Degenerate lower-dimensional tori under the Bryuno condition

Guido Gentile

Dipartimento di Matematica, Università di Roma Tre, Roma, I-00146, Italy

Abstract. *We study the problem of conservation of maximal and lower-dimensional invariant tori for analytic convex quasi-integrable Hamiltonian systems. In the absence of perturbation the lower-dimensional tori are degenerate, in the sense that the normal frequencies vanish, so that the tori are neither elliptic nor hyperbolic. We show that if the perturbation parameter is small enough, for a large measure subset of any resonant submanifold of the action variable space, under some generic non-degeneracy conditions on the perturbation function, there are lower-dimensional tori which are conserved. They are characterised by rotation vectors satisfying some generalised Bryuno conditions involving also the normal frequencies. We also show that, again under some generic assumptions on the perturbation, any torus with fixed rotation vector satisfying the Bryuno condition is conserved for most values of the perturbation parameter in an interval small enough around the origin. According to the sign of the normal frequencies and of the perturbation parameter the torus becomes either hyperbolic or elliptic or of mixed type.*

1. Introduction

It is well known that in quasi-integrable analytic Hamiltonian systems KAM invariant tori are conserved under conditions on the rotation vectors milder than the usual Diophantine condition originally introduced by Kolmogorov [38]. A more general condition was introduced by Bryuno in Refs. [8] and [9], and it is nowadays known as the *Bryuno condition*. Among the most exhaustive studies in this direction we cite those by Rüssmann (for a recent review see Ref. [49]). In some related problems, such as Siegel's problem (in the analytic framework), one knows that the Bryuno condition is a necessary and sufficient condition for the dynamics to be conjugated to the linear one, as the work by Yoccoz has shown [59]. In the case of area-preserving maps, the same result has been explicitly verified for the standard map. For more general maps, as well as for higher dimensional systems, it is not clear whether a result of this kind holds: the Bryuno condition is sufficient, but the necessary part seems not to hold entirely. For Siegel's problem, Yoccoz also proved that a deep relationship exists between the radius of convergence of the linearising function and the so-called Bryuno function. An analogous relationship between the radius of convergence of the conjugating function and the Bryuno function has been found for the standard map by combining the results of Davie [14] with those of Berretti and Gentile [2].

We mention also the work by Ecalle and Vallet [15], where it is shown that, under the Bryuno condition, all analytic resonant vector fields and diffeomorphisms admit an analytic correction which make them linearisable (as conjectured by Gallavotti [18], and proved under the usual Diophantine condition by Eliasson [17], hence by Gentile and Mastropietro [29] with techniques more similar to those we use in the present paper). Note that for such a problem the rotation vector is fixed and no value of the perturbation parameter has to be excluded, so that the problem rather simplifies, as all the difficulties related to estimating the measure of the allowed values for the parameters disappear. For instance there is not the difficulty of including the correction in the original vector field with a different unperturbed rotation vector, as in the case of the KAM theorem for isochronous systems (cf. Ref. [1] for a discussion within the formalism used here).

Extensions of the Bryuno condition to other contexts, such as the reducibility of skew-products on $\mathbb{T}^d \times \mathrm{SL}(2,\mathbb{R})$, has been provided recently by Lopes Dias [42], for Bryuno base flows on $\mathbb{T}^2$, and by Gentile [25], for Bryuno base flows in any dimension d. Existence and properties of quasi-periodic solutions for dissipative systems in the presence of a quasi-periodic forcing with Bryuno frequency vector has been studied in Ref. [26] and [27].

In this paper we consider a problem of lower-dimensional tori similar to that considered in Refs. [46] and [49], with the main difference being that the normal frequencies vanish in the absence of perturbation. Such a problem has been explicitly considered in a series of papers, such as Refs. [55], [34], [10], [19] and [28]. We refer to the latter for an introduction, and for a review of the existing results. In particular we start by considering the same class of Hamiltonian systems with $d \geq 2$ degrees of freedom considered in Refs. [19] and [28], originally introduced in Ref. [54],

$$\mathcal{H} = \frac{1}{2}\mathbf{A}\cdot\mathbf{A} + \frac{1}{2}\mathbf{B}\cdot\mathbf{B} + \varepsilon f(\boldsymbol{\alpha},\boldsymbol{\beta}), \tag{1.1}$$

where $(\boldsymbol{\alpha},\mathbf{A}) \in \mathbb{T}^r \times \mathbb{R}^r$ and $(\boldsymbol{\beta},\mathbf{B}) \in \mathbb{T}^s \times \mathbb{R}^s$ are conjugate action-angle variables, with $r+s=d$, and $\cdot$ denotes the inner product both in $\mathbb{R}^r$ and in $\mathbb{R}^s$. The perturbation $f(\boldsymbol{\alpha},\boldsymbol{\beta})$ is assumed to be real analytic, so that, if we write

$$f(\boldsymbol{\alpha},\boldsymbol{\beta}) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^r} \mathrm{e}^{i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}} f_{\boldsymbol{\nu}}(\boldsymbol{\beta}), \tag{1.2}$$

there exist positive constants F_0, F_1 and κ_0 such that $|\partial_{\boldsymbol{\beta}}^q f_{\boldsymbol{\nu}}(\boldsymbol{\beta})| \leq q! F_0 F_1^q \mathrm{e}^{-\kappa_0|\boldsymbol{\nu}|}$ for all $\boldsymbol{\nu} \in \mathbb{Z}^r$, all $\boldsymbol{\beta} \in \mathbb{T}^s$ and all $q \in \mathbb{Z}_+$. For $\boldsymbol{\beta}_0$ a stationary point of $f_{\mathbf{0}}(\boldsymbol{\beta})$ we call $a_1,\ldots,a_s$ the eigenvalues of the matrix $\partial_{\boldsymbol{\beta}}^2 f_{\mathbf{0}}(\boldsymbol{\beta}_0)$. The case of maximal tori is recovered by setting $r=d$. The general case of Hamiltonians describing perturbations of any convex systems will be briefly discussed in Appendix A4, even if the full discussion is deferred to Ref. [21]. Here we prefer to concentrate ourselves to the simpler model (1.1), in order to distinguish between the more relevant features of the renormalisation group techniques and the more technical intricacies pertaining rather to problems of spectral analysis and matrix algebra.

For $\varepsilon = 0$ the system described by the Hamiltonian (1.1) is integrable. Any solution of the form $(\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{A},\mathbf{B}) = (\boldsymbol{\alpha}_0 + \boldsymbol{\omega}t, \boldsymbol{\beta}_0, \mathbf{A}_0, \mathbf{B}_0)$, with $\boldsymbol{\omega} = \mathbf{A}_0$ having rationally independent components, fills densely a lower-dimensional torus with rotation vector $\boldsymbol{\omega}$. We call $\mathbf{A}$ and $\mathbf{B}$ the *non-resonant* and *resonant*, respectively, action variables. With a shift of the resonant action variables we can always assume $\mathbf{B}_0 = \mathbf{0}$. The so-called *normal frequencies*, that is the frequencies describing the dynamics of the $(\boldsymbol{\beta},\mathbf{B})$-variables, vanish for $\varepsilon = 0$, so that the considered unperturbed torus is neither elliptic nor hyperbolic (nor of mixed type). We refer to such a situation by saying that one has a *degenerate torus*.

The frequency map $\mathbf{A} \mapsto \boldsymbol{\omega}(\mathbf{A})$ is a local diffeomorphism (in our case it is trivially the identity), so that the condition $\mathbf{B} = \mathbf{0}$ defines an r-dimensional manifold $\mathbb{M}_r$ (*resonant submanifold*), which is determined by the space of the non-resonant action variable $\mathbf{A}$; we call the latter the *non-resonant action variable space*.

We are interested in two different problems.

(A) One can fix the perturbation parameter (small enough) and study for which rotation vectors some invariant tori are conserved, in the spirit of the KAM theorem for maximal tori, and as done in most of the papers on such a subject, as Refs. [43], [44], [39], [40], [16], [46], [11], [49], [5], [36], and many others.

(B) Either one can look at a lower-dimensional invariant torus with fixed rotation vector, and study the dependence of such a torus on the perturbation parameter. For instance this has been done, with the techniques used here, in Refs. [19] and [28].

The same twofold program has been followed, under the usual Diophantine condition, in Refs. [23] and [24] in the study of the quasi-periodic solutions and of the spectrum for a class of two-level systems in a strong quasi-periodic external field.

About problem (A) we find the analogue of Rüssmann's [49] and Pöschel's [46] result for systems with distinct[1] normal frequencies of order 1. In addition, our result applies also to the case of non-distinct normal frequencies, provided that they are all different from zero and of order ε, that is provided degeneracy is removed to first order. In particular this means that our result does not follow from the works available in literature: in principle one could think to perform a canonical transformation which introduces normal frequencies of order $\sqrt{|\varepsilon|}$, while keeping the perturbation to order ε (as explicitly done in Refs. [55] and [12]), but in this way the normal frequencies are still required to be distinct in order to apply Pöschel's result, whereas we do not need such a condition. On the other hand the case of possibly non-distinct normal frequencies (of order 1) has been dealt with in Refs. [3], [4], [60], [57] and [58] only under the usual Diophantine condition.

Theorem 1. *Consider the Hamiltonian (1.1). Suppose $\boldsymbol{\beta}_0$ to be such that $\partial_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0) = 0$, and assume that the eigenvalues of the matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ are all different from zero (that is $a_i \neq 0$ for all $i = 1, \dots, s$). Let $\mathcal{A} \subset \mathbb{R}^r$ be any open set of the non-resonant action variable space. Then for any $\delta > 0$ there are ε_0 small enough and a subset $\mathcal{A}_* \subset \mathcal{A}$ such that if $|\varepsilon| < \varepsilon_0$ and $\boldsymbol{\omega} \in \mathcal{A}_*$ the system described by the Hamiltonian (1.1) admits a lower-dimensional torus of the form*

$$\begin{cases} \boldsymbol{\alpha} = \boldsymbol{\psi} + \mathbf{a}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon), \\ \boldsymbol{\beta} = \boldsymbol{\beta}_0 + \mathbf{b}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon), \\ \mathbf{A} = \boldsymbol{\omega} + \left(\boldsymbol{\omega} \cdot \partial_{\boldsymbol{\psi}}\right) \mathbf{a}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon), \\ \mathbf{B} = \left(\boldsymbol{\omega} \cdot \partial_{\boldsymbol{\psi}}\right) \mathbf{b}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon), \end{cases} \tag{1.3}$$

with the functions $\mathbf{a}$ and $\mathbf{b}$ vanishing at $\varepsilon = 0$, analytic and periodic in $\boldsymbol{\psi}$, and the Lebesgue measure of the set $\mathcal{A} \setminus \mathcal{A}^*$ is less than δ. The parameterisation in (1.3) is such that $\boldsymbol{\psi} = \boldsymbol{\psi}_0 + \boldsymbol{\omega} t$ describes a linear flow on $\mathbb{T}^r$. In the case of maximal tori ($r = d$) the subset of

[1] We prefer using the word 'distinct', instead of 'non-degenerate', just to avoid confusion, as we are calling the normal frequencies 'degenerate' when they vanish for $\varepsilon = 0$.

phase space which is filled by invariant tori has complement whose Lebesgue measure is less than $C\delta$, for some positive constant C.

We can interpret Theorem 1 by saying that, in the presence of perturbations, the resonant tori are destroyed in general, but some of them survive. They are determined by the stationary points of the *potential function* $f_{\mathbf{0}}(\boldsymbol{\beta})$. Let $\boldsymbol{\beta}_0$ one of these stationary points: under the (generic) non-degeneracy conditions assumed on the eigenvalues of the matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$, we can say that, in correspondence of such a point $\boldsymbol{\beta}_0$, there is a conserved invariant lower-dimensional torus. The latter is either *hyperbolic* or *elliptic* or *of mixed type* according to the signs of the eigenvalues and of the perturbation parameter ε: it is elliptic if the eigenvalues have the same sign as ε, hyperbolic if they have opposite sign with respect to ε, and of mixed type otherwise. The rotation vector of such a torus will be found to satisfy some Diophantine conditions involving also the normal frequencies. In particular we shall find that any maximal torus with rotation vector $\boldsymbol{\omega}$ which is a Bryuno vector in $\mathbb{R}^d$ is conserved provided that ε is small enough (depending on $\boldsymbol{\omega}$). By *Bryuno vector* in $\mathbb{R}^r$ we mean a vector $\boldsymbol{\omega} \in \mathbb{R}^r$ such that

$$\sum_{n=0}^{\infty} \frac{1}{2^n} \log \frac{1}{\inf\limits_{0<|\boldsymbol{\nu}|\le 2^n} |\boldsymbol{\omega}\cdot\boldsymbol{\nu}|} < \infty. \tag{1.4}$$

For further properties of Bryuno vectors we refer to Section 2. Note that in Refs. [46] and [49] slightly different (but equivalent) conditions are found for the rotation vectors of the surviving elliptic tori, which can be expressed in terms of a suitable approximation function, first introduced by Rüssmann [48]; cf. the quoted references for further details.

Also concerning problem (B) our result does not exist in literature, and it represents the natural extension of Ref. [19] and [28] to the case of more general rotation vectors. In this case we still need the condition for the normal frequencies to be distinct, as in Ref. [28]. Such a condition could be weakened by assuming only that degeneracy is removed to some finite perturbation order; cf. Ref. [20]. Though, we shall impose on the rotation vector only the Bryuno condition (1.4), a condition much weaker than Kolmogorov's Diophantine condition for maximal tori and Mel′nikov condition for elliptic lower-dimensional tori, as usually assumed (cf. Refs. [55], [34] and [35]).

Theorem 2. *Let $\boldsymbol{\omega}$ be a vector in $\mathbb{R}^r$ satisfying the Diophantine condition (1.4). Suppose $\boldsymbol{\beta}_0$ to be such that $\partial_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0) = 0$, and assume that the eigenvalues of the matrix $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ are all different from zero and pairwise distinct (that is $a_i \neq 0$ for all $i = 1, \ldots, s$ and $a_i \neq a_j$ for all $1 \le i < j \le s$). Then there exists ε_0 and a set $\mathcal{E} \subset (-\varepsilon_0, \varepsilon_0)$, with a density point at the origin, such that for all $\varepsilon \in \mathcal{E}$ there is a lower-dimensional torus for the system described by the Hamiltonian (1.1) with rotation vector $\boldsymbol{\omega}$, which can be parameterised as (1.3), with $\boldsymbol{\psi} \in \mathbb{T}^d$ and the functions $\mathbf{a}$ and $\mathbf{b}$ vanishing at $\varepsilon = 0$, analytic and periodic in $\boldsymbol{\psi}$.*

A density point for $\mathcal{E}$ at the origin means that the relative Lebesgue measure of the set $\mathcal{E} \cap (-\varepsilon, \varepsilon)$, that is $\text{meas}(\mathcal{E} \cap (-\varepsilon, \varepsilon))/2\varepsilon$, tends to 1 as $\varepsilon \to 0$. We shall say also, in such a case, that the set $\mathcal{E}$ has large relative (Lebesgue) measure in $(-\varepsilon_0, \varepsilon_0)$.

If $\varepsilon > 0$ we can require $a_i \neq a_j$ for $i \neq j$ only for positive eigenvalues. If all eigenvalues are positive then the corresponding torus is elliptic; if all eigenvalues are negative then the corresponding torus is hyperbolic. In the first case the allowed values of ε form a Cantor set in $[0, \varepsilon_0)$ with large relative measure, in the second one all values in $[0, \varepsilon_0)$ are allowed. The obvious analogue holds for $\varepsilon < 0$.

For both theorems we shall give the proof in the case in which all eigenvalues of $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ are strictly positive and $\varepsilon > 0$, which is the difficult case. All the other cases can be obtained with trivial adaptations of the proof. Note also that the case of maximal tori can be obtained as a byproduct by setting $r = d$ in the following.

Notice that if do not require that degeneracy of the normal frequencies be removed to first order then the problem can become much harder. For instance if no condition at all is imposed on the perturbation only partial results exist, and only for the case $s = 1$ and $\boldsymbol{\omega}$ a Diophantine rotation vector; cf. Refs. [10] and [11] (see also Ref. [20]).

Finally we mention two related items which should deserve to be investigated. In [52] and [53], Sevryuk considers integrable Hamiltonian systems satisfying nondegeneracy conditions much weaker than convexity, and gives a discussion on the persistence under perturbation of maximal and lower dimensional tori based on a method introduced by Herman: suitable parameters are first introduced to remove degeneracies and then eliminated by using results on Diophantine approximations of dependent quantities. A natural question is in how far similar methods could apply in the current setting. Also the issue of quasi-periodic bifurcation (cf. for instance Refs. [31] and [6] and the surveys [7], [32] and [33]) could be addressed for the models considered in the this paper. In our case the perturbation parameter changes the normal linear behaviour of the lower-dimensional tori, which is parabolic at $\varepsilon = 0$; it would be interesting to study other bifurcation phenomena.

The paper is organised as follows. In Section 2 we introduce the Bryuno vectors, and we briefly review some properties of theirs, which will be used in the forthcoming analysis. Then Sections 3 and 4 are devoted to the proof of Theorem 1 and of Theorem 2, respectively. In principle, the proofs heavenly rely, both for notations and results, on Ref. [28], and, in those sections, we confine ourselves to give full details only for the parts which are really different. However, in order to make the paper self-contained, in Appendices A1 and A2 we briefly recall the basic notions of the tree formalism, with the aim of introducing notations and symbols which are used all along the present paper; thus, the reader who is not acquainted with Ref. [28] will find there all the technical ingredients to follow the discussion. Some more technical aspects of the proof of Theorem 1 are deferred to Appendix A3. Finally in Appendix A4 we briefly discuss how the analysis can be adapted to deal with more general convex Hamiltonian systems.

2. The Bryuno condition

Given $\boldsymbol{\omega} \in \mathbb{R}^2$ set $\omega \equiv \min\{|\omega_1|, |\omega_2|\}/\max\{|\omega_1|, |\omega_2|\}$. Let $\{q_n\}_{n=0}^{\infty}$ be the denominators of the convergents of ω.

The *Bryuno function* $\mathcal{B}(\omega)$ is defined as the solution of the functional equation [59]

$$\begin{cases} \mathcal{B}(\omega + 1) = \mathcal{B}(\omega), \\ \mathcal{B}(\omega) = -\log\omega + \omega\, \mathcal{B}(1/\omega), \quad \text{if } \omega \in (0,1). \end{cases} \tag{2.1}$$

Define

$$D(\boldsymbol{\omega}) \equiv \sum_{n=0}^{\infty} \frac{\log q_{n+1}}{q_n}. \tag{2.2}$$

Then it is easy to show that $\mathcal{B}(\omega) < \infty$ if and only if $D(\boldsymbol{\omega}) < \infty$ [59].

Given $\boldsymbol{\omega} \in \mathbb{R}^r$ and $n \in \mathbb{Z}_+$ set

$$\alpha_n(\boldsymbol{\omega}) = \inf_{0<|\boldsymbol{\nu}|\le 2^n} |\boldsymbol{\omega}\cdot\boldsymbol{\nu}| , \tag{2.3}$$

and define the *generalised Bryuno function* as

$$B(\boldsymbol{\omega}) = \sum_{n=0}^{\infty} \frac{1}{2^n} \log \frac{1}{\alpha_n(\boldsymbol{\omega})}. \tag{2.4}$$

Definition 1. *We shall call* $\mathfrak{B}_r = \{\boldsymbol{\omega} \in \mathbb{R}^r : B(\boldsymbol{\omega}) < \infty\}$ *the set of Bryuno vectors in* $\mathbb{R}^r$*. For any open set* $\Omega \subset \mathbb{R}^r$ *we call* $\mathfrak{B}_r(\Omega)$ *the set of Bryuno vectors in* Ω*.*

The reason for this terminology relies on the following result.

Lemma 1. *For* $r = 2$ *one has* $\boldsymbol{\omega} \in \mathfrak{B}_2$ *if and only if* $D(\boldsymbol{\omega}) < \infty$*.*

Proof. Given $\boldsymbol{\omega} \in \mathbb{R}^2$ assume for notation simplicity $0 < \omega_2 < \omega_1$. Call $\alpha = \omega_2/\omega_1$, and set $\boldsymbol{\omega} = \omega_1\boldsymbol{\omega}_0$, with $\boldsymbol{\omega}_0 = (1,\alpha)$, so that $0 < \alpha < 1$ and $\log\alpha_n(\boldsymbol{\omega}) = \log\omega_1 + \log\alpha_n(\boldsymbol{\omega}_0)$. Consider the sequence of convergents $\{p_n/q_n\}_{n=1}^{\infty}$ for α [50]; one has $1/2q_{n+1} < |\alpha q_n - p_n| < 1/q_{n+1}$, and $|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}| > |\alpha q_n - p_n|$ for all $|\nu_2| < q_{n+1}$, hence for all $|\boldsymbol{\nu}| < 2q_{n+1}$.

For each $n \ge 0$ define r_n and s_n such that $2^{r_n-1} < 2q_n \le 2^{r_n}$ and $r_n + s_n + 1 = r_{n+1}$. Hence for all $r_n \le r' \le r_n + s_n$ one has $\alpha_{r'}(\boldsymbol{\omega}_0) = |\alpha q_n - p_n|$, which implies

$$\frac{1}{4}\left(\frac{\log q_{n+1}}{q_n}\right) \le \sum_{r'=r_n}^{r_n+s_n} \frac{1}{2^{r'}} \log\frac{1}{\alpha_{r'}(\boldsymbol{\omega}_0)} \le \frac{\log 2}{q_n} + \frac{\log q_{n+1}}{q_n}, \tag{2.5}$$

so that, by using that $\sum_{n=0}^{\infty} q_n^{-1} < \infty$, one obtains that there exist two positive constants C_1 and C_2 such that $D(\boldsymbol{\omega})/4 - C_1 \le B(\boldsymbol{\omega}) \le D(\boldsymbol{\omega}) + C_2$, and the assertion follows. ■

The sequence $\{\alpha_n(\boldsymbol{\omega})\}_{n=1}^{\infty}$ is non-increasing, so that it converges to 0 monotonically as $n \to \infty$. By taking possibly a subsequence, we can always suppose $\alpha_{n+1}(\boldsymbol{\omega}) < \alpha_n(\boldsymbol{\omega})$, strictly.

Definition 2. *Set* $\mathbb{Z}_*^r = \mathbb{Z}^r \setminus \{\mathbf{0}\}$*, and define*

$$n(\boldsymbol{\nu}) = \left\{n \in \mathbb{Z}_+ : 2^{n-1} < |\boldsymbol{\nu}| \le 2^n\right\} = \inf\left\{n \in \mathbb{N} : |\boldsymbol{\nu}| \le 2^n\right\} \tag{2.6}$$

for any $\boldsymbol{\nu} \in \mathbb{Z}_*^r$*.*

For all $\boldsymbol{\nu} \in \mathbb{Z}_*^r$ one has, by definition, $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}| \ge \alpha_{n(\boldsymbol{\nu})}(\boldsymbol{\omega})$ and $2^{n(\boldsymbol{\nu})-1} < |\boldsymbol{\nu}| \le 2^{n(\boldsymbol{\nu})}$.

Definition 3. *Given a non-increasing sequence* $\{\alpha_n^*\}_{n=0}^{\infty}$ *converging to* 0*, define*

$$B^* = \sum_{n=0}^{\infty} \frac{1}{2^n} \log\frac{1}{\alpha_n^*}, \qquad \Gamma_p^* \equiv \sum_{n=0}^{\infty} \alpha_n^* 2^{np}, \tag{2.7}$$

for $p \in \mathbb{Z}_+$. One has $\Gamma^*_{p+1} > \Gamma^*_p$ for all $p \in \mathbb{Z}_+$ such that Γ^*_p is finite.

Lemma 2. *Let $\Omega \subset \mathbb{R}^r$ be an open set, and let $\{\alpha^*_n\}_{n=0}^\infty$ be a decreasing sequence converging to zero such that one has $B^* < \infty$ and $\Gamma^*_r = C_0$ for some finite constant C_0. Call $\Omega(C_0)$ the subset of Bryuno vectors in $\mathfrak{B}_r(\Omega)$ such that $\alpha_n(\boldsymbol{\omega}) \geq \alpha^*_n$ for all $n \geq 1$. Then the Lebesgue measure of the set $\Omega^c(C_0) = \Omega \setminus \Omega(C_0)$ is bounded proportional to C_0.*

Proof. The measure of the set $\Omega^c(C_0)$ can be bounded by

$$\begin{aligned} \mathrm{meas}(\Omega^c(C_0)) &\leq \text{const.} \sum_{n=0}^{\infty} \sum_{2^{n-1} < |\boldsymbol{\nu}| \leq 2^n} \frac{\alpha^*_n}{|\boldsymbol{\nu}|} \\ &\leq \text{const.} \sum_{n=0}^{\infty} \alpha^*_n 2^{nr} 2^{-(n-1)} \leq \text{const.}\, \Gamma^*_{r-1} \leq \text{const.}\, C_0, \end{aligned} \tag{2.8}$$

so that the assertion follows. ∎

The *Diophantine vectors*, that is the vectors satisfying the usual Diophantine condition

$$|\boldsymbol{\omega} \cdot \boldsymbol{\nu}| > \frac{C_0}{|\boldsymbol{\nu}|^\tau}, \tag{2.9}$$

for all $\boldsymbol{\nu} \in \mathbb{Z}^r_*$ and for suitable positive constants C_0 and τ, are a particular case of Bryuno vectors, with $\alpha_n(\boldsymbol{\omega}) \geq 2^{-n\tau} C_0$. In such a case in order to have the convergence of the sum in (2.8), hence to apply Lemma 2, one must have $\tau > r - 1$, which is the condition for the set of Diophantine vectors to have full measure.

The condition $\Gamma^*_r = C_0$ motivates us to introduce a new sequence $\{\gamma_n(\boldsymbol{\omega})\}_{n=1}^\infty$, with $\gamma_n(\boldsymbol{\omega}) = C_0^{-1} \alpha_n(\boldsymbol{\omega})$, such that, by setting $\alpha^*_n = C_0 \gamma^*_n$ and defining

$$\begin{aligned} &\overline{\Gamma}_p(\boldsymbol{\omega}) = \sum_{n=0}^{\infty} \gamma_n(\boldsymbol{\omega}) 2^{np}, \qquad \overline{\Gamma}^*_p = \sum_{n=0}^{\infty} \gamma^*_n 2^{np} = 1, \\ &\overline{B}(\boldsymbol{\omega}) = \sum_{n=0}^{\infty} \frac{1}{2^n} \log \frac{1}{\gamma_n(\boldsymbol{\omega})}, \qquad \overline{B}^* = \sum_{n=0}^{\infty} \frac{1}{2^n} \log \frac{1}{\gamma^*_n}, \end{aligned} \tag{2.10}$$

one has $\overline{\Gamma}_r(\boldsymbol{\omega}) \geq \overline{\Gamma}^*_r$ and $\overline{B}(\boldsymbol{\omega}) \leq \overline{B}^*$ for all $\Omega \subset \mathbb{R}^r$ and all $\boldsymbol{\omega} \in \Omega(C_0)$.

Note that if $|\boldsymbol{\omega} \cdot \boldsymbol{\nu}| < C_0 \gamma_n(\boldsymbol{\omega})$ then $|\boldsymbol{\nu}| > 2^n$. This is easily checked by contradiction: if $|\boldsymbol{\nu}| \leq 2^n$ then $|\boldsymbol{\omega} \cdot \boldsymbol{\nu}| \geq C_0 \gamma_n(\boldsymbol{\omega})$.

3. Fixing the perturbation parameter: proof of Theorem 1

In this section we deal with the construction of the function $\mathbf{h}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon) = (\mathbf{a}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon), \mathbf{b}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon))$ of Theorem 1. We follow very closely Ref. [28] (and Ref. [24]), by confining ourselves to show where the analysis differs. Also notations which are not defined below are meant the same as in Ref. [28].

In particular the formal expansion is described through the *tree formalism* introduced in Ref. [28], Section 2. The reader can refer to that paper for the notations about trees, and for the definitions of mode labels, momenta, node factors and propagators. However, in order to make self-contained the present paper, we briefly recall notations and definitions in Appendices A1 and A2. Only for the proofs of the more technical lemmata some reference is made to Ref. [28], just to avoid repetitions of identical arguments.

The first step consists in providing a diagrammatic representation of the formal series expansion for the function $\mathbf{h} = \mathbf{h}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon)$. This is recalled in Appendix A1. However the formal expansion is plagued by the small divisors problem, which prevents us from controlling the convergence. Then a different convergent expansion has to be looked for. This is envisaged through multiscale analysis techniques typical of the renormalisation group theory: the lines of the trees receive a further label – the scale label –, and the propagators of the lines are recursively defined in function of the scales. This leads to a different diagrammatic representation, in terms of trees which a variant of the previous ones, and which we call renormalised trees. The basic steps in assigning the scale labels and constructing the renormalised trees are described in Appendix A2. The recursive definition of the new propagators will be given here, in order to outline the differences with respect to Ref. [28]. For a reader who is not acquainted with Ref. [28], a preliminary reading of Appendices A1 and A2 would be useful before attacking the following discussion.

The multiscale decomposition of the propagators is performed as in Ref. [28], Section 5, by using the C^∞ non-decreasing function defined as

$$\chi(x) = \begin{cases} 1, & \text{if } |x| < C_0^2/4, \\ 0, & \text{if } |x| > C_0^2, \end{cases} \tag{3.1}$$

with the only difference that now χ_n for $n \geq 0$ is defined as $\chi_n(x) = \chi(\beta^{-2}(\gamma_n^*)^{-2}(\boldsymbol{\omega})x)$, with $\beta = 1/4$ and the sequence $\{\gamma_n^*\}_{n=0}^\infty$ introduced in Section 2. We set also $\psi_n(x) = 1 - \chi_n(x)$ for $n \geq 0$.

Next, given a tree θ, with each line ℓ of θ we associate a *scale* label $[n_\ell]$, with $n_\ell \in \{-1\} \cup \{0\} \cup \mathbb{N}$, and a *renormalised propagator*, or simply propagator, g_ℓ (still to be defined). Then we say that the propagator of the line ℓ is on scale $[n_\ell]$.

We introduce the notions of *clusters* and *self-energy clusters* as in Ref. [28], Section 5; cf. also Appendix A2. The *self-energy value* $\mathcal{V}_T(x; \varepsilon, \boldsymbol{\omega})$ and the *tree value* $\mathrm{Val}(\theta)$ are defined in terms of node factors and renormalised propagators, according to Ref. [28], equations (5.8) and (5.11). Therefore one defines the self-energy value as

$$\mathcal{V}_T(x; \varepsilon, \boldsymbol{\omega}) = \frac{\varepsilon^k}{k!} \Big(\prod_{\mathfrak{v} \in V(T)} F_{\mathfrak{v}} \Big) \Big(\prod_{\ell \in \Lambda(T)} g_\ell \Big),$$

where $V(T)$ and $\Lambda(T)$ denote the number of nodes and lines, respectively, of the self-energy cluster T, and $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}$, if $\boldsymbol{\nu}$ is the momentum of the line entering T, and the tree value as

$$\mathrm{Val}(\theta) = \frac{\varepsilon^k}{k!} \Big(\prod_{\mathfrak{v} \in V(\theta)} F_{\mathfrak{v}} \Big) \Big(\prod_{\ell \in \Lambda(\theta)} g_\ell \Big),$$

where $V(\theta)$ and $\Lambda(\theta)$ denote the number of nodes and lines, respectively, of the tree θ; cf. Appendix A1 for the definition of the node factors $F_{\mathfrak{v}}$.

Of course one has $\mathrm{d}\mathcal{V}_T/\mathrm{d}\boldsymbol{\omega} = \partial_x \mathcal{V}_T \partial_{\boldsymbol{\omega}} x + \partial_{\boldsymbol{\omega}} \mathcal{V}_T$, and $\mathrm{d}\mathcal{V}_T/\mathrm{d}\varepsilon = \partial_\varepsilon \mathcal{V}_T$. Note that here and henceforth, with respect to Ref. [28], we are making explicit the dependence of all quantities on $\boldsymbol{\omega}$, as we are interested also in changing $\boldsymbol{\omega}$ for fixed ε.

In terms of the self-energy values we can define the *self-energy matrices*

$$\begin{aligned} \mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega}) &= \sum_{p=0}^{n} \mathcal{M}^{[n]}(x;\varepsilon,\boldsymbol{\omega}), \\ \mathcal{M}^{[n]}(x;\varepsilon,\boldsymbol{\omega}) &= \left(\prod_{p=0}^{n} \chi_p(\Delta^{[p]}(x;\varepsilon,\boldsymbol{\omega})) \right) \sum_{k=1}^{\infty} \sum_{T \in \mathcal{S}^{\mathcal{R}}_{k,n-1}} \mathcal{V}_T(x;\varepsilon,\boldsymbol{\omega}), \end{aligned} \tag{3.2}$$

where $\mathcal{S}^{\mathcal{R}}_{k,n}$ denotes the set of *renormalised self-energy clusters* of degree k and scale $[n]$ (renormalised means that they do not contain any other self-energy clusters). Such matrices are formally Hermitian (cf. Lemma 2 in Ref. [28]), so that they admit d real eigenvalues, which we denote by $\lambda_1^{[n]}(x;\varepsilon,\boldsymbol{\omega})$, …, $\lambda_d^{[n]}(x;\varepsilon,\boldsymbol{\omega})$.

The matrices $\mathcal{M}^{[n]}(x;\varepsilon,\boldsymbol{\omega})$ in (3.2) can be written as

$$\mathcal{M}^{[n]}(x;\varepsilon,\boldsymbol{\omega}) = \begin{pmatrix} \mathcal{M}^{[n]}_{\alpha\alpha}(x;\varepsilon,\boldsymbol{\omega}) & \mathcal{M}^{[n]}_{\alpha\beta}(x;\varepsilon,\boldsymbol{\omega}) \\ \mathcal{M}^{[n]}_{\beta\alpha}(x;\varepsilon,\boldsymbol{\omega}) & \mathcal{M}^{[n]}_{\beta\beta}(x;\varepsilon,\boldsymbol{\omega}) \end{pmatrix}, \tag{3.3}$$

where the labels α and β run over $\{1,\ldots,r\}$ and $\{r+1,\ldots,d\}$, respectively.

With respect to Ref. [28], we slightly change the definition of the *propagator divisors* for $n \geq 0$ (cf. Definition 6 in Ref. [28]); see also Ref. [21]. We set

$$\Delta^{[n]}(x;\varepsilon,\boldsymbol{\omega}) = \left(\frac{1}{d} \sum_{j=1}^{d} \frac{1}{(x^2 - \underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega}))^2} \right)^{-1/2}, \tag{3.4}$$

and define recursively the *propagators* $g_\ell = g^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon,\boldsymbol{\omega})$, by setting

$$g^{[n]}(x;\varepsilon,\boldsymbol{\omega}) = \left(\prod_{p=0}^{n-1} \chi_p(\Delta^{[p]}(x;\varepsilon,\boldsymbol{\omega})) \right) \psi_n(\Delta^{[n]}(x;\varepsilon,\boldsymbol{\omega})) \left(x^2 - \mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega}) \right)^{-1}, \tag{3.5}$$

where the *self-energies* $\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})$ are defined implicitly as $\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega}) = \lambda_j^{[n]}(\sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,$ $\boldsymbol{\omega})$ for $n \geq 1$, with $\underline{\lambda}_j^{[0]}(\varepsilon,\boldsymbol{\omega}) = \varepsilon a_{j-r}$ for $j = r+1,\ldots,d$ and $\underline{\lambda}_j^{[0]}(\varepsilon,\boldsymbol{\omega}) = 0$ for $j = 1,\ldots,r$. The function $x \to \lambda_j^{[n]}(x;\varepsilon,\boldsymbol{\omega})$ will be showed to be a contraction, so that the definition will turn out to be well-posed.

Therefore if a line ℓ is on scale $[n]$, with $n \geq 1$, such that $g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon,\boldsymbol{\omega}) \neq 0$ one has

$$\begin{aligned} &\min_{1\leq j\leq d} \left| (\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell)^2 - \underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega}) \right| \geq \frac{C_0^2}{4\sqrt{d}} \beta^2 (\gamma_n^*)^2, \\ &\min_{1\leq j\leq d} \left| (\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell)^2 - \underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega}) \right| \leq C_0^2 \beta^2 (\gamma_p^*)^2, \qquad 0 \leq p \leq n-1, \end{aligned} \tag{3.6}$$

and β is chosen in such a way to make uninfluential the small changes of the propagators when shifting the lines in order to exploit the cancellations discussed in Ref. [28], Appendix

A3 (see the proof of Lemma 5 below). If a line ℓ is on scale [0] and $g^{[0]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon,\boldsymbol{\omega})\neq 0$ the condition (3.6) has to be replaced with $\min_{1\le j\le d}|(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell)^2-\underline{\lambda}_j^{[0]}(\varepsilon,\boldsymbol{\omega})|\ge C_0^2\beta^2(\gamma_0^*)^2/4\sqrt{d}$.

The advantage of the definition (3.4) for the propagator divisors is that on the one hand, as (3.6) shows, the propagator divisor behaves essentially the minimum among the quantities $x^2-\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})$, and on the other hand it is a smooth function of both ε and $\boldsymbol{\omega}$ and no discontinuity appears when the self-energies cross each others.

Note also that, contrary to what was done in Ref. [28] (and to what will be done in next section), we do not distinguish here ultraviolet and infrared resummations. This would be possible, but it would not introduce any further simplification. In fact, the main reason why it was introduced in Ref. [28], was to avoid self-energy crossings (a problem which is absent in the present case, as we shall see): in the ultraviolet region the small divisors can be bounded as the un-resummed ones, and in the infrared region they are well-separated. In next section, we shall proceed as in Ref. [28], as we shall need for the self-energies to be distinct in order to deal with the second Mel'nikov conditions.

We define the renormalised expansion for $\mathbf{h}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)=(\mathbf{a}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon),\mathbf{b}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon))$, by setting

$$\begin{aligned}&\mathbf{h}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)=\sum_{\boldsymbol{\nu}\in\mathbb{Z}^r}\mathrm{e}^{i\boldsymbol{\nu}\cdot\boldsymbol{\psi}}\,\mathbf{h}_{\boldsymbol{\nu}}(\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon),\qquad \mathbf{h}_{\boldsymbol{\nu}}(\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)\equiv\mathbf{h}_{\boldsymbol{\nu}},\\ &\mathbf{h}_{\boldsymbol{\nu}}=(h_{\boldsymbol{\nu},1},\ldots,h_{\boldsymbol{\nu},d}),\qquad h_{\boldsymbol{\nu},\gamma}=\sum_{k=1}^{\infty}\sum_{\theta\in\Theta^{\mathcal{R}}_{k,\nu,\gamma}}\mathrm{Val}(\theta),\end{aligned}\tag{3.7}$$

where the set of trees $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu},\gamma}$ and the tree values are defined as in Ref. [28], Definition 5 and equation (5.11); see also Appendix A2 below.

We shall impose the following Diophantine conditions:

$$\begin{aligned}&\left|\boldsymbol{\omega}\cdot\boldsymbol{\nu}\pm\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})}\right|\ge C_0\gamma^*_{n(\boldsymbol{\nu})},\\ &\left|\boldsymbol{\omega}\cdot\boldsymbol{\nu}\pm\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})}\pm\sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})}\right|\ge C_0\gamma^*_{n(\boldsymbol{\nu})}\end{aligned}\tag{3.8}$$

for all $i,j=1,\ldots,d$, for all $\boldsymbol{\nu}\in\mathbb{Z}^r_*$ and for all $n\ge 0$. We shall refer to conditions (3.8) as to the *first Mel′nikov conditions* (first line) and the *second Mel′nikov conditions* (second line).

Definition 4 *Given $C_0\in\mathbb{R}_+$ and an open set $\Omega\subset\mathbb{R}^r$ call $\Omega_*(C_0)\subset\Omega$ the set of Bryuno vectors in $\Omega(C_0)$ satisfying all the conditions (3.8).*

Hence the vectors $\boldsymbol{\omega}\in\Omega_*(C_0)$ verify the condition $\gamma_n(\boldsymbol{\omega})\ge\gamma_n^*$ for all $n\ge 0$, with $C_0\gamma_n^*=\alpha_n^*$ and the sequence $\{\alpha_n^*\}_{n=0}^{\infty}$ defined as in Lemma 2, and the first and second Mel′nikov conditions (3.8).

Lemma 3. *Call $N_n(\theta)$ the set of lines in $\Lambda(\theta)$ which are on scale $[n]$. One has*

$$N_n(\theta)\le K\,2^{-n}M(\theta),\qquad M(\theta)=\sum_{\mathfrak{v}\in V(\theta)}|\boldsymbol{\nu}_{\mathfrak{v}}|,\tag{3.9}$$

for a suitable constant K. One can take $K = 2$.

Proof. First of all note that one can have $N_n(\theta) \geq 1$ only if $M(\theta) \geq 2^{n-1}$. Indeed if a line ℓ is on scale $[n]$ then there exists $i \in \{1, \ldots, d\}$ such that

$$C_0 \gamma^*_{n-1} > C_0 \beta \gamma^*_{n-1} \geq \left| |\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell| - \sqrt{\underline{\lambda}_i^{[n-1]}(\varepsilon, \boldsymbol{\omega})} \right| > C_0 \gamma^*_{n(\boldsymbol{\nu}_\ell)}, \tag{3.10}$$

so that $n(\boldsymbol{\nu}_\ell) \geq n$. Then one must have $|\boldsymbol{\nu}_\ell| > 2^{n(\boldsymbol{\nu}_\ell)-1} \geq 2^{n-1}$, hence $M(\theta) \geq |\boldsymbol{\nu}_\ell| > 2^{n-1}$, thence $K 2^{-n} M(\theta) \geq 1$ if $K \geq 2$.

Then one proves the bound $N_n(\theta) \leq \max\{2^{-n} K M(\theta) - 1, 0\}$ for all $n \geq 0$, by induction on the number of vertices of the tree. The only case which requires a different discussion with respect to Ref. [28], Appendix A3, is the one in which the root line ℓ is on scale $[n]$ and exits a cluster on scale $[n_T]$, which has only one entering line, say ℓ', on scale $[n']$, with $n' \geq n$. In such a case one has $n_T < n$ of course, and, for suitable i and j,

$$\begin{aligned} \left| |\boldsymbol{\omega} \cdot \boldsymbol{\nu}| - \sqrt{\underline{\lambda}_i^{[n-1]}(\varepsilon, \boldsymbol{\omega})} \right| &\leq C_0 \beta \gamma^*_{n-1}, \\ \left| |\boldsymbol{\omega} \cdot \boldsymbol{\nu}'| - \sqrt{\underline{\lambda}_j^{[n-1]}(\varepsilon, \boldsymbol{\omega})} \right| &\leq C_0 \beta \gamma^*_{n-1}, \end{aligned} \tag{3.11}$$

where $\boldsymbol{\nu} = \boldsymbol{\nu}_\ell$ and $\boldsymbol{\nu}' = \boldsymbol{\nu}_{\ell'}$, so that, for suitable $\eta, \eta' \in \{\pm 1\}$,

$$\left| \boldsymbol{\omega} \cdot (\boldsymbol{\nu} - \boldsymbol{\nu}') + \eta \sqrt{\underline{\lambda}_i^{[n-1]}(\varepsilon, \boldsymbol{\omega})} + \eta' \sqrt{\underline{\lambda}_j^{[n-1]}(\varepsilon, \boldsymbol{\omega})} \right| < C_0 \gamma^*_{n-1}, \tag{3.12}$$

which by the Diophantine conditions (3.8) implies $n(\boldsymbol{\nu} - \boldsymbol{\nu}') \geq n$, hence one finds $M(T) \geq |\boldsymbol{\nu} - \boldsymbol{\nu}'| > 2^{n-1}$, if $M(T) = \sum_{\mathfrak{v} \in V(T)} |\boldsymbol{\nu}_\mathfrak{v}|$. Call θ' the tree having ℓ' as root line. Then by the inductive hypothesis $N_n(\theta) = 1 + N_n(\theta') \leq 1 + \max\{2^{-n} K M(\theta') - 1, 0\}$. If the maximum is 0 the bound is trivially satisfied, because in such a case $N_n(\theta) = 1$ and we have seen that in order to have a line on scale $[n]$ one needs $M(\theta) > 2^{n-1}$. Otherwise one has $N_n(\theta) \leq 1 + 2^{-n} K M(\theta') - 1 \leq 2^{-n} K M(\theta) - 1 + (1 - 2^{-n} K M(T))$, where $2^{-n} K M(T) \geq 1$ by the inequality $|\boldsymbol{\nu} - \boldsymbol{\nu}'| > 2^{n-1}$, provided that one takes $K \geq 2$. ∎

Lemma 4. *Call $N_n(T)$ the set of lines in $\Lambda(T)$ which are on scale $[n]$, for $n \leq n_T$. One has*

$$M(T) = \sum_{\mathfrak{v} \in V(T)} |\boldsymbol{\nu}_\mathfrak{v}| > 2^{n_T - 1}, \qquad N_n(T) \leq K \, 2^{-n} M(T), \tag{3.13}$$

with the same constant K as in (3.9).

Proof. The first bound in (3.13) can be proved by *reductio ad absurdum* as in Ref. [28], while the proof of the second one is based on the same argument used for proving Lemma 3 (cf. Ref. [28], Appendix A3, for further details). ∎

Another difference with respect to Ref. [28] relies in discussing the change of scale of the lines when performing the cancellations inside the families $\mathcal{F}_T$, when looking for bounds on the entries of the matrices $\mathcal{M}^{[n]}(x; \varepsilon, \boldsymbol{\omega})$.

Lemma 5. *Assume that the propagators $g^{[p]}(x; \varepsilon, \boldsymbol{\omega})$ can be uniformly bounded for all $0 \leq p \leq n-1$ as*

$$\left| g^{[p]}(x; \varepsilon, \boldsymbol{\omega}) \right| \leq K_1 C_0^{-2} (\gamma^*_p)^{-K_2}, \tag{3.14}$$

for some p-independent constants K_1 and K_2. Assume also that ε is small enough. Then, with the notations (3.3), one has

$$\begin{aligned} \left\| \mathcal{M}^{[n]}_{\alpha\alpha}(x;\varepsilon,\boldsymbol{\omega}) \right\| &\le B e^{-\kappa_1 2^n} \min\{\varepsilon^2, \varepsilon x^2\}, \\ \left\| \mathcal{M}^{[n]}_{\alpha\beta}(x;\varepsilon,\boldsymbol{\omega}) \right\| &\le B e^{-\kappa_1 2^n} \min\{\varepsilon^2, \varepsilon^{3/2}|x|\}, \\ \left\| \mathcal{M}^{[n]}_{\beta\beta}(x;\varepsilon,\boldsymbol{\omega}) \right\| &\le B e^{-\kappa_1 2^n} \varepsilon^2, \end{aligned} \tag{3.15}$$

for suitable n-independent constants B and κ_1.

Proof. Again we only discuss the differences with respect to Ref. [28]. First we show that no cancellation is needed for self-energy clusters T with $C_0\gamma^*_{n(M(T))} \le 4|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$, if $\boldsymbol{\nu}$ is the momentum flowing through the entering line of T. Note that we can extract from the self-energy value a factor $e^{-\kappa_0 M(T)/4} \le e^{-\kappa_0 2^{n(M(T))}/8}$. If we set $2^{-n}\log 1/\alpha^*_n = a_n$, we have $\lim_{n\to\infty} a_n = 0$ (because $B^* < \infty$), hence for $\boldsymbol{\omega}\cdot\boldsymbol{\nu}$ small enough $e^{-\kappa_0 2^{n(M(T))}/8} \le (C_0\gamma^*_{n(M(T))})^{\kappa_0/8a_{n(M(T))}} \le C_0^2(\gamma^*_{n(M(T))})^2 \le 16|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|^2$.

Hence we need the cancellations only for self-energy clusters T with $C_0\gamma^*_{n(M(T))} > 4|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$ if $\boldsymbol{\nu}$ is the momentum flowing through the entering line of T. In such a case one can reason as follows. For any line $\ell\in\Lambda(T)$ and for any $n\le n_\ell$ one has, by the Diophantine conditions (3.8), $||\boldsymbol{\omega}\cdot\boldsymbol{\nu}^0_\ell| - \sqrt{\underline{\lambda}^{[n]}_j(\varepsilon,\boldsymbol{\omega})}| \ge C_0\gamma^*_{n(\boldsymbol{\nu}^0_\ell)}$, if $\boldsymbol{\nu}^0_\ell$ is defined as

$$\boldsymbol{\nu}^0_\ell = \sum_{\substack{\mathfrak{w}\in V(T)\\ \mathfrak{w}\preceq\mathfrak{v}}} |\boldsymbol{\nu}_{\mathfrak{v}}|, \qquad \ell\equiv\ell_{\mathfrak{v}}. \tag{3.16}$$

On the other hand one has $|\boldsymbol{\nu}^0_\ell| \le M(T)$, so that $C_0\gamma^*_{n(\boldsymbol{\nu}^0_\ell)} \ge C_0\gamma^*_{n(M(T))} > 4|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$. Therefore we can bound

$$2\left||\boldsymbol{\omega}\cdot\boldsymbol{\nu}^0_\ell| - \sqrt{\underline{\lambda}^{[n]}_j(\varepsilon,\boldsymbol{\omega})}\right| \ge \left||\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell| - \sqrt{\underline{\lambda}^{[n]}_j(\varepsilon,\boldsymbol{\omega})}\right| \ge \frac{1}{2}\left||\boldsymbol{\omega}\cdot\boldsymbol{\nu}^0_\ell| - \sqrt{\underline{\lambda}^{[n]}_j(\varepsilon,\boldsymbol{\omega})}\right|. \tag{3.17}$$

This implies the following. When considering a family $\mathcal{F}_T$, a line $\ell\in\Lambda(T)$ with momentum $\boldsymbol{\nu}_\ell$ can be on a scale $[n_\ell]$ such that $g^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}^0_\ell;\varepsilon,\boldsymbol{\omega}) = 0$. But in such a case it is obtained, by shifting the external lines of T, from a line with non-vanishing propagator, that is from a line for which (3.6) holds. Then, even if the bounds (3.6) can fail to hold, one still obtains bounds of the same form with the only difference that β^2 is replaced with $\beta^2/4$ in the first line and with $4\beta^2$ in the second line. In particular for $\beta = 1/4$ the inequalities (3.10) and (3.12) are still satisfied for all self-energy clusters in the family $\mathcal{F}_T$.

This shows that one can reason in Ref. [28] to deduce the bounds (3.15), which are of algebraic nature, and are due to symmetry properties of the self-energy matrices, that is $\mathcal{M}^{[\le n]T}(x;\varepsilon,\boldsymbol{\omega}) = \mathcal{M}^{[\le n]}(-x;\varepsilon,\boldsymbol{\omega})$ and $\mathcal{M}^{[\le n]\dagger}(x;\varepsilon,\boldsymbol{\omega}) = \mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega})$. ∎

Lemma 6. *Assume that the propagators $g^{[p]}(x;\varepsilon,\boldsymbol{\omega})$ can be uniformly bounded for all $0\le p\le n-1$ as in (3.14), for some p-independent constants K_1 and K_2. Assume also that ε is small enough. The matrices $\mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega})$ are differentiable in x, and one has*

$$\begin{aligned} &\left\| \mathcal{M}^{[\le n]}(x';\varepsilon,\boldsymbol{\omega}) - \mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega}) - \partial_x\mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega})\,(x'-x) \right\| = \varepsilon^2\, o(|x'-x|), \\ &\left\| \partial_x\mathcal{M}^{[\le n]}(x;\varepsilon,\omega) \right\| \le B\varepsilon^2, \end{aligned} \tag{3.18}$$

for a suitable positive constant B. Moreover $\lambda_j^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega})$ is also differentiable for all $j = 1,\ldots,d$, and one has

$$\left|\partial_x \lambda_j^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega})\right| \leq A\varepsilon^2, \tag{3.19}$$

for a suitable constant A.

Proof. The proof of (3.18) can be performed as in Ref. [28]; cf. in particular Section 6. Then property (3.19) follows from general properties of Hermitian matrices. One can refer to Ref. [28], Appendix A4, in the case in which the eigenvalues a_i are all distinct. Otherwise one can apply the results on non-analytic Hermitian matrices discussed in Ref. [37], Chapter 2, Section 6: one can rely on Rellich's theorem [47] to deduce differentiability of the eigenvalues and on Lidskiĭ's theorem [41] to obtain a bound on the derivative. ■

Lemma 7. *Assume that the propagators $g^{[p]}(x;\varepsilon,\boldsymbol{\omega})$ can be uniformly bounded for all $0 \leq p \leq n-1$ as in (3.14), for some p-independent constants K_1 and K_2. Assume also that ε is small enough. The self-energies $\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega})$ satisfy for all $0 \leq p \leq n$ and all $1 \leq j \leq d$ the closeness property*

$$\left|\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega}) - \underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})\right| \leq D\, \mathrm{e}^{-\kappa_1 2^p}\varepsilon^2 \tag{3.20}$$

and one has

$$\left|\lambda_j^{[p]}(x;\varepsilon,\boldsymbol{\omega})\right| \leq D\min\{\varepsilon^2, \varepsilon x^2\}, \qquad j = 1,\ldots,r, \tag{3.21}$$

for a suitable positive constant D.

Proof. The proof that $|\lambda_j^{[p]}(x;\varepsilon,\boldsymbol{\omega}) - \lambda_j^{[p-1]}(x;\varepsilon,\boldsymbol{\omega})| \leq A_1 \mathrm{e}^{-\kappa_1 2^p}\varepsilon^2$, for some constant A_1, can be performed as in Ref. [28], by using the bounds (3.15) and the fact that the matrices $\mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega})$ are Hermitian (see Ref. [28], Lemma 2). Then (3.20) follows immediately from the definition: indeed one has

$$\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega}) - \underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega}) = \lambda_j^{[p]}\big(\sqrt{\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega}\big) - \lambda_j^{[p-1]}\big(\sqrt{\underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega}\big), \tag{3.22}$$

so that, by using also (3.19),

$$\begin{aligned}
\Delta_j^{[p]}(\varepsilon,\boldsymbol{\omega}) &\equiv |\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega}) - \underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})| \\
&\leq |\lambda_j^{[p]}\big(\sqrt{\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega}\big) - \lambda_j^{[p]}\big(\sqrt{\underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega}\big)| \\
&\qquad + |\lambda_j^{[p]}\big(\sqrt{\underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega}\big) - \lambda_j^{[p-1]}\big(\sqrt{\underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega}\big)| \\
&\leq A\varepsilon^2 \left|\sqrt{\underline{\lambda}_j^{[p]}(\varepsilon,\boldsymbol{\omega})} - \sqrt{\underline{\lambda}_j^{[p-1]}(\varepsilon,\boldsymbol{\omega})}\right| + A_1 \mathrm{e}^{-\kappa_1 2^p}\varepsilon^2 \\
&\leq A'\varepsilon^{3/2}\Delta_j^{[p]}(\varepsilon,\boldsymbol{\omega}) + A_1 \mathrm{e}^{-\kappa_1 2^p}\varepsilon^2,
\end{aligned} \tag{3.23}$$

for some constant A', which gives (3.20).

Also the proof of (3.21) can be done as in Ref. [28]. ■

Lemma 8. *Assume that the propagators $g^{[p]}(x;\varepsilon,\boldsymbol{\omega})$ can be uniformly bounded for all $0 \leq p \leq n-1$ as in (3.14), for some p-independent constants K_1 and K_2. Assume also that ε is small enough. If $g^{[n]}(x;\varepsilon,\boldsymbol{\omega}) \neq 0$ then one has*

$$\min_{j=1,\ldots,d}\left|x^2 - \lambda_j^{[n]}(x;\varepsilon,\boldsymbol{\omega})\right| \geq \frac{1}{2}\min_{j=1,\ldots,d}\left|x^2 - \underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})\right|. \tag{3.24}$$

The same holds if $g^{[n]}(x;\varepsilon,\boldsymbol{\omega})=0$ but (3.6) are satisfied with β^2 replaced with $\beta^2/4$ in the first line and with $4\beta^2$ in the second line.

Proof. We can write for all $j'=1,\ldots,d$

$$\begin{aligned}\left|x^2-\lambda_{j'}^{[n]}(x;\varepsilon,\boldsymbol{\omega})\right| &\geq \left|x^2-\lambda_{j'}^{[n]}(\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega});\varepsilon,\boldsymbol{\omega})+\lambda_{j'}^{[n]}(\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega});\varepsilon,\boldsymbol{\omega})-\lambda_{j'}^{[n]}(x;\varepsilon,\boldsymbol{\omega})\right| \\ &\geq \left|x^2-\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})\right|-\left|\lambda_{j'}^{[n]}(\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega});\varepsilon,\boldsymbol{\omega})-\lambda_{j'}^{[n]}(x;\varepsilon,\boldsymbol{\omega})\right|, \end{aligned} \tag{3.25}$$

where we have used that $\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})=\lambda_{j'}^{[n]}(\sqrt{\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})};\varepsilon,\boldsymbol{\omega})$. Hence if $j'>r$ and $x>0$ we can bound

$$\left|\lambda_{j'}^{[n]}(x;\varepsilon,\boldsymbol{\omega})-\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})\right|\leq A\varepsilon^2\left||x|-\sqrt{\underline{\lambda}_{j'}^{[n1]}(\varepsilon,\boldsymbol{\omega})}\right|\leq A'\varepsilon^{3/2}\left|x^2-\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})\right|, \tag{3.26}$$

by Lemma 6 and in particular (3.19). If $j'>r$ and $x<0$ we can apply the same argument by using the symmetry property that $\lambda_{j'}^{[n]}(-x;\varepsilon,\boldsymbol{\omega})$ belongs to the spectrum if $\lambda_{j'}^{[n]}(x;\varepsilon,\omega)$ does (because $\mathcal{M}^{[\leq n]}(-x;\varepsilon,\boldsymbol{\omega})=(\mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega}))^T$, see Lemma 2, (ii), in Ref. [28]; cf. also the comments at the end of the proof of Lemma 5 above). If $j'\leq r$ then

$$\left|\lambda_{j'}^{[n]}(x;\varepsilon,\boldsymbol{\omega})-\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})\right|=\left|\lambda_{j'}^{[n]}(x;\varepsilon,\boldsymbol{\omega})\right|\leq A\varepsilon x^2=A\varepsilon\left|x^2-\underline{\lambda}_{j'}^{[n]}(\varepsilon,\boldsymbol{\omega})\right|, \tag{3.27}$$

by (3.21). By inserting (3.26) or (3.27) into (3.25), and choosing j' as the value of j minimising $|x^2-\lambda_j^{[p]}(x;\varepsilon,\boldsymbol{\omega})|$, then the assertion follows. ∎

Lemma 9. *Let $\boldsymbol{\omega}\in\Omega_*(C_0)$ and assume that ε is small enough, say $|\varepsilon|<\varepsilon_0$. Then the series (3.7) admits the bound $|h_{\boldsymbol{\nu},\gamma}|<H\,\mathrm{e}^{-\kappa|\boldsymbol{\nu}|}|\varepsilon|$ for suitable positive constants κ and H. One has $\varepsilon_0=O(C_0^2(\gamma_{m_0}^*)^2)$ with m_0 depending on κ_0.*

Proof. One can proceed as in Ref. [28]. Here we outline only the differences. As a consequence of Lemma 8, we can prove inductively that for all $n\geq 0$ the propagators with scales $[n]$ are bounded proportionally to $(\gamma_n^*)^{-K_2}$, and one finds, in particular, $K_2=2$. Then the product of propagators can be bounded by relying on Lemma 3 for the lines on scale $[n]$, with $n>m_0$, and bounding with $(\gamma_{m_0}^*)^{-2k}$ the propagators of all lines on scale $[n]$ for $n\leq m_0$. Therefore for any tree $\theta\in\Theta_{k,\boldsymbol{\nu},\gamma}^{\mathcal{R}}$ one has

$$\begin{aligned}\prod_{\ell\in\Lambda(\theta)}\left|g^{[n_\ell]}\right| &\leq (\text{const.})^k\,C_0^{-2k}\left(\frac{1}{\gamma_{m_0}^*}\right)^{2k}\prod_{n=m_0+1}^{\infty}\left(\frac{1}{\gamma_n^*}\right)^{2N_n(\theta)} \\ &\leq (\text{const.})^k\,C_0^{-2k}\left(\frac{1}{\gamma_{m_0}^*}\right)^{2k}\exp\left(K|\boldsymbol{\nu}|\sum_{n=m_0+1}\frac{1}{2^n}\log\frac{1}{\gamma_n^*}\right),\end{aligned} \tag{3.28}$$

and one can choose $m_0=m_0(\kappa_0)$, so the last exponential is less than $\mathrm{e}^{\kappa_0|\boldsymbol{\nu}|/4}$. By making use of the bound $\prod_{\mathfrak{v}\in V(\theta)}\mathrm{e}^{-\kappa|\boldsymbol{\nu}_\mathfrak{v}|}\leq\mathrm{e}^{-\kappa|\boldsymbol{\nu}|}$, this produces an overall factor $\mathrm{e}^{-\kappa_0|\boldsymbol{\nu}|/2}$. This completes the proof, and it gives $\varepsilon_0=O(C_0^2(\gamma_{m_0}^*)^2)$. ∎

Note that for Diophantine vectors satisfying the bound (2.8) one has $m_0=\tau\,O(\log 1/\kappa_0)$, and one obtains $\varepsilon_0=O(C_0^2)$, for fixed τ and C_0.

If we are interested in studying the conservation of a maximal torus with rotation vector $\boldsymbol{\omega}$ satisfying the Bryuno condition $B(\boldsymbol{\omega})<\infty$, we can use directly the sequence $\{\gamma_n(\boldsymbol{\omega})\}_{n=0}^{\infty}$ for the multiscale decomposition, without introducing a further sequence $\{\gamma_n^*\}_{n=0}^{\infty}$. Then the result stated in Lemma 9 holds with $\gamma_{m_0}(\boldsymbol{\omega})$ replacing $\gamma_{m_0}^*$.

An important remark is that, in the case of perturbations which are trigonometric polynomials of degree N in the bound (3.9) one can bound $M(\theta)\le kN$, and as consequence the product of propagators in (3.28) can be bounded as

$$\prod_{\ell\in\Lambda(\theta)}\left|g_\ell^{[n_\ell]}\right|\le\exp\left(2KNk\sum_{n=0}^{\infty}\frac{1}{2^n}\log\frac{1}{\alpha_n(\boldsymbol{\omega})}\right)=\mathrm{e}^{4NkB(\boldsymbol{\omega})},\tag{3.29}$$

which implies $\varepsilon_0=O(\mathrm{e}^{-4NB(\boldsymbol{\omega})})$. We can compare this result with the one found in Ref. [30] for maximal tori, where a bound of this kind with the factor 4 replaced with the likely optimal 2 was obtained. With the techniques described in this paper some further work is necessary in order to reach the factor 2; cf. for instance Ref. [2]. On the other hand an advantage with respect to Ref. [30], which relies on using Lie transforms for Hamiltonian flows, is that our techniques apply, essentially unchanged, not only to the case of flows, but also to the case of diffeomorphisms, as done for instance in Refs. [2] and [22], where the case of the standard map was explicitly treated.

Lemma 10. *For all $\boldsymbol{\omega},\boldsymbol{\omega}'\in\Omega_*(C_0)$ in the same direction $\mathbf{v}$ one has*

$$\begin{aligned}&\Big\|\mathcal{M}^{[\le n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')-\mathcal{M}^{[\le n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})-\\&\qquad\partial_{\mathbf{v}}\mathcal{M}^{[\le n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})\cdot(\boldsymbol{\omega}'-\boldsymbol{\omega})\Big\|=\varepsilon^2|\boldsymbol{\nu}|\,o(|\boldsymbol{\omega}'-\boldsymbol{\omega}|),\\&\Big\|\partial_{\mathbf{v}}\mathcal{M}^{[\le n]}(x;\varepsilon,\omega)\Big\|\le B\varepsilon^2|\boldsymbol{\nu}|,\end{aligned}\tag{3.30}$$

for a suitable constant B, and, as a consequence,

$$\left|\partial_{\mathbf{v}}\lambda_j^{[n]}(x;\varepsilon,\omega)\right|\le A\varepsilon^2|\boldsymbol{\nu}|,\qquad\left|\partial_{\mathbf{v}}\underline{\lambda}_j^{[n]}(\varepsilon,\omega)\right|\le A\varepsilon^2|\boldsymbol{\nu}|,\tag{3.31}$$

for a suitable constant A.

The proof is deferred to Appendix A3. Note that in fact it is enough to prove (3.30), because then property (3.31) follows by general properties of Hermitian matrices and by the definition of the self-energies $\underline{\lambda}_j^{[n]}(\varepsilon,\omega)$; cf. the comments in the proof of Lemma 6.

Lemma 11. *The Lebesgue measure of the set $\Omega\setminus\Omega_*(C_0)$ is bounded proportionally to C_0.*

Proof. Let us start with the first conditions in (3.8). We can reason as in Ref. [24], and write $\boldsymbol{\omega}=\alpha\boldsymbol{\nu}/|\boldsymbol{\nu}|+\boldsymbol{\beta}$, with $\boldsymbol{\beta}\cdot\boldsymbol{\nu}=0$. Then we define $\alpha(t)$, $t\in[-1,1]$, such that $F(t)=\alpha(t)|\boldsymbol{\nu}|\pm\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,(\alpha(t),\boldsymbol{\beta}))}=tC_0\gamma_{n(\boldsymbol{\nu})}^*$, so that $\mathrm{d}F/\mathrm{d}t=|\boldsymbol{\nu}|(1+O(\sqrt{\varepsilon}))\mathrm{d}\alpha/\mathrm{d}t=C_0\gamma_{n(\boldsymbol{\nu})}^*$; cf. the proofs of Lemma 3 and Lemma 4 in Ref. [24] for further details. Given $p\ge1$ we define $\Omega^{[p]}$ as the sets of $\boldsymbol{\omega}\in\Omega$ satisfying the conditions (3.8) for $n\le p$; we also set $\Omega^{[0]}=\Omega$. For each n, for fixed $\boldsymbol{\nu}$ and i, we call $I_n(i,\boldsymbol{\nu})$ the sets of $\boldsymbol{\omega}\in\Omega^{[n-1]}$ such that either $|\boldsymbol{\omega}\cdot\boldsymbol{\nu} \pm \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})}| < C_0\gamma^*_{n(\boldsymbol{\nu})}$ or $|\boldsymbol{\omega}\cdot\boldsymbol{\nu} \pm \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})} \pm \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})}| < C_0\gamma^*_{n(\boldsymbol{\nu})}$. We define in the same way the sets $J_n(i,\boldsymbol{\nu})$, with the only difference that the width of the sets is $2C_0\gamma^*_{n(\boldsymbol{\nu})}$ instead of $C_0\gamma^*_{n(\boldsymbol{\nu})}$. By the closeness property of Lemma 7, there is some $n_1(\boldsymbol{\nu}) = O(\log\log 1/\gamma^*_{n(\boldsymbol{\nu})})$ such that all the sets $I_n(i,\boldsymbol{\nu})$ fall inside $J_{n_1(\boldsymbol{\nu})}$ for $n \ge n_1(\boldsymbol{\nu})$. Therefore for all $\boldsymbol{\nu}\in\mathbb{Z}^r_*$, all $i = r+1,\dots,d$, and all $n \le n_1(\boldsymbol{\nu})$ we have to exclude all values of $\boldsymbol{\omega}\in\Omega^{[n-1]}$ which fall inside the set $J_n(i,\boldsymbol{\nu})$; we refer to Ref. [23], Section 7, for details. Note that $\boldsymbol{\omega}\in\mathfrak{B}_r$ implies $n_1(\boldsymbol{\nu}) \le Cn(\boldsymbol{\nu})$, for some constant C. Hence we can bound the measure of the set of excluded values by a constant times

$$\begin{aligned}\text{const.} \sum_{i=r+1}^{d} C_0 \sum_{\boldsymbol{\nu}\in\mathbb{Z}^r} \sum_{n=1}^{n_1(\boldsymbol{\nu})} \frac{\gamma^*_{n(\boldsymbol{\nu})}}{|\boldsymbol{\nu}|} &\le \text{const.}\, sC_0 \sum_{n=0}^{\infty} 2^{n(r-1)}\gamma^*_n \log\log 1/\gamma^*_n \\ &\le \text{const.}\, sC_0 \sum_{n=0}^{\infty} n2^{n(r-1)}\gamma^*_n,\end{aligned} \tag{3.32}$$

which is bounded proportionally to C_0 by Lemma 2.

Analogously one discusses the other conditions in (3.8). Simply one defines $F(t) = \alpha(t)|\boldsymbol{\nu}| \pm \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,(\alpha(t),\boldsymbol{\beta}))} \pm \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,(\alpha(t),\boldsymbol{\beta}))} = tC_0\gamma^*_{n(\boldsymbol{\nu})}$, so that again one has $\mathrm{d}F/\mathrm{d}t = |\boldsymbol{\nu}|(1+O(\sqrt{\varepsilon}))\mathrm{d}\alpha/\mathrm{d}t = C_0\gamma^*_{n(\boldsymbol{\nu})}$, and one can proceed as before. ∎

To complete the proof of Theorem 1 we have to prove the last assertion about maximal tori, that is that for $r = d$ most of phase space is filled by invariant tori.

We summarise what we have found so far. The invariant tori are determined by the corresponding rotation vectors $\boldsymbol{\omega}$. For $\boldsymbol{\omega}\in\Omega_*(C_0)$ we can parameterise the invariant torus with rotation vector $\boldsymbol{\omega}$ as

$$\begin{cases} \boldsymbol{\alpha} = \boldsymbol{\psi} + \mathbf{a}(\boldsymbol{\psi},\boldsymbol{\omega},\varepsilon), \\ \mathbf{A} = \boldsymbol{\omega} + (\boldsymbol{\omega}\cdot\partial_{\boldsymbol{\psi}})\,\mathbf{a}(\boldsymbol{\psi},\boldsymbol{\omega},\varepsilon), \end{cases} \tag{3.33}$$

with $\boldsymbol{\psi}\in\mathbb{T}^r$. Moreover the function $\mathbf{a}$ is analytic in $\boldsymbol{\psi}$ (as the Fourier coefficients decay exponentially), while it is defined only on a Cantorian set of values $\boldsymbol{\omega}$.

For each value of $\boldsymbol{\psi}$ we can consider the map $\boldsymbol{\omega}\mapsto\mathbf{A}(\boldsymbol{\omega})$, defined in (3.33). For ε small enough in (3.33) one has $|\mathbf{A}-\boldsymbol{\omega}| = |(\boldsymbol{\omega}\cdot\partial_{\boldsymbol{\psi}})\,\mathbf{a}(\boldsymbol{\psi},\boldsymbol{\omega},\varepsilon)| \le R|\varepsilon|$ for some constant R. Call $\Omega_*(C_0,d)$ the open set obtained from $\Omega_*(C_0)$ by excluding all vectors within a distance d from the boundary of Ω, i.e. $\Omega_*(C_0,d) = \{\boldsymbol{\omega}\in\Omega(C_0) : d(\boldsymbol{\omega},\partial\Omega) \ge d\}$, with $d(\boldsymbol{\omega},\partial\Omega) = \min_{\boldsymbol{\omega}'\in\partial\Omega}|\boldsymbol{\omega}-\boldsymbol{\omega}'|$, and define $\mathcal{A}_*(C_0)$ as the image of $\Omega_*(C_0,R|\varepsilon|)$ of such a map (note that the latter is not just the inverse of the frequency map, rather it is a perturbation of it). Then the measure of the complement of the action variable space filled by invariant tori is given by

$$\mathrm{meas}(\mathcal{A}^c_*(C_0)) = \int_{\mathcal{A}^c_*(C_0)} \mathrm{d}\mathbf{A} = \int_{\Omega^c_*(C_0,R|\varepsilon|)} \mathrm{d}\boldsymbol{\omega}\,|\det\partial_{\boldsymbol{\omega}}\mathbf{A}|\,, \tag{3.34}$$

provided that the Jacobian in the last integral is well defined (that is the map $\boldsymbol{\omega}\to\mathbf{A}(\boldsymbol{\omega})$ is smooth enough, at least in the sense of Whitney) and is uniformly bounded. This turns out to be the case, as the following result shows.

Lemma 12. *The solutions of the equations of motion* $\mathbf{h}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)$ *are differentiable in* $\boldsymbol{\omega}$ *in the sense of Whitney for* $\boldsymbol{\omega}\in\Omega_*(C_0,R|\varepsilon|)$.

Proof. The proof (for any value of $r \le d$, not necessarily $r = d$) can be performed as for Lemma 10, with the only difference that now we have to deal with the renormalised expansion for $h_{\boldsymbol{\nu},\gamma}$ instead of the matrices $\mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega})$, hence with trees instead of self-energy clusters. The condition $d(\boldsymbol{\omega},\partial\Omega) \geq R|\varepsilon|$ yields that the actions variables $\mathbf{A}$ remain in Ω for all values of $\boldsymbol{\psi}$. ∎

As a consequence we can bound the Jacobian in (3.34) by using (3.33), which gives $\partial_{\boldsymbol{\omega}}\mathbf{A} = \mathbb{1} + \partial_{\boldsymbol{\omega}}(\boldsymbol{\omega}\cdot\partial_{\boldsymbol{\psi}})\,\mathbf{a}(\boldsymbol{\psi},\boldsymbol{\omega},\varepsilon)$, and Lemma 12, which assures that the last derivative (in the sense of Whitney) is bounded proportionally to ε. Therefore we can bound $\mathrm{meas}(\mathcal{A}_*^{c}(C_0))$ proportionally to C_0 by Lemma 11, and by taking $C_0 = O(\sqrt{|\varepsilon|})$ (which is allowed by Lemma 9), we obtain the last assertion of Theorem 1. Cf. also Refs. [13] and [45], where the usual Diophantine conditions were considered in the analytic and differentiable case, respectively.

4. Fixing the rotation vector: proof of Theorem 2

In the following we assume that $\boldsymbol{\omega}$ is fixed, and that it satisfies the Bryuno Diophantine condition $B(\boldsymbol{\omega}) < \infty$, with $B(\boldsymbol{\omega})$ defined in (2.4). Set $\gamma_n = C_0^{-1}\alpha_n(\boldsymbol{\omega})$.

Let $\varepsilon \in (\varepsilon_0/4, \varepsilon_0]$ and set $\lambda_d^{[0]} = \varepsilon a_s$ and $\varepsilon_0 a_s = \Lambda_0$. Define $n_0 \in \mathbb{N}$ such that $C_0\gamma_{n_0+1} < 2\sqrt{\Lambda_0} \leq C_0\gamma_{n_0}$. We set

$$\gamma_n^* = \begin{cases} \gamma_n, & n < n_0, \\ \gamma_n 2^{-n(r+1)}, & n \geq n_0, \end{cases} \tag{4.1}$$

and, by using the sequence $\{\gamma_n^*\}_{n=0}^{\infty}$, we proceed as in Section 3, for constructing the multi-scale decomposition of the propagators. Though, we define $\Delta^{[n]}(x;\varepsilon,\boldsymbol{\omega}) = \Delta^{[0]}(x;\varepsilon,\boldsymbol{\omega})$ for $n \leq n_0$.

The main difference is that we shall need the following Diophantine conditions:

$$\begin{aligned} &\left|\boldsymbol{\omega}\cdot\boldsymbol{\nu} \pm \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})}\right| \geq C_0\gamma_{n(\boldsymbol{\nu})}^*, \\ &\left|\boldsymbol{\omega}\cdot\boldsymbol{\nu} \pm \sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})} \pm \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})}\right| \geq C_0\gamma_{n(\boldsymbol{\nu})}^* \end{aligned} \tag{4.2}$$

for all $i,j = 1,\ldots,d$, for all $\boldsymbol{\nu} \in \mathbb{Z}_*^r$ such that $n(\boldsymbol{\nu}) \geq n_0$ and for all $n \geq n_0$. We do not impose any conditions like (4.2) for $n \leq n_0$, because for such scales one has $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}| > 2\sqrt{\Lambda_0}$, so that we can bound $|(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 - \underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})|$ with $(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2/2$ for all $i = 1,\ldots,d$. In the same way we have excluded in (4.2) the values of $\boldsymbol{\nu} \in \mathbb{Z}_*^r$ such that $n(\boldsymbol{\nu}) \leq n_0$. Hence, at the price of adding a factor 2^2 in the bound of each propagator, we can confine ourselves to impose (4.3) only for $\boldsymbol{\nu}$ such that $n(\boldsymbol{\nu}) \geq n_0$ and for $n \geq n_0$.

Then we can prove the following result.

Lemma 13. *Call $N_n(\theta)$ the set of lines in $\Lambda(\theta)$ which are on scale $[n]$. One has*

$$N_n(\theta) \leq K\,2^{-n}M(\theta), \qquad M(\theta) = \sum_{\mathfrak{v}\in V(\theta)} |\boldsymbol{\nu}_{\mathfrak{v}}|, \tag{4.3}$$

for a suitable constant K. One can take $K = 2$.

Proof. The proof proceeds exactly as for Lemma 3, with the only difference that we have to deal in a different way with the lines on scales $n < n_0$ and those with scales $n \geq n_0$. The same was done in Ref. [28]. ■

In the same way the following result is proved.

Lemma 14. *Call $N_n(T)$ the set of lines in $\Lambda(T)$ which are on scale $[n]$. One has*

$$M(T) = \sum_{\mathfrak{v}\in V(T)} |\boldsymbol{\nu}_{\mathfrak{v}}| > 2^{n_T-1}, \qquad N_n(T) \le K\, 2^{-n/2} M(T), \tag{4.4}$$

for a suitable constant K.

Then Lemma 5 is replaced with the following one.

Lemma 15. *Assume that the propagators $g^{[p]}(x;\varepsilon,\boldsymbol{\omega})$ can be uniformly bounded for all $0 \le p \le n-1$ as (3.14), for some p-independent constant K. Assume also that ε_0 is small enough. With the notations (3.3) one has*

$$\begin{aligned} \left\| \mathcal{M}^{[n]}_{\alpha\alpha}(x;\varepsilon,\boldsymbol{\omega}) \right\| &\le B\mathrm{e}^{-\kappa_1 2^{n/2}} \min\{\varepsilon^2, \varepsilon x^2\}, \\ \left\| \mathcal{M}^{[n]}_{\alpha\beta}(x;\varepsilon,\boldsymbol{\omega}) \right\| &\le B\mathrm{e}^{-\kappa_1 2^{n/2}} \min\{\varepsilon^2, \varepsilon^{3/2}|x|\}, \\ \left\| \mathcal{M}^{[n]}_{\beta\beta}(x;\varepsilon,\boldsymbol{\omega}) \right\| &\le B\mathrm{e}^{-\kappa_1 2^{n/2}} \varepsilon^2, \end{aligned} \tag{4.5}$$

for all $n \in \mathbb{N}$ and for suitable n-independent constants B and κ_1.

Therefore we can prove the following estimates. The proof is the same as for Lemma 6, as one easily realizes that it works for fixed values of ε and $\boldsymbol{\omega}$.

Lemma 16. *Let $\boldsymbol{\omega}$ satisfy the Diophantine condition (1.4) and assume that ε is small enough, say $|\varepsilon| < \varepsilon_0$. Then the series (3.7) admits the bound $|h_{\boldsymbol{\nu},\gamma}| < A\,\mathrm{e}^{-\kappa|\boldsymbol{\nu}|}|\varepsilon|$ for suitable positive constants κ and A. One has $\varepsilon_0 = O(C_0^2(\gamma^*_{m_0})^2)$ with m_0 depending on $\boldsymbol{\omega}$ and κ_0.*

With respect to Section 3 the first differences appear when dealing with Whitney extensions of the matrices $\mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega})$: indeed now $\boldsymbol{\omega}$ is assumed to be fixed, while ε is the free parameter. We define $\mathcal{E}_{n_0} \equiv (\varepsilon_0/4, \varepsilon_0]$ and for $n > n_0$, recursively, $\mathcal{E}_n = \mathcal{E}_{n-1} \setminus \mathcal{E}^o_n$, where $\mathcal{E}^o_n$ is the set of values of $\varepsilon \in \mathcal{E}_n$ such that the conditions (4.2) are violated. We define also $\mathcal{E}_* = \cap_{n=n_0}^{\infty} \mathcal{E}_n$.

Lemma 17. *For all $n \ge 0$ and all $\varepsilon, \varepsilon' \in \mathcal{E}_n$ one has*

$$\begin{aligned} &\left\| \mathcal{M}^{[\le n]}(x;\varepsilon',\boldsymbol{\omega}) - \mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega}) - \partial_\varepsilon \mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega})\,(\varepsilon'-\varepsilon) \right\| = \varepsilon\, o(|\varepsilon'-\varepsilon|), \\ &\left\| \partial_\varepsilon \mathcal{M}^{[\le n]}(x;\varepsilon,\boldsymbol{\omega}) \right\| \le B, \end{aligned} \tag{4.6}$$

and, as a consequence,

$$\begin{aligned} B' &\le \left| \partial_\varepsilon \underline{\lambda}^{[n]}_j(\varepsilon,\boldsymbol{\omega}) \right| \le B, \qquad r+1 \le j \le d, \\ B' &\le \left| \partial_\varepsilon \left(\underline{\lambda}^{[n]}_i(\varepsilon,\boldsymbol{\omega}) \pm \underline{\lambda}^{[n]}_j(\varepsilon,\boldsymbol{\omega}) \right) \right| \le B, \qquad r+1 \le j < i \le d, \end{aligned} \tag{4.7}$$

for suitable positive constants B and B'.

Proof. The proof can be performed as for Lemma 10, with the parameter ε now playing the role of the parameters $\boldsymbol{\omega}$. We do not give the details, which, however, have been worked out in Ref. [28]. Again the upper bound (4.7) follows from (4.6); cf. analogous comments in the proof of Lemma 6. To obtain the lower bound one has to use also that $\lambda_j(x;\varepsilon,\boldsymbol{\omega}) = a_j\varepsilon + O(\varepsilon^2)$, with $a_j \neq 0$ and $a_i \neq a_j$ for $i,j = r+1,\ldots,d$; again cf. Ref. [28] for details. ■

Lemma 17 implies that the matrices $\mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega})$ can be extended in $(0,\varepsilon_0)$ to smooth C^1 functions (Whitney extensions). Again a closeness property of the self-energies, which reads

$$\left|\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega}) - \underline{\lambda}_j^{[n-1]}(\varepsilon,\boldsymbol{\omega})\right| \leq B\mathrm{e}^{-\kappa_1 2^{n/2}}\varepsilon^2, \tag{4.8}$$

follows from Lemma 15. As before the bounds (4.8) can be improved for the first r self-energies, and give $|\lambda_j^{[n]}(x;\varepsilon,\boldsymbol{\omega})| \leq A\min\{\varepsilon^2,\varepsilon x^2\}$, $j = 1,\ldots r$. What really changes with respect to the previous case is the estimate of the set of allowed values of ε, which explains why we have required the stronger condition on $\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)$ that its eigenvalues are non-degenerate. The following result holds.

Lemma 18. *The Lebesgue measure of the set $(0,\varepsilon_0)\setminus\mathcal{E}^*$ is bounded proportionally to some value $G(\varepsilon_0)$, with $G(\varepsilon) = o(\varepsilon)$.*

Proof. As in the proof of Lemma 12 we start with the first conditions in (4.2). By setting $\varepsilon = \varepsilon(t)$, with $t \in [-1,1]$, and defining $F(t) = \boldsymbol{\omega}\cdot\boldsymbol{\nu} \pm \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon(t),\boldsymbol{\omega})} = tC_0\gamma^*_{n(\boldsymbol{\nu})}$, one has $|\mathrm{d}F/\mathrm{d}t| = |\partial_\varepsilon\sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})}|\,|\mathrm{d}\varepsilon/\mathrm{d}t| = C_0\gamma^*_{n(\boldsymbol{\nu})}$, so that, by using that $\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega}) = a_j\varepsilon + O(\varepsilon^2)$, one finds $|\mathrm{d}\varepsilon/\mathrm{d}t| \leq BC_0\sqrt{\varepsilon}\gamma^*_{n(\boldsymbol{\nu})}$ for some constant B. Again, for fixed $\boldsymbol{\nu}$ and i, by the closeness property of the self-energies, we can impose only the conditions corresponding to the scales up to $n_1(\boldsymbol{\nu}) = O(\log\log 1/\gamma^*_{n(\boldsymbol{\nu})})$, at the price of enlarging the sets of excluded values (by a factor 2). Hence the measure of the set of excluded values $\varepsilon \in \mathcal{E}^{[n-1]}$, $n \geq n_0$, found by imposing the first conditions (4.2) can be bounded by

$$\begin{aligned} &\text{const.} \sum_{i=r+1}^{d} C_0\sqrt{\varepsilon_0} \sum_{n=n_0}^{\infty} \sum_{n'=n_0}^{n_1(\boldsymbol{\nu})} \sum_{2^{n-1}<|\boldsymbol{\nu}|\leq 2^n} \gamma^*_n \\ &\qquad \leq \text{const.}\, s\, C_0\sqrt{\varepsilon_0} \sum_{n=n_0}^{\infty} \gamma^*_n\, 2^{nr} \log\log\frac{1}{\gamma^*_n} \equiv G(\varepsilon_0), \end{aligned} \tag{4.9}$$

where we have used that n_0 is uniquely determined by ε_0. Therefore we have

$$\begin{aligned} \frac{G(\varepsilon_0)}{\varepsilon_0} &\leq \text{const.}\, \frac{1}{\gamma_{n_0}} \sum_{n=n_0}^{\infty} \gamma^*_n(\boldsymbol{\omega})\, 2^{nr} \log\log\frac{1}{\gamma^*_n} \\ &\leq \text{const.}\, \frac{1}{\gamma_{n_0}(\boldsymbol{\omega})} \sum_{n=n_0}^{\infty} n\,\gamma^*_{n_0} 2^{-n(r+1)} 2^{nr}, \end{aligned} \tag{4.10}$$

which tends to 0 as $n_0 \to \infty$ (that is as $\varepsilon_0 \to 0$).

The estimate of the measure of the set of excluded values $\varepsilon \in \mathcal{E}^{[n-1]}$, $n \geq n_0$, found by imposing the second conditions (4.2) can be obtained by reasoning in the same way. In such a case we need a lower bound on $\partial_\varepsilon(\sqrt{\underline{\lambda}_i^{[n]}(\varepsilon,\boldsymbol{\omega})} \pm \sqrt{\underline{\lambda}_j^{[n]}(\varepsilon,\boldsymbol{\omega})})$, which requires $|a_i - a_j| > a$ for all $i \neq j$ and for some constant a; cf. Ref. [28], Appendix A2, for a similar discussion. ■

Appendix A1. Trees and formal expansion

In this appendix we briefly recall the diagrammatic formalism used to express the function $\mathbf{h} = \mathbf{h}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon)$ in terms of trees. Let us denote with h_γ, $\gamma = 1,\ldots,d$, the components of $\mathbf{h}$, and with $h_{\boldsymbol{\nu},\gamma}$, $\gamma = 1,\ldots,d$, the components of the Fourier coefficients $\mathbf{h}_{\boldsymbol{\nu}}$ of the Fourier expansion in $\boldsymbol{\psi}$ for $\mathbf{h}$. Hence the first r components of $\mathbf{h}$ define the function $\mathbf{a} = \mathbf{a}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon)$, while the last s components define the function $\mathbf{b} = \mathbf{b}(\boldsymbol{\psi}, \boldsymbol{\beta}_0, \boldsymbol{\omega}, \varepsilon)$.

A tree θ is defined as a partially ordered set of points and *lines* connecting the points. The lines are consistently oriented toward the *root*, which is the leftmost point $\mathfrak{r}$. If a line ℓ connects two points $\mathfrak{v}_1, \mathfrak{v}_2$ and is oriented from $\mathfrak{v}_2$ to $\mathfrak{v}_1$, we write $\mathfrak{v}_2 \prec \mathfrak{v}_1$, and we shall set $\mathfrak{v}_2' = \mathfrak{v}_1$ and $\ell_{\mathfrak{v}_2} = \ell$; we shall say also that ℓ exits $\mathfrak{v}_2$ and enters $\mathfrak{v}_1$. More generally we write $\mathfrak{v}_2 \prec \mathfrak{v}_1$ if $\mathfrak{v}_1$ is on the path of lines connecting $\mathfrak{v}_2$ to the root. The root admits only one entering line, which is called the *root line*. The points different from the root will be called the *nodes* of the tree: by construction the number of nodes equals the number of lines. For an example of tree cf. Ref. [28], Section 2: note, however, that drawing pictures for trees instead of simply relying on the abstract definition is by no means mandatory, even if it helps following the forthcoming construction of the function $\mathbf{h}$.

Each line from $\mathfrak{v}$ to $\mathfrak{v}'$ carries a pair η of *component* labels $\eta = (\gamma, \gamma')$ ranging in $\{1,\ldots,d\}$. The labels γ and γ' should be regarded as associated with $\mathfrak{v}$ and $\mathfrak{v}'$, respectively; hence with each node $\mathfrak{v}$ with $p_{\mathfrak{v}}$ entering lines $\ell_1,\ldots,\ell_{p_{\mathfrak{v}}}$ one can associate $p_{\mathfrak{v}}+1$ labels $\gamma_0,\gamma_1,\ldots,\gamma_{p_{\mathfrak{v}}}$, with $\gamma_0 = \gamma_{\ell_{\mathfrak{v}}}$ and $\gamma_j = \gamma'_{\ell_j}$. Also the root line carries two such labels and the one associated with the left extreme of the root line will be called the *root label.*

Fixed any line $\ell_{\mathfrak{v}}$ in θ, we shall say that the subset of θ containing $\ell_{\mathfrak{v}}$ as well as all nodes $\mathfrak{w} \preceq \mathfrak{v}$ and all lines connecting them is a *subtree* of θ with root $\mathfrak{v}'$; of course a subtree is a tree (by construction).

Given a tree, we shall denote by $V(\theta)$ the set of nodes and by $\Lambda(\theta)$ the set of lines. The number $k = |V(\theta)|$ of nodes in the tree θ, equal to the number $|\Lambda(\theta)|$ of lines, will be called the *order* of θ. We call a node with one entering line and $\mathbf{0}$ mode label a *trivial node.*

With each node $\mathfrak{v} \in V(\theta)$ we associate a *harmonic* or *mode* label $\boldsymbol{\nu}_{\mathfrak{v}} \in \mathbb{Z}^r$. With any line $\ell = \ell_{\mathfrak{v}}$ we associate, besides the above mentioned pair $\eta_\ell = (\gamma_\ell, \gamma'_\ell)$ of labels assuming values in $\{1,\ldots,d\}$, a *momentum* $\boldsymbol{\nu}_\ell \in \mathbb{Z}^r$ defined as

$$\boldsymbol{\nu}_\ell \equiv \boldsymbol{\nu}_{\ell_{\mathfrak{v}}} = \sum_{\substack{\mathfrak{w} \in V(\theta) \\ \mathfrak{w} \preceq \mathfrak{v}}} \boldsymbol{\nu}_{\mathfrak{w}}. \tag{A1.1}$$

We shall consider only trees not containing trivial nodes with the entering line with $\mathbf{0}$ momentum. This is a consequence of the derivation of the Lindstedt series – see Ref. [19] and [28] for further details, – and it simply means that eventually the function $\mathbf{h}$ will turn out to be expressed in terms of trees where no such nodes appear.

We call *degree* $P(\theta)$ of a tree the order of the tree minus the number of $\mathbf{0}$ momentum lines, so that $|V(\theta)| - P(\theta)$ is their number.

We call $\Theta_{\boldsymbol{\nu},k,\gamma}$ the set of trees θ whose root line $\ell_{\mathfrak{v}_0}$ has momentum $\boldsymbol{\nu}$, root label γ and have order k, i.e. with $|V(\theta)| = k$ nodes, while we call $\Theta^o_{\boldsymbol{\nu},k,\gamma}$ the set of trees of degree k, i.e. with $P(\theta) = k$. Note that $\Theta_{\boldsymbol{\nu},k,\gamma} \neq \Theta^o_{\boldsymbol{\nu},k,\gamma}$.

Each tree θ "decorated" by labels in the described way will have a *value* which is defined in terms of a product of several factors. With each node $\mathfrak{v}$ we associate a *node factor*

$$F_{\mathfrak{v}} = \Big(\prod_j \partial_{\gamma_j} \Big) f_{\boldsymbol{\nu}_{\mathfrak{v}}}(\boldsymbol{\beta}_0), \tag{A1.2}$$

where the following conventions have been used: $f_{\boldsymbol{\nu}}(\boldsymbol{\beta})$ is the Fourier coefficient with label $\boldsymbol{\nu}$ in the Fourier expansion in the angle $\boldsymbol{\alpha}$ for the perturbation $f(\boldsymbol{\alpha},\boldsymbol{\beta})$ (see (1.2) for notations); the labels γ_j are the $p_{\mathfrak{v}}+1$ labels associated with the extreme $\mathfrak{v}$ of the $p_{\mathfrak{v}}$ lines entering the node $\mathfrak{v}$ and of the line exiting it; the derivatives ∂_γ, with $\gamma = 1,\ldots,r$, have to be interpreted as factors $(i\boldsymbol{\nu}_{\mathfrak{v}})_\gamma$, while $\partial_\gamma = \partial/\partial\beta_{\gamma-r}$ for $\gamma = r+1,\ldots,d$ (recall that $d = r+s$). Hence $F_{\mathfrak{v}}$ is a tensor of rank $p_{\mathfrak{v}}+1$. With each line ℓ carrying labels $\eta_\ell = (\gamma_\ell, \gamma_{\ell'})$ and momentum $\boldsymbol{\nu}_\ell$ we associate a matrix, called *propagator*,

$$\begin{aligned} G_\ell &\equiv \delta_{\gamma_\ell,\gamma'_\ell} \frac{1}{(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell)^2}, && \text{if } \boldsymbol{\nu}_\ell \neq \mathbf{0}, \\ G_\ell &\equiv -\varepsilon^{-1} \left(\partial^2_{\boldsymbol{\beta}} f_{\mathbf{0}}(\boldsymbol{\beta}_0)\right)^{-1}_{\gamma_\ell,\gamma'_\ell} \chi(\gamma_\ell,\gamma'_\ell > r), && \text{if } \boldsymbol{\nu}_\ell = \mathbf{0}, \end{aligned} \tag{A1.3}$$

where $\chi(\gamma_\ell,\gamma'_\ell > r)$ is 1 if both γ_ℓ and γ'_ℓ are strictly greater than r, and 0 otherwise.

The propagators (A1.3) are matrices which can be written as

$$G_\ell = \begin{pmatrix} G_{\ell,\alpha\alpha} & G_{\ell,\alpha\beta} \\ G_{\ell,\beta\alpha} & G_{\ell,\beta\beta} \end{pmatrix}, \tag{A1.4}$$

where the labels α and β run over $\{1,\ldots,r\}$ and $\{r+1,\ldots,d\}$, respectively. This motivates the writing (3.3).

Given the definitions (A1.2) and (A1.3), we define a *value function* Val, which with each tree θ of order k associates a *tree value*

$$\mathrm{Val}(\theta) = \frac{\varepsilon^k}{k!} \Big(\prod_{\mathfrak{v}\in V(\theta)} F_{\mathfrak{v}} \Big) \Big(\prod_{\ell\in\Lambda(\theta)} G_\ell \Big), \tag{A1.5}$$

where, by the definitions, all labels γ_i associated with the nodes appear twice because they appear also in the propagators: we make in (A1.5) the *summation convention* that repeated γ labels associated with nodes and lines are summed over, with the exception of the label γ associated with the root. Therefore (A1.5) is a number labeled by $\gamma = 1,\ldots,d$, i.e. $\mathrm{Val}(\theta)$ is a vector.

The trees can be drawn in various ways: we can limit the arbitrariness by demanding that the length of the segments representing the lines is 1 and that the angles between the lines are irrelevant. This arbitrariness reflects the fact that trees are uniquely identified by the abstract definition given here, and drawing trees is just a way to visualise the construction. We adopt the convention that trees are drawn on a plane and carry an identifier label, that we call *number label* which distinguishes the lines from each other even if we ignore the other labels attached to them. Furthermore two trees that can be transformed into each other by permuting the lines entering the same node $\mathfrak{v}$ are considered identical.

Then the formal expansion for $\mathbf{h}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)=(\boldsymbol{\alpha}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon),\boldsymbol{\beta}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon))$ is defined as

$$\begin{aligned}&\mathbf{h}(\boldsymbol{\psi},\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)=\sum_{\boldsymbol{\nu}\in\mathbb{Z}^r}\mathrm{e}^{i\boldsymbol{\nu}\cdot\boldsymbol{\psi}}\,\mathbf{h}_{\boldsymbol{\nu}}(\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon),\qquad \mathbf{h}_{\boldsymbol{\nu}}(\boldsymbol{\beta}_0,\boldsymbol{\omega},\varepsilon)\equiv\mathbf{h}_{\boldsymbol{\nu}},\\&\mathbf{h}_{\boldsymbol{\nu}}=(h_{\boldsymbol{\nu},1},\ldots,h_{\boldsymbol{\nu},d}),\qquad h_{\boldsymbol{\nu},\gamma}=\sum_{k=1}^{\infty}\sum_{\theta\in\Theta^o_{k,\nu,\gamma}}\mathrm{Val}(\theta),\end{aligned}\tag{A1.6}$$

where $\Theta^o_{k,\boldsymbol{\nu},\gamma}$ is defined after (A1.1). Note that $h_{\boldsymbol{\nu},\gamma}$ is a (formal) power series in ε: this follows from (A1.3) and (A1.5), and by the observation that, because of the absence of trivial nodes whose entering line has zero momentum, the number of lines with zero momentum, hence with propagators proportional to $1/\varepsilon$, is strictly less than the order k (so that the degree is always positive).

A different expansion will be introduced in Appendix A2. This is due to the fact that the series (A1.6) is likely to be divergent – at least we cannot prove that it converges; see related remarks in Refs. [19] and [28]. Hence we shall be forced to give a summation criterion of the formal series, which leads to the renormalised expansion (3.7). The backbone of the renormalised expansion in terms of trees is the same as that given here for the formal expansion: essentially the main difference is that – as shown in Section 3 – the propagators are changed, and some further constraints have to imposed on the trees.

Appendix A2. Renormalised trees and renormalised expansion

The renormalised trees are defined as the trees in Appendix A1, with the following differences. Besides the momentum $\boldsymbol{\nu}_\ell$, each line ℓ carries also a scale label $[n_\ell]$, where $n_\ell=-1$ if $\boldsymbol{\nu}_\ell=\mathbf{0}$ and $n_\ell\in\mathbb{Z}_+=\{\mathbf{0}\}\cup\mathbb{N}$ otherwise.

A cluster on scale $[n]$ is a maximal set of nodes and lines connecting them with propagators on scale $[p]$, $p\le n$, and such that at least one line is on scale exactly $[n]$. Each cluster has at most one exiting line, while it has $m_T\ge 0$ entering lines (the notions of entering and exiting lines are well defined, as the trees are oriented; recall that each line carries an arrow pointing toward the root of the tree): such lines are called the external lines of the cluster.

A self-energy scale is a cluster T such that the following properties hold: (i) it has one exiting line and only one entering lines, (ii) both the external lines carry the same momentum, so that

$$\sum_{\mathfrak{v}\in V(T)}\boldsymbol{\nu}_{\mathfrak{v}}=\mathbf{0},\tag{A2.1}$$

and (iii) no line along the path of lines connecting the external lines is on scale $[-1]$. Note that the last property, even if not explicitly mentioned in the Definition 4 of Ref. [28], was nonetheless used along the proofs of the aforementioned paper.

Then a *renormalised tree* is a tree in which no self-energy clusters (as well as no trivial nodes) appear. The node factors are still defined by (A1.2), whereas the (renormalised) propagators g_ℓ are defined recursively by (3.5). Then we can define the value of a renormalised tree θ as

$$\mathrm{Val}(\theta) = \frac{\varepsilon^k}{k!}\Big(\prod_{\mathfrak{v}\in V(\theta)} F_v\Big)\Big(\prod_{\ell\in\Lambda(\theta)} g_\ell\Big), \tag{A2.2}$$

and the renormalised expansion for $\mathbf{h}$ is given by (3.7).

As already remarked in Section 3, here we are not distinguishing between ultraviolet and infrared scales, and we have defined the propagator divisors according to (3.4) for all $n \ge 0$. In fact, the self-energies can all be bounded as x^2 as far as the scale is less than n_0, if $\varepsilon a_s \approx C_0^2 2^{-2n_0}$ (ultraviolet region). The two approaches are equivalent. In Section 3 we prefer to follow the first one, as it allows do treat all scales in the same way, so leading to more uniform, hence simpler, notations. In Section 4 we adopt the second approach, simply to make the discussion as similar as possible to that of Ref. [28]. In such a way, a reader acquainted with that paper, can follow easily the changes required in order to extend the results to Bryuno rotation vectors.

Appendix A3. Proof of Lemma 10

If we consider two rotation vectors $\boldsymbol{\omega},\boldsymbol{\omega}' \in \Omega(C_0)$, they are characterised by the respective sequences $\{\gamma_n(\boldsymbol{\omega})\}_{n=0}^\infty$ and $\{\gamma_n(\boldsymbol{\omega}')\}_{n=0}^\infty$. For all $n \ge 0$ one has $\gamma_n(\boldsymbol{\omega}) \ge \gamma_n^*$ and $\gamma_n(\boldsymbol{\omega}') \ge \gamma_n^*$.

We introduce some shortened notations, by setting

$$\begin{aligned}
\Xi_n(\boldsymbol{\nu},\boldsymbol{\omega}) &= \prod_{p=0}^{n} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})),\\
\Psi_n(\boldsymbol{\nu},\boldsymbol{\omega}) &= \left(\prod_{p=0}^{n-1} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}))\right)\psi_n(\Delta^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})),
\end{aligned} \tag{A3.1}$$

and

$$\begin{aligned}
\Xi_{n,s}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega}) &= \Big(\prod_{p=0}^{s-1} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}'))\Big)\Big(\chi_s(\Delta^{[s]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')) - \chi_s(\Delta^{[s]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}))\Big)\\
&\quad\Big(\prod_{p=s+1}^{n} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}))\Big), \qquad 0 \le s \le n, \qquad\qquad (A3.2)\\
\Psi_{n,s}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega}) &= \Big(\prod_{p=0}^{s-1} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}'))\Big)\Big(\chi_s(\Delta^{[s]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')) - \chi_s(\Delta^{[s]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}))\Big)\\
&\quad\Big(\prod_{p=s+1}^{n-1} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}))\Big)\psi_n(\Delta^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})), \qquad 0 \le s \le n-1,\\
\Psi_{n,n}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega}) &= \Big(\prod_{p=0}^{n-1} \chi_p(\Delta^{[p]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}'))\Big)\Big(\psi_n(\Delta^{[n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')) - \psi_n(\Delta^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}))\Big),
\end{aligned}$$

where all products have to be meant as 1 when containing no factor.

Finally we define the Hermitian matrices

$$D_n(\boldsymbol{\nu},\boldsymbol{\omega}) = (\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 - \mathcal{M}^{[\le n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}), \tag{A3.3}$$

with $\mathcal{M}^{[\leq n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})$ given by (3.3). In the obvious way one defines also $D_n(\boldsymbol{\nu},\boldsymbol{\omega}')$.

Note that one has

$$\sum_{p=0}^{n-1}(\gamma_p^*)^{-m} \leq n\,(\gamma_{n-1}^*)^{-m} < (\gamma_n^*)^{-(m+1)}, \tag{A3.4}$$

for all $m\in\mathbb{N}$.

Lemma A1. *One has*

$$\chi_n(x') - \chi_n(x) = b(x)\,(x'-x) + o(|x'-x|), \qquad |b(x)| \leq \Phi\,(\gamma_n^*)^{-2}|x'-x|, \tag{A3.5}$$

for a suitable positive constant Φ*.*

Proof. We can write

$$|\chi_n(x')-\chi_n(x)| \leq \beta^{-2}(\gamma_n^*)^{-2}\,|x'-x|\int_0^1 \mathrm{d}t\,\partial_x\chi(x(t)), \tag{A3.6}$$

where $x(t)=\beta^{-2}(\gamma_n^*)^{-2}(x+t(x'-x))$ and $|\partial_x\chi(x(t))|\leq\text{const}$. ■

By noting that $\psi_n = 1-\chi_n$, we see that Lemma A1 yields the same bounds as (A3.6) also if we replace χ_n with ψ_n.

Lemma A2. *For* $\boldsymbol{\omega},\boldsymbol{\omega}'\in\Omega_*(C_0)$ *in the same direction* $\mathbf{v}$ *assume that the bounds (3.28) hold for all* $n'\leq n$*. Then one has*

$$\begin{aligned}&\|g^{[n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}') - g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}) - \partial_{\mathbf{v}}g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})\,(\boldsymbol{\omega}'-\boldsymbol{\omega})\| = (\gamma_n^*)^{-\delta}|\boldsymbol{\nu}|\,o(|\boldsymbol{\omega}'-\boldsymbol{\omega}|),\\ &\|\partial_{\mathbf{v}}g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})\| \leq G\,(\gamma_n^*)^{-\delta}|\boldsymbol{\nu}|,\end{aligned} \tag{A3.7}$$

for suitable positive constants G *and* δ*.*

Proof. By using the definition (3.5) we have

$$\begin{aligned}g^{[n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}') - g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}) &= \Psi_n(\boldsymbol{\nu},\boldsymbol{\omega}')D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega}') - \Psi_n(\boldsymbol{\nu},\boldsymbol{\omega})D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega})\\ &= -\Psi_n(\boldsymbol{\nu},\boldsymbol{\omega})D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega}')\left(D_n(\boldsymbol{\nu},\boldsymbol{\omega}')-D_n(\boldsymbol{\nu},\boldsymbol{\omega})\right)D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega})\\ &\quad+\sum_{p=0}^{n}\Psi_{n,p}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega})D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega}'),\end{aligned} \tag{A3.8}$$

where we can write

$$\begin{aligned}&D_n(\boldsymbol{\nu},\boldsymbol{\omega}')-D_n(\boldsymbol{\nu},\boldsymbol{\omega})\\ &\quad= (\boldsymbol{\omega}'\cdot\boldsymbol{\nu}+\boldsymbol{\omega}\cdot\boldsymbol{\nu})((\boldsymbol{\omega}'-\boldsymbol{\omega})\cdot\boldsymbol{\nu}) - \mathcal{M}^{[\leq n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}') + \mathcal{M}^{[\leq n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})\end{aligned} \tag{A3.9}$$

so that we obtain

$$\begin{aligned}&D_n(\boldsymbol{\nu},\boldsymbol{\omega}')-D_n(\boldsymbol{\nu},\boldsymbol{\omega})\\ &\quad= \Big(2\boldsymbol{\nu}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) - \partial_{\boldsymbol{\omega}}\mathcal{M}^{[\leq n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu},\varepsilon,\boldsymbol{\omega})\Big)\cdot(\boldsymbol{\omega}'-\boldsymbol{\omega}) + ((\boldsymbol{\omega}'-\boldsymbol{\omega})\cdot\boldsymbol{\nu})^2 + \varepsilon^2|\boldsymbol{\nu}|\,o(|\boldsymbol{\omega}'-\boldsymbol{\omega}|),\end{aligned} \tag{A3.10}$$

by the assumed estimate (3.28).

If $\Psi_n(\boldsymbol{\nu},\boldsymbol{\omega}')\neq 0$ we can bound the last sum in (A3.8) by

$$\begin{aligned}\sum_{p=0}^{n}\Psi_{n,p}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega})\left\|D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega}')\right\| &\le \text{const.}\,(\gamma_n^*)^{-2}\sum_{p=0}^{n}(\gamma_p^*)^{-2}\,|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}| \\ &\le \text{const.}\,(\gamma_n^*)^{-5}|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|, \end{aligned}\qquad (A3.11)$$

where we have used Lemma A1 to bound $\Psi_{n,p}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega})$, and (A3.4) to perform the sum over $p=0,\ldots,n$. Note that in order to profitably use the bound (A3.5) we have to use that the bounds (3.28) and the consequent (3.29) imply analogous bounds also for the propagator divisors $\Delta^{[n]}(x;\varepsilon,\boldsymbol{\omega})$, without the factor ε^2: this follows from the fact that (3.4) defines functions which are smooth in ε and $\boldsymbol{\omega}$.

Still if $\Psi_n(\boldsymbol{\nu},\boldsymbol{\omega}')\neq 0$ we can bound in (A3.8) also the matrices $D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega}')$ and $D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega})$ both proportionally to $(\gamma_n^*)^{-2}$, so that (A3.7) follows.

If $\Psi_n(\boldsymbol{\nu},\boldsymbol{\omega}')=0$ call α_n and α_n' the eigenvalues with minimum absolute value of $D_n(\boldsymbol{\nu},\boldsymbol{\omega})$ and $D_n(\boldsymbol{\nu},\boldsymbol{\omega}')$, respectively. If $|\alpha_n'|\ge|\alpha_n|$ we can proceed as in the previous case, and we obtain the same bound.

Finally if $\Psi_n(\boldsymbol{\nu},\boldsymbol{\omega}')=0$ and $|\alpha_n'|<|\alpha_n|$, we can write

$$g^{[n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')-g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})=-g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}). \qquad (A3.12)$$

Moreover, by the assumed bound (3.28) we have that the difference $D_n(\boldsymbol{\nu},\boldsymbol{\omega}')-D_n(\boldsymbol{\nu},\boldsymbol{\omega})$ is given as in (A3.10), so that $\|D_n(\boldsymbol{\nu},\boldsymbol{\omega}')-D_n(\boldsymbol{\nu},\boldsymbol{\omega})\|\le 4|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|$. Hence the difference between the eigenvalues of $D_n(\boldsymbol{\nu},\boldsymbol{\omega})$ and $D_n(\boldsymbol{\nu},\boldsymbol{\omega}')$ is bounded by $C|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|$, for some constant C; this is again a consequence of Lidskiĭ's theorem. Therefore for $|\alpha_n|\ge 2C|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|$ we can bound $|\alpha_n'|\ge|\alpha_n|-C|\boldsymbol{\nu}|\,|\boldsymbol{\omega}-\boldsymbol{\omega}'|\ge|\alpha_n|/2$, and (A3.7) follows once more by reasoning as before, whereas for $|\alpha_n|<2C|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|$ we can bound in (A3.12)

$$\begin{aligned}\left\|g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})\right\| &\le \left\|D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega})\right\|\le\frac{\left\|D_n^{-1}(\boldsymbol{\nu},\boldsymbol{\omega})\right\|}{|\alpha_n|}|\alpha_n| \\ &\le \frac{1}{|\alpha_n|^2}2C|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|\le\text{const.}(\gamma_n^*)^{-2}|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|,\end{aligned}\qquad (A3.13)$$

and (A3.7) follows also in such a case. ∎

Now we can prove Lemma 10. The proof is by induction on n. One can reason as in Ref. [24]. More precisely one writes

$$\begin{aligned}&\mathcal{M}^{[\le n]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')-\mathcal{M}^{[\le n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}) \\ &=\sum_{p=0}^{n}\Xi_p(\boldsymbol{\nu},\boldsymbol{\omega})\left(M^{[p]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')-M^{[p]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})\right) \\ &\qquad +\sum_{p=0}^{\nu}\sum_{s=0}^{p}\Xi_{p,s}(\boldsymbol{\nu},\boldsymbol{\omega}',\boldsymbol{\omega})\,M^{[p]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}'),\end{aligned}\qquad (A3.14)$$

where the matrices $M^{[p]}(x;\varepsilon,\boldsymbol{\omega})$ are defined as in Ref. [28], formula (5.9). In the first term of the sum the difference $M^{[p]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')-M^{[p]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})$ can be written as a sum over self-energy clusters T of differences of self-energy values $\mathcal{V}_T(\boldsymbol{\omega}'\cdot\boldsymbol{\nu})-\mathcal{V}_T(\boldsymbol{\omega}\cdot\boldsymbol{\nu})$, computed with rotation vectors $\boldsymbol{\omega}'$ and $\boldsymbol{\omega}$, respectively. The latter can be written as a sum of several contributions, each of which is given by the product of a factor $\mathcal{A}(\boldsymbol{\omega}')$ depending on $\boldsymbol{\omega}'$ but no on $\boldsymbol{\omega}$ times a factor[2] $\mathcal{B}(\boldsymbol{\omega})$ depending on $\boldsymbol{\omega}$ but not on $\boldsymbol{\omega}'$ times a difference of propagators $\Delta_\ell(\boldsymbol{\omega}',\boldsymbol{\omega})=g^{[n_\ell]}(\boldsymbol{\omega}'\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega}')-g^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\boldsymbol{\omega})$, with $n_\ell<n$.

The difference $\Delta_\ell(\boldsymbol{\omega}',\boldsymbol{\omega})$ can be bounded according to Lemma A2, proportionally to $|\boldsymbol{\nu}|\,|\boldsymbol{\omega}'-\boldsymbol{\omega}|$, by using that $n_\ell<n$ and the inductive hypothesis.

Moreover the decomposition $\mathcal{A}(\boldsymbol{\omega}')\Delta_\ell(\boldsymbol{\omega}',\boldsymbol{\omega})\mathcal{B}(\boldsymbol{\omega})$ can be made in such a way that the factor $\mathcal{A}(\boldsymbol{\omega}')$ corresponds to a connected subset T_0 of T, while the factor $\mathcal{B}(\boldsymbol{\omega})$ is the product of factorising factors $\mathcal{B}_i(\boldsymbol{\omega})$ corresponding to subsets T_i of T containing lines preceding the lines of T_0 (again we refer to Ref. [19] and [24] for details). The only factor which requires some care is that corresponding to the subset, say T_1, connected to the entering line of T. But this can be easily discussed as in deriving Lemma 4. Indeed one can prove by induction (on the number of nodes) that, given a subset T_1 with the considered structure, one has $N_p(T_1)\le K\,2^{-p}M(T_1)$ for all $p<n$ (note the absence of the summand -1 with the respect the analogous inductive assumption one makes for $N_h(\theta)$ and $N_h(T)$).

Then, by taking into account also the factor $(\gamma^*_{n_\ell})^{-\delta}$ possibly arising from $\Delta_\ell(\boldsymbol{\omega}',\boldsymbol{\omega})$, each factor $\mathcal{B}_i(\boldsymbol{\omega})$ can be bounded by $B_1^{k_i}\mathrm{e}^{-k|\boldsymbol{\nu}_i|/2}$, where k_i is the number of vertices in T_i, $\boldsymbol{\nu}_i$ is the momentum of the line ℓ_i connecting T_i to T_0, and B_1 is some positive constant (the proof proceeds as for Lemma 9). The factor $\mathcal{A}(\boldsymbol{\omega}')$ can be bounded in the same way, and, by taking into account also the factor $(\gamma^*_{n_\ell})^{-\delta}$ possibly arising from $\Delta_\ell(\boldsymbol{\omega}',\boldsymbol{\omega})$ and the factors $\mathrm{e}^{-k|\boldsymbol{\nu}_i|/2}$, it can be bounded by $B_2^{k_0}$, where k_0 is the number of vertices in T_0, and B_2 is some other positive constant (again the proof proceeds as for Lemma 9, but with κ replaced with $\kappa/2$). By writing $\mathcal{A}(\boldsymbol{\omega}')=\mathcal{A}(\boldsymbol{\omega})+(\mathcal{A}(\boldsymbol{\omega}')-\mathcal{A}(\boldsymbol{\omega}))$ one can iterate the construction above for the difference $\mathcal{A}(\boldsymbol{\omega}')-\mathcal{A}(\boldsymbol{\omega})$. The only difference with respect to the previous case is that now the factor $\kappa/2$ is replaced with $\kappa/4$.

All the other terms of the double sum in (A3.14) can be discussed in a similar way, by relying once more on Lemma A2. We omit the details, which can be worked out as in Ref. [24].

Therefore the property (3.28) of Lemma 10 is proved.

Appendix A4. Extensions and generalisations

The analysis performed in this article applies to more general Hamiltonians of the form

$$\mathcal{H}=\mathcal{H}_0(\mathbf{I})+\varepsilon f(\mathbf{I},\boldsymbol{\varphi}), \tag{A4.1}$$

where $(\mathbf{I},\boldsymbol{\varphi})\in\mathbb{A}\times\mathbb{T}^d$, with $\mathbb{A}\subset\mathbb{R}^d$, are conjugate action-angle variables, ε is a real parameter, and the functions $\mathcal{H}_0$ and f are assumed to be analytic in their arguments. We assume also convexity on $\mathcal{H}_0$ and a non-degeneracy condition on f which will be specified later. Here we confine ourselves to sketch the basic arguments: full details will be published elsewhere [21]. Note that in principle weaker conditions on the unperturbed Hamiltonian could be considered; we refer to Ref. [51] for a survey of results under the usual Diophantine conditions.

[2] No relation with the Bryuno function $\mathcal{B}(\omega)$ introduced in Section 2.

We say that a vector $\boldsymbol{\omega}_* \in \mathbb{R}^d$ is s-resonant if it satisfies s resonance conditions $\boldsymbol{\nu}_i \cdot \boldsymbol{\omega}_* = 0$, for s linearly independent integer vectors $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_s$.

With respect to (A4.1), the Hamiltonian (1.1) has a very special form. Even by considering Hamiltonians of the form

$$\mathcal{H} = \frac{1}{2}\mathbf{I}\cdot\mathbf{I} + \varepsilon f(\boldsymbol{\varphi}), \tag{A4.2}$$

it is not always possible to reduce to (1.1). In fact if $\mathcal{I}\colon \mathbb{R}^d \to \mathbb{R}^d$ is the linear operator which transforms $\boldsymbol{\omega}_*$ into $(\boldsymbol{\omega}, \mathbf{0}) \in \mathbb{R}^r \times \mathbb{R}^s$, where $\boldsymbol{\omega}$ has rationally independent components, then the action variables $(\mathbf{A}, \mathbf{B})$ are mixed together and also terms of the form $A_i B_j$ appear.

Though, it is easy to extend the analysis to such a case. And with a little further work, we can consider also Hamiltonians with any unperturbed Hamiltonian $\mathcal{H}_0$ satisfying a convexity property (so that the eigenvalues of the matrix $\det \partial^2_{\mathbf{I}}\mathcal{H}_0$ are all strictly positive). The frequency map $\mathbf{I} \to \boldsymbol{\omega} = \partial\mathcal{H}_0/\partial\mathbf{I}$ is a local diffeomorphism, so that if we fix $\mathbf{I}_0$ in such a way that the corresponding rotation vector $\boldsymbol{\omega}(\mathbf{I})$ is s-resonant, we can find an immersed r-dimensional manifold $\mathbb{M}_r$, with $r = d - s$, containing $\mathbf{I}_0$, on which the s resonance conditions are satisfied. We shall call $\mathbb{M}_r$ a *resonant manifold*

Under the action of the symplectic transformation given by the lift $\mathcal{S}$ of $\mathcal{I}$, we can pass to new coordinates, which we continue to denote with the same symbols, such that in the new coordinates the rotation vector has became $(\boldsymbol{\omega}, \mathbf{0})$. For simplicity we still call $\mathbb{M}_r$ the resonant manifold in the new coordinates.

As we are interested in local properties (in the action variables) we can assume that a system of coordinates adapted to $\mathbb{M}_r$ has been fixed, so that we can write $\mathbf{I} = (\mathbf{A}, \mathbf{B})$ in such a way that $\mathbf{B} = \mathbf{0}$ identifies $\mathbb{M}_r$. For $\varepsilon = 0$ a motion on $\mathbb{M}_r$ is determined by fixing $\mathbf{B} = \mathbf{0}$ and $\mathbf{A} = \mathbf{A}_0$ in such a way that the conjugated angles $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ move according to the law $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \to (\boldsymbol{\alpha} + \boldsymbol{\omega} t, \boldsymbol{\beta})$, with $\boldsymbol{\omega}$ uniquely determined by $\mathbf{A}_0$. This means that the unperturbed lower-dimensional tori can be characterized by the rotation vectors $\boldsymbol{\omega} \in \mathbb{R}^r$ depending on the action variables $\mathbf{A}$. Hence for $\varepsilon \neq 0$ the Hamiltonian describing the system can be written as

$$\mathcal{H} = \mathcal{H}_0(\mathbf{A}, \mathbf{B}) + \varepsilon f(\mathbf{A}, \mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \tag{A4.3}$$

so that the subset of $\mathbb{M}_r \times \mathbb{T}^r$ whose unperturbed invariant tori can be continued under perturbation can be characterised by the set of allowed values of $\boldsymbol{\omega}$, provided the map $\mathbf{A} \mapsto \boldsymbol{\omega}(\mathbf{A})$ is a local diffeomorphism, which is true under our hypotheses.

When the perturbation depends also on the action variables, as in (A4.3), of course one needs both equations for action and angle variables:

$$\begin{cases} \dot{\mathbf{A}} = -\varepsilon \partial_{\boldsymbol{\alpha}} f(\mathbf{A}, \mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \\ \dot{\mathbf{B}} = -\varepsilon \partial_{\boldsymbol{\beta}} f(\mathbf{A}, \mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \\ \dot{\boldsymbol{\alpha}} = \partial_{\mathbf{A}}\mathcal{H}_0(\mathbf{A}, \mathbf{B}) + \varepsilon \partial_{\mathbf{A}} f(\mathbf{A}, \mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \\ \dot{\boldsymbol{\beta}} = \partial_{\mathbf{B}}\mathcal{H}_0(\mathbf{A}, \mathbf{B}) + \varepsilon \partial_{\mathbf{B}} f(\mathbf{A}, \mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\beta}). \end{cases} \tag{A4.4}$$

The main difference with respect to the analysis in Sections 3 and 4 is that the propagators are of the form (3.4), with $(x^2 - \mathcal{M}^{[\le n]}(x; \varepsilon, \omega))^{-1}$ replaced with $(ix - \mathcal{M}^{[\le n]}(x; \varepsilon, \omega))^{-1}$, where we can write $\mathcal{M}^{[\le n]}(x; \varepsilon, \omega) = \mathcal{C} + \mathcal{N}^{[\le n]}(x; \varepsilon, \omega)$, if the matrix $\mathcal{C}$ is given by

$$\mathcal{C} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \partial^2_{\mathbf{A}}\mathcal{H}_0 & \partial_{\mathbf{A}}\partial_{\mathbf{B}}\mathcal{H}_0 & 0 & 0 \\ \partial_{\mathbf{A}}\partial_{\mathbf{B}}\mathcal{H}_0 & \partial^2_{\mathbf{B}}\mathcal{H}_0 & 0 & 0 \end{pmatrix}, \tag{A4.5}$$

and the matrix $\mathcal{N}^{[\leq n]} \equiv \mathcal{N}^{[\leq n]}(x;\varepsilon,\omega)$ is such that, by extracting the dominant order, and setting $x = 0$, one has

$$
\begin{aligned}
\mathcal{N}^{[\leq n]}(0;\varepsilon,\boldsymbol{\omega}) &= \mathcal{N}^{[0]}(0;\varepsilon,\boldsymbol{\omega}) + O(\varepsilon^2), \\
\mathcal{N}^{[0]}(0;\varepsilon,\boldsymbol{\omega}) &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\varepsilon\partial_{\boldsymbol{\beta}}\partial_{\mathbf{A}}f & -\varepsilon\partial_{\boldsymbol{\beta}}\partial_{\mathbf{B}}f & 0 & -\varepsilon\partial_{\boldsymbol{\beta}}^2 f \\ \varepsilon\partial_{\mathbf{A}}^2 f & \varepsilon\partial_{\mathbf{A}}\partial_{\mathbf{B}}f & 0 & \varepsilon\partial_{\mathbf{A}}\partial_{\boldsymbol{\beta}}f \\ \varepsilon\partial_{\mathbf{B}}\partial_{\mathbf{A}}f & \varepsilon\partial_{\mathbf{B}}^2 f & 0 & \varepsilon\partial_{\mathbf{B}}\partial_{\boldsymbol{\beta}}f \end{pmatrix},
\end{aligned} \tag{A4.6}
$$

whereas the other terms depending on x which are not negligible with respect with the dominant ones appear as

$$
\mathcal{N}_1^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega}) \equiv \begin{pmatrix} O(\varepsilon^2 x) & O(\varepsilon^2 x) & O(\varepsilon^2 x^2) & O(\varepsilon^2 x) \\ 0 & 0 & O(\varepsilon^2 x) & 0 \\ 0 & 0 & 0(\varepsilon^2 x) & 0 \\ 0 & 0 & 0(\varepsilon^2 x) & 0 \end{pmatrix}. \tag{A4.7}
$$

Some deep relations turn out to exist between the matrices $\mathcal{M}^{[\leq n]}(x;\varepsilon,\boldsymbol{\omega})\,E$ and their transposed, if E denotes the standard symplectic matrix. Then, by using these relations, one can bound the propagators in terms of the eigenvalues of a suitable symplectic matrix S: for the latter, besides d harmless eigenvalues of order 1 (in ε and x) there are r eigenvalues proportional to x^2, while the other s eigenvalues are of the form

$$
x^2 - \lambda_j^{[0]}(x,\varepsilon,\boldsymbol{\omega}) + O(\varepsilon x) + O(\varepsilon^2), \tag{A4.8}
$$

with $\lambda_j^{[0]}(x,\varepsilon,\boldsymbol{\omega}) = \varepsilon a_{j-r}(\boldsymbol{\omega})$, $j = r+1,\ldots,d$, if $a_1(\boldsymbol{\omega}),\ldots,a_s(\boldsymbol{\omega})$ are the dominant terms of the normal frequencies (which depend also on $\mathbf{A}$, hence on $\boldsymbol{\omega}$, in this case). The aforementioned non-degeneracy condition on f is that the functions $a_j(\boldsymbol{\omega})$ are all strictly positive. The dependence on $\boldsymbol{\omega}$ does not introduce any further difficulties, and in fact Whitney differentiability in $\boldsymbol{\omega}$ (as it appears in the subsequent iterative steps) would be enough to carry on the analysis.

Hence, the situation is very similar to that which has been considered in the previous Sections 3 and 4, and one can reason essentially as there. Of course notations become more cumbersome, because one has to keep trace also of the action variables (which cannot be any more expressed trivially in terms of the angle variables), and more sophisticated diagrammatic rules have to be envisaged; again see Ref. [21] for details. But the basic estimates and arguments are the same, and the same conclusions hold.

Acknowledgments. I am indebted to A. Giuliani for many discussions and comments.

References

[1] 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.

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

[3] J. Bourgain, *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices **11** (1994), 475–497.

[4] J. Bourgain, *On Melnikov's persistency problem*, Math. Rev. Lett. **4** (1997), no. 4, 445–458.

[5] J. Bricmont, A. Kupiainen, A. Schenkel, *Renormalization group and the Melnikov problem for PDE's*, Comm. Math. Phys. **221** (2001), no. 1, 101–140.

[6] H.W. Broer, H. Hanßmann, J. You, *Bifurcations of normally parabolic tori in Hamiltonian systems*, Nonlinearity **18** (2005), no. 4, 1735–1769.

[7] H.W. Broer, G.B. Huitema, M.B. Sevryuk, *Quasi-periodic motions in families of dynamical systems. Order amidst chaos*, Lecture Notes in Mathematics 1645, Springer, Berlin, 1996.

[8] A.D. Bryuno, *Analytic form of differential equations. I*, Trudy Moskov. Mat. Obšč. **25** (1971), 119–262; english translation in Trans. Moscow Math. Soc. **25** (1973), 131–288.

[9] A.D. Bryuno, *Analytic form of differential equations. II*, Trudy Moskov. Mat. Obšč. **26** (1972), 199–239; english translation in Trans. Moscow Math. Soc. **26** (1974), 199–239.

[10] Ch.-Q. Cheng, *Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems*, Comm. Math. Phys. **177** (1996), no. 3, 529–559.

[11] Ch.-Q. Cheng, *Lower-dimensional invariant tori in the regions of instability for nearly integrable Hamiltonian systems*, Comm. Math. Phys. **203** (1999), no. 2, 385–419.

[12] Ch.-Q. Cheng, Sh. Wang, *The surviving of lower-dimensional tori from a resonant torus of Hamiltonian systems*, J. Differential Equations **155** (1999), no. 2, 311–326.

[13] L. Chierchia, G. Gallavotti *Smooth prime integrals for quasi-integrable Hamiltonian systems*, Nuovo Cimento B (11) **67** (1982), no. 2, 277–295.

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

[15] J. Ecalle, B. Vallet, *Bruno Correction and linearization of resonant vector fields and diffeomorphisms*, Math. Z. **229** (1998), no. 2, 249–318.

[16] L.H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **15** (1988), no. 1, 115–147.

[17] L.H. Eliasson, *Hamiltonian systems with linear normal form near an invariant torus*, Nonlinear dynamics (Bologna, 1988), 11–29, World Sci. Publishing, Teaneck, NJ, 1989.

[18] G. Gallavotti, *A criterion of integrability for perturbed nonresonant harmonic oscillators. "Wick ordering" of the perturbations in classical mechanics and invariance of the frequency spectrum*, Comm. Math. Phys. **87** (1982/83), no. 3, 365–383.

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

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

[21] G. Gallavotti, G. Gentile, A. Giuliani, unpublished.

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

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

[24] G. Gentile, *Pure point spectrum for two-level systems in a strong quasi-periodic field*, J. Statist. Phys. **115** (2004), no. 5-6, 1605-2620.

[25] G. Gentile, *Resummation of perturbation series and reducibility for Bryuno skew-product flows*, J. Statist. Phys., in press.

[26] 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.

[27] G. Gentile, M. Bartuccelli, J. Deane, *Quasi-periodic attractors, Borel summability and the Bryuno condition for strongly dissipative systems*, J. Math. Phys., in press.

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

[29] 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.

[30] A. Giorgilli, U. Locatelli, *On classical series expansions for quasi-periodic motions*, Math. Phys. Electron. J. **3** (1997), Paper 5, 25 pp.

[31] H. Hanßmann, *The quasi-periodic centre-saddle bifurcation*, J. Differential Equations **142** (1998), no. 2, 305–370.

[32] H. Hanßmann, *Hamiltonian torus bifurcations related to simple singularities*, Dynamic systems and applications, Vol. 4, 679–685, Dynamic, Atlanta, 2004.

[33] H. Hanßmann, *A survey on bifurcations of invariant tori*, New advances in celestial mechanics and Hamiltonian systems, 109–121, Kluwer/Plenum, New York, 2004.

[34] À. Jorba, R. de la Llave, M. Zou, *Lindstedt series for lower-dimensional tori*, in *Hamiltonian systems with three or more degrees of freedom* (S'Agaró, 1995), 151–167, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 533, Ed. C. Simó, Kluwer Acad. Publ., Dordrecht, 1999.

[35] À. Jorba, J. Villanueva, *On the persistence of lower-dimensional invariant tori under quasi-periodic perturbations*, J. Nonlinear Sci. **7** (1997), no. 5, 427–473.

[36] À. Jorba, J. Villanueva, *On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems*, Nonlinearity **10** (1997), no. 4, 783–822.

[37] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer, Berlin, 1995.

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

[39] S.B. Kuksin, *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Akademiya Nauk SSSR. Funktsional′nyĭ Analiz i ego Prilozheniya **21** (1987), no. 3, 22–37.

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

[41] V.B. Lidskiĭ, *On the characteristic numbers of the sum and product of symmetric matrices*, Dokl. Akad. Nauk SSSR (N.S.) **75** (1950), 769–772.

[42] J. Lopes Dias, *A normal form theorem for Brjuno skew-systems through renormalization*, J. Differential Equations, in press.

[43] V.K. Mel′nikov, *On certain cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function*, Dokl. Akad. Nauk SSSR **165** (1965), 1245–1248; english translation in Soviet Math. Dokl. **6** (1965), 1592–1596.

[44] V.K. Mel′nikov, *A certain family of conditionally periodic solutions of a Hamiltonian systems*, Dokl. Akad. Nauk SSSR **181** (1968), 546–549; english translation in Soviet Math. Dokl. **9** (1968), 882-886.

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

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

[47] F. Rellich, *Perturbation theory of eigenvalue problems*, Gordon and Breach Science Publishers, New York-London-Paris, 1969.

[48] 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., 357, New York Acad. Sci., New York, 1980.

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

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

[51] M.B. Sevryuk, *The classical KAM theory at the dawn of the twenty-first century*, Mosc. Math. J. **3** (2003), no. 3, 1113–1144.

[52] M.B. Sevryuk, *Partial preservation of frequencies in KAM theory*, Nonlinearity **19** (2006), no. 5, 1099–1140.

[53] M.B. Sevryuk, *Invariant tori in quasi-periodic non-autonomous dynamical systems via Hermn's method*, Discrete Contin. Dynam. Systems, in press.

[54] W. Thirring, *A course in mathematical physics. Vol. I. Classical dynamical systems*, Springer, New York-Vienna, 1978.

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

[56] H. Whitney, *Analytic extensions of differential functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934), no. 1, 63–89.

[57] J. Xu, *Persistence of elliptic lower-dimensional invariant tori for small perturbation of degenerate integrable Hamiltonian systems*, J. Math. Anal. Appl. **208** (1997), no. 2, 372–387.

[58] J. Xu, J. You, *Persistence of lower dimensional tori under the first Melnikov's non-resonance condition*, J. Math. Pures Appl. (9) **80** (2001), no. 10, 1045–1067.

[59] J.-Ch. Yoccoz, *Théorème de Siegel, nombres de Bruno et polinômes quadratiques*, Astérisque **231** (1995), 3–88.

[60] J. You, *Perturbations of lower-dimensional tori for Hamiltonian systems*, J. Differential Equations **152** (1999), no. 1, 1–29.