, Oxford, 1949.

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

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

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

[62] N.N. Krasovskiĭ, *Stability of motion. Applications of Lyapunov's second method to differential systems and equations with delay*, Stanford University Press, Stanford, 1963.

[63] S.B. Kuksin, *Nearly integrable infinite-dimensional Hamiltonian systems*, Lecture Notes in Mathematics Vol. 1556, Springer-Verlag, Berlin, 1993.

[64] A.J. Lichtenberg, M.A. Lieberman, *Regular and chaotic dynamics*, Springer, Berlin (1992).

[65] R.S. MacKay, *Renormalisation in area-preserving maps*, Advanced Series in Nonlinear Dynamics Vol. 6. World Scientific Publishing, River Edge, NJ, 1993.

[66] V.K. Melnikov, *On the stability of a center for time-periodic perturbations*, Trudy Moskov. Mat. Obšč. **12** (1963), 3–52. English translation: Trans. Moscow Math. Soc. **12** (1963), 1–57.

[67] V.K. Melnikov, *A certain family of conditionally periodic solutions of a Hamiltonian system*, Dokl. Akad. Nauk SSSR **181** (1968), 546–549; English translation: Soviet Math. Dokl. **9** (1968), 882–886.

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

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

[70] L.M. Perko, *Global families of limit cycles of planar analytic systems*, Trans. Amer. Math. Soc. **322** (1990), no. 2, 627–656.

[71] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste. Vol. III*, Gauthier-Villars, Paris, 1899.

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

[73] J. Pöschel, *On elliptic lower-dimensional tori in Hamiltonian systems*, Math. Z. **202** (1989), no. 4, 559–608.

[74] M.V. Puiseux, *Recherches sur les fonctions algébriques*, J. Math. Pures Appl. **15** (1850), 365–480.

[75] H. Rüssmann, *On the one-dimensional Schrödinger equation with a quasiperiodic potential*, Nonlinear dynamics (Internat. Conf., New York, 1979), pp. 90–107, Ann. New York Acad. Sci. Vol. 357, New York Academy of Sciences, New York, 1980.

[76] H. Rüssmann, *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regul. Chaotic Dynam. **6** (2001), 119–204.

[77] W.M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics Vol. 785, Springer, Berlin, 1980.

[78] A.D. Sokal, *An improvement of Watson's theorem on Borel summability*, J. Math. Phys. **21** (1980), no. 2, 261–263.

[79] D.V. Treshchëv, *A mechanism for the destruction of resonance tori in Hamiltonian systems*, Mat. Sb. **180** (1989), no. 10, 1325–1346. English translation: Math. USSR-Sb. **68** (1991), no. 1, 181–203.

[80] C.E. Wayne, *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. **127** (1990), no. 3, 479–528.