



# Resonant tori of arbitrary codimension for quasi-periodically forced systems

Livia Corsi and Guido Gentile

**Abstract.** We consider a system of rotators subject to a small quasi-periodic forcing. We require the forcing to be analytic and satisfy a time-reversibility property and we assume its frequency vector to be Bryuno. Then we prove that, without imposing any non-degeneracy condition on the forcing, there exists at least one quasi-periodic solution with the same frequency vector as the forcing. The result can be interpreted as a theorem of persistence of lower-dimensional tori of arbitrary codimension in degenerate cases.

**Mathematics Subject Classification.** 34C45, 34C23, 37J40, 70H08, 70H33.

## 1. Introduction

In this paper we deal with the problem of persistence of lower-dimensional tori for nearly integrable Hamiltonian systems without assumptions of non-degeneracy on the perturbation. Consider an analytic Hamiltonian function of the form

$$H(\varphi, J) = H_0(J) + \varepsilon f(\varphi, J), \quad (1.1)$$

where  $(J, \varphi) \in \mathbb{R}^n \times \mathbb{T}^n$  are conjugate action-angle variables, with  $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ , and  $\varepsilon \in \mathbb{R}$ . The function  $H_0$  will be referred to as the unperturbed Hamiltonian, while  $f$  will be called the perturbation. For  $\varepsilon = 0$  the phase space is foliated into  $n$ -dimensional invariant tori; on each torus the action  $J$  is constant and the motion is quasi-periodic with frequency vector  $\omega_0(J) := \partial_J H_0(J)$ .

We say that a frequency vector  $\omega_0 \in \mathbb{R}^n$  is resonant with multiplicity  $r$  if there is a rank  $r$  subgroup  $G$  of  $\mathbb{Z}^n$  such that  $\omega_0 \cdot \nu = 0$  for all  $\nu \in G$  and  $\omega_0 \cdot \nu \neq 0$  for all  $\nu \in \mathbb{Z}^n \setminus G$ . It is a long standing conjecture that, in nearly integrable Hamiltonian systems with Hamiltonian function of the form (1.1),

---

Partially supported by the European Research Council under FP7 “Hamiltonian PDEs and small divisor problems: a dynamical systems approach”.

under a convexity assumption on  $H_0$ , for any  $r \leq n - 1$  and for most families of unperturbed tori whose frequency vectors are resonant with multiplicity  $r$ , at least  $r + 1$  tori survive any perturbation for  $\varepsilon$  small enough [8, 20]. When this happens one says that the surviving torus has codimension  $r$ . In the case  $r = n - 1$ , where the surviving tori reduce to closed orbits, this has been shown by Bernstein and Katok [2]. The case  $r < n - 1$  is harder, because of the presence of small divisors. For  $r = 1$  the conjecture has been proved by Cheng [6], while only partial results exist for  $1 < r < n - 1$ , requiring non-degeneracy hypotheses on the perturbation [7, 16, 24].

A somewhat different but related problem is the following. Let the frequency vector  $\omega_0$  be resonant with multiplicity  $r$  and fix  $J_0$  such that  $\omega_0(J_0) = \omega_0$ . Then, after a suitable change of coordinates, one can rewrite (1.1) as

$$H(\alpha, \beta, A, B) = H_0(A, B) + \varepsilon f(\alpha, \beta, A, B), \quad (1.2)$$

where  $(A, B) \in \mathbb{R}^d \times \mathbb{R}^r$  and  $(\alpha, \beta) \in \mathbb{T}^d \times \mathbb{T}^r$  are such that  $d + r = n$  and  $\omega_0(A_0, B_0) = (\omega, 0)$ , with  $\omega \in \mathbb{R}^d$  non-resonant; not to overwhelm the notation, we are denoting with the same symbol the Hamiltonian obtained by composing the original Hamiltonian with the change of coordinates. Assume some Diophantine condition on  $\omega$  and consider the particular family of tori of the unperturbed system with that frequency vector. Then look at the fate of the family when the perturbation is switched on. For Hamiltonian functions (1.2), again assuming  $H_0$  to be convex, in general one expects only that at least one torus in the family survives any perturbation  $f$  for  $\varepsilon$  small enough. In fact, the persistence of one such torus is known to hold for any  $r$  only if one assumes a non-degeneracy condition on  $f$ . More precisely, the perturbation averaged on the unperturbed torus is taken to have a non-degenerate maximum [24]; see also [14, 19], where the analyticity properties of the tori in the perturbation parameter are investigated. By weakening the non-degeneracy assumptions on the perturbation partial results have been obtained for  $r = 1$  in [15, 25].

Despite the fact that assumptions on the perturbations are commonly believed to be unnecessary, so far the only result existing in the literature for Hamiltonian functions (1.2) is in the case of tori with codimension 1 and is due to Cheng [5]. An analogous result of persistence of at least one lower-dimensional torus of codimension 1, without assuming any hypothesis on the perturbation, has been recently proved for a suitable class of non-convex  $H_0$  with a saddle point [9, 21, 22].

In the present paper we consider a system with “partially isochronous” unperturbed Hamiltonian and perturbation depending only on the angles, namely

$$H(\alpha, \beta, A, B) = \omega \cdot A + \frac{1}{2}B^2 + \varepsilon f(\alpha, \beta), \quad (1.3)$$

and, under a mild Diophantine condition on  $\omega \in \mathbb{R}^d$ , we prove the existence of at least one quasi-periodic solution with frequency vector  $\omega$  for the corresponding Hamilton equations, requiring on  $f$  a parity condition with respect to the variable  $\alpha$ ; see Hypothesis 2 below. Such assumption can be seen as a time-reversibility condition on the Hamiltonian system.

For non-convex Hamiltonians (1.3) with  $r = 1$ , the same result of persistence of at least one invariant torus has been proved in [11] without assuming Hypothesis 2 on the perturbation; the result has been extended to more general one-dimensional systems in [12]. Strictly speaking, Cheng's result does not imply the result in [11, 12], because the unperturbed Hamiltonian is not convex; however Cheng's method can be adapted to such a case; see [13] for an explicit implementation. Moreover in [11, 12] the frequency vector is allowed to satisfy the Bryuno condition, a weaker condition with respect to the standard Diophantine condition considered in both [5] and [13]; in [12] non-Hamiltonian systems have been considered as well: in that case further conditions have to be assumed in general on the perturbation (in the Hamiltonian case such conditions are automatically satisfied).

For the general case (1.1) with  $r > 1$ , the problem of persistence of at least one lower-dimensional torus without requiring any non-degeneracy assumption on the perturbation is still open. As far as we know, Theorem 1 below is the first result in that direction.

The time reversibility hypothesis we assume on the perturbation is a symmetry property, not a non-degeneracy condition, and aims to ensure a suitable cancellation (precisely the one in Lemma 3.5) that we need in the proof. Indeed the rest of the proof could be easily adapted to the general case, whereas at the moment we are not able to obtain a result analogous to Lemma 3.5 without the reversibility hypothesis, except when  $r = 1$ . Ultimately this is related to a commutator between two matrices appearing in the small divisor; see Remark 3.7. In the non-degenerate case the contribution of the commutator is negligible, when  $r = 1$  the matrices are just numbers and in the reversible case the matrices are equal to each other: thus in all cases above either the commutator is negligible or it vanishes. In the general case there is no reason for the commutator to vanish; therefore in principle it is even possible that there are Diophantine frequencies for which no lower-dimensional torus persists, when  $r \geq 1$ , for degenerate and non-reversible perturbations.

### 1.1. Statement of the results

Let us write the Hamiltonian (1.3) in the non-autonomous form

$$H(\beta, B, t) = \frac{1}{2}B^2 + \varepsilon f(\omega t, \beta), \quad (1.4)$$

where  $(B, \beta) \in \mathbb{R}^r \times \mathbb{T}^r$  are conjugate action-angle variables,  $\omega \in \mathbb{R}^d$  is the frequency vector,  $f$  is analytic in the complexified torus  $\mathbb{T}_\xi^{d+r} := \{\varphi \in \mathbb{C}^{d+r} : \operatorname{Re} \varphi_k \in \mathbb{T}, |\operatorname{Im} \varphi_k| \leq \xi\}$  and  $\varepsilon \in \mathbb{R}$  is a small parameter (*perturbation parameter*). Without loss of generality, we may and shall assume that  $\omega$  has rationally independent components and  $\varepsilon \geq 0$ .

The corresponding Hamilton equations can be written as

$$\ddot{\beta} = -\varepsilon \partial_\beta f(\omega t, \beta) \quad (1.5)$$

and hence describe an  $r$ -dimensional system of rotators with a quasi-periodic forcing.

We consider  $d > 1$  in the following, since the periodic case  $d = 1$  is well understood; see [10] and references within. For  $\omega \in \mathbb{R}^d$  define

$$\alpha_m(\omega) := \inf_{0 < |\nu| \leq 2^m} |\omega \cdot \nu|, \quad \mathfrak{B}(\omega) := \sum_{m=0}^{\infty} \frac{1}{2^m} \log \frac{1}{\alpha_m(\omega)}, \quad (1.6)$$

where  $\cdot$  denotes the inner product in  $\mathbb{R}^d$  and  $|\nu| := |\nu_1| + \dots + |\nu_d|$ , if  $\nu_1, \dots, \nu_d$  are the components of  $\nu$ . We assume the following hypotheses on  $\omega$  and  $f$ .

**Hypothesis 1.**  $\omega \in \mathbb{R}^d$ , with  $d > 1$ , satisfies the condition  $\mathfrak{B}(\omega) < \infty$ .

**Hypothesis 2.**  $f$  is even in  $\alpha$ , that is  $f(-\alpha, \beta) = f(\alpha, \beta)$ .

Hypothesis 1 is a weak Diophantine condition on the frequency vector, known in the literature as the *Bryuno condition*. Hypothesis 2 is a *time reversibility* condition on the forcing. We shall prove the following result.

**Theorem 1.** *Consider the Hamiltonian (1.4) and assume Hypotheses 1 and 2. Then for  $\varepsilon$  small enough there exists at least one quasi-periodic solution to the corresponding Hamilton equation (1.5) with frequency vector  $\omega$ .*

### 1.2. Informal presentation of the proof

We look for quasi-periodic solutions to (1.5) with frequency vector  $\omega$  (*response solutions*), so we split

$$\beta(t) = \beta_0 + b(\omega t), \quad [b(\cdot)]_0 := \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} d\alpha b(\alpha) = 0,$$

and separate (1.5) into the so-called “range” and “bifurcation” equations

$$(\omega \cdot \partial_\alpha)^2 b + \varepsilon (\partial_\beta f(\alpha, \beta_0 + b) - [\partial_\beta f(\cdot, \beta_0 + b(\cdot))]_0) = 0, \quad (1.7a)$$

$$[\varepsilon \partial_\beta f(\cdot, \beta_0 + b(\cdot))]_0 = 0. \quad (1.7b)$$

The range equation (1.7a) can be seen as an implicit function equation of the form

$$D_\omega b + \varepsilon F(\beta_0 + b) = 0, \quad (1.8)$$

where the differential operator  $D_\omega := (\omega \cdot \partial_\alpha)^2$  is diagonal on the Fourier basis  $\{e^{i\nu \cdot \alpha}\}_{\nu \in \mathbb{Z}^d}$ . As it is well known, the standard implicit function theorem fails to apply since the inverse of  $D_\omega$  is unbounded (small divisor problem). In order to overcome this difficulty, one can implement a fast iterative scheme (KAM, Nash–Moser or Renormalisation Group). At any iterative step one has to impose Diophantine conditions involving the parameter  $\beta_0$ , so as to control the small divisors: at the  $n$ -step one requires  $\beta_0$  to be restricted to some set  $\mathcal{C}_n \subset \mathbb{T}^r$ . Eventually one obtains a function  $b = b(\alpha, \varepsilon, \beta_0)$  which solves the range equation for  $\beta_0$  in a set  $\mathcal{C} = \bigcap_{n=0}^\infty \mathcal{C}_n$ . Then one passes to the bifurcation equation (1.7b). Formally (i.e. by assuming that a solution to (1.7a) exists for  $\beta_0$  in a non-empty set  $\mathcal{C}$ ), by relying on the variational structure of the Hamilton equation one has the identity

$$[\varepsilon \partial_\beta f(\cdot, \beta_0 + b(\cdot, \varepsilon, \beta_0))]_0 = -\partial_{\beta_0} L(\varepsilon, \beta_0), \quad (1.9)$$

where  $L(\varepsilon, \beta_0)$  is the average with respect to  $\alpha$  of the Lagrangian computed along  $b(\alpha, \varepsilon, \beta_0)$ ; see for instance [1, 3, 23]. However, without requiring any assumption on the perturbation, it is difficult to have any control on the set  $\mathcal{C}$ . Even assuming that the solution can be extended to the torus  $\mathbb{T}^r$ , so as to ensure  $L(\varepsilon, \beta_0)$  to have a critical point, it is by no way obvious that such a critical point falls inside  $\mathcal{C}$ . In fact, the non-degeneracy condition usually assumed in the literature aims exactly to make the left hand side of (1.9) to vanish for a suitable value of  $\beta_0$  for which the Diophantine conditions are satisfied. So, if no non-degeneracy condition is assumed on the perturbation, we have to proceed in a different way.

In the present paper, we use a Renormalisation Group approach; see for instance [4, 17]. We proceed as follows. Choose a subsequence  $\alpha_{m_n}(\omega)$ ,  $n \geq 0$ , in such a way that it is strictly decreasing, and define an integer vector  $\nu \in \mathbb{Z}^d \setminus \{0\}$  to be on scale  $n \geq 1$  if  $\alpha_{m_{n+1}}(\omega) \leq |\omega \cdot \nu| < \alpha_{m_n}(\omega)$  and on scale  $n = 0$  if  $|\omega \cdot \nu| \geq \alpha_{m_0}(\omega)$  (actually, for technical reasons, we shall use a smooth partition). Roughly, the higher the scale of  $\nu$  is, the smaller the quantity  $\alpha_n(\omega)$  associated with  $\nu$ . Introducing the notation

$$b = \sum_{n \geq 0} b_n, \quad b_{\geq m} = \sum_{n \geq m} b_n, \quad b_n = \sum_{\nu \text{ on scale } n} e^{i\nu \cdot \omega t} b_{n,\nu},$$

the idea is to solve the range equation “scale by scale”, by fixing iteratively the “components”  $b_n$  in terms of the “corrections”  $b_{\geq n+1}$ , i.e.  $b_n = b_n(b_{\geq n+1})$ . Starting from scale  $n = 0$  we obtain from (1.8)

$$(\omega \cdot \nu)^2 b_{0,\nu} - [\varepsilon F(\beta_0 + b_0 + b_{\geq 1})]_{\nu} = 0,$$

so that linearising at  $b = 0$  we have

$$(\omega \cdot \nu)^2 b_{0,\nu} - [\varepsilon \partial F(\beta_0) b_0]_{\nu} - [\varepsilon \partial F(\beta_0) b_{\geq 1}]_{\nu} + O(b^2) = 0,$$

with  $\partial$  denoting derivative with respect to the argument. Then we correct the differential operator by adding the diagonal part (with respect to the Fourier basis) of  $-\varepsilon \partial F(\beta_0)$ ; denoting  $\mathcal{M}_0 := \text{diag}_{\nu}(\varepsilon \partial F(\beta_0))$  and  $\mathcal{N}_0 := \varepsilon \partial F(\beta_0) - \mathcal{M}_0$ , we write

$$\mathcal{D}_0(\nu) b_{0,\nu} = [\mathcal{N}_0 b_0]_{\nu} + [\varepsilon \partial F(\beta_0) b_{\geq 1}]_{\nu} + O(b^2), \quad \mathcal{D}_0(\nu) := (\omega \cdot \nu)^2 \mathbf{1} - \mathcal{M}_0,$$

where  $\mathbf{1}$  is the  $r \times r$  identity matrix. If the equation can be solved, we obtain  $b_0$  expressed in terms of  $b_{\geq 1}$ . Iterating, at the  $n$ -th step we write  $b = b_{\leq n-1}(b_{\geq n}) + b_{\geq n}$ , where

$$b_{\leq n-1}(b_{\geq n}) := \sum_{m \leq n-1} b_{m,n}(b_{\geq n}),$$

with each  $b_{m,n}(b_{\geq n})$  recursively expressed in terms of  $b_{\geq n}$ ; an explicit computation gives  $b_{n-1,n}(x) = b_{n-1}(x)$  and  $b_{m,n}(x) = b_m(b_{m+1,n}(x) + \dots + b_{n-1,n}(x) + x)$  for  $0 \leq m \leq n-2$ . Hence, for  $\nu$  on scale  $n$ , (1.8) reduces to

$$(\omega \cdot \nu)^2 b_{n,\nu} - [\varepsilon F(\beta_0 + b_{\leq n-1}(b_{\geq n}) + b_{\geq n})]_{\nu} = 0.$$

We linearise at  $b_{\geq n} = 0$ , so as to obtain

$$(\omega \cdot \nu)^2 b_{n,\nu} - [\varepsilon \partial F(\beta_0 + b_{\leq n-1}(0))(\mathbf{1} + \partial b_{\leq n-1}(0)) (b_n + b_{\geq n+1})]_{\nu} + O(b_{\geq n}^2) = 0,$$

and add the correction  $\text{diag}_\nu(-\varepsilon\partial F(\beta_0 + b_{\leq n-1}(0))(\mathbb{1} + \partial b_{\leq n-1}(0)))$  to the differential operator. In other words, the corrections to  $(\omega \cdot \nu)^2$  at each step are diagonal with respect to the Fourier indices  $\nu$ . Moreover they turn out to depend on  $\nu$  only through  $\omega \cdot \nu$ , so that we can denote the correction  $-\mathcal{M}_n(\omega \cdot \nu; \varepsilon, \beta_0)$ . In conclusion, at each step  $n$ , we have to deal with an equation of the form

$$\mathcal{D}_n(\nu) b_{n,\nu} = [\mathcal{N}_n b_n]_\nu + [\varepsilon\partial F(\beta_0 + b_{\leq n-1}(0))(\mathbb{1} + \partial b_{\leq n-1}(0)) b_{\geq n+1}]_\nu + O(b_{\geq n}^2), \tag{1.10}$$

where  $\mathcal{N}_n := \varepsilon\partial F(\beta_0 + b_{\leq n-1}(0))(\mathbb{1} + \partial b_{\leq n-1}(0)) - \mathcal{M}_n(\omega \cdot \nu; \varepsilon, \beta_0)$  and the operator  $\mathcal{D}_n(\nu)$  has the form

$$\mathcal{D}_n(\nu) = (\omega \cdot \nu)^2 \mathbb{1} - \mathcal{M}_n(\omega \cdot \nu; \varepsilon, \beta_0). \tag{1.11}$$

By using a version of the Siegel–Bryuno bound (see Lemma 2.1), we show that the convergence of such iterative scheme would follow if one had a bound like

$$\|(\omega \cdot \nu)^2 \mathbb{1} - \mathcal{M}_n(\omega \cdot \nu; \varepsilon, \beta_0)\| \geq \frac{(\omega \cdot \nu)^2}{2}, \tag{1.12}$$

with  $\|\cdot\|$  denoting the  $r$ -dimensional  $L^2$  operator norm. A second order Taylor expansion of (1.11) at  $\omega \cdot \nu = 0$  gives

$$(\omega \cdot \nu)^2 \mathbb{1} - (\mathcal{M}_n(0; \varepsilon, \beta_0) + \partial \mathcal{M}_n(0; \varepsilon, \beta_0)(\omega \cdot \nu) + O(\varepsilon(\omega \cdot \nu)^2)). \tag{1.13}$$

The time reversibility assumption in Hypothesis 2 ensures that  $\mathcal{M}_n(\omega \cdot \nu; \varepsilon, \beta_0)$  is even in  $\omega \cdot \nu$ , so that the linear term in (1.13) vanishes identically and hence one has to control  $\mathcal{M}_n(0; \varepsilon, \beta_0)$  only. Formally (i.e. assuming the bound (1.12)) one can show that, up to corrections,  $\mathcal{M}_n(0; \varepsilon, \beta_0) = -\partial_{\beta_0}^2 L_n(\varepsilon, \beta_0)$ , where  $L_n(\varepsilon, \beta_0)$  is the average of the Lagrangian computed along the  $n$ -step approximate solution  $b^{\leq n} := b_{\leq n}(0)$ . If one could take  $\beta_0$  as a maximum of  $L_n(\varepsilon, \beta_0)$ , the eigenvalues of  $\mathcal{M}_n(0; \varepsilon, \beta_0)$  would be all non-positive, so implying the bound (1.12). However, we are not really allowed to proceed that way since  $L_n(\varepsilon, \beta_0)$  is not defined for all  $\beta_0$  and in any case  $\beta_0$  cannot be fixed at a different value at every intermediate step  $n$ .

The idea is to define recursively an auxiliary function  $\bar{b}(\alpha; \varepsilon, \beta_0)$ , obtained recursively from (1.10) by modifying the operators (1.11) in such a way that a bound like (1.12) automatically holds. To this aim we replace recursively  $\mathcal{M}_n(0; \varepsilon, \beta_0)$  with  $-\partial_{\beta_0}^2 \bar{L}_n(\varepsilon, \beta_0) \bar{\xi}_n(\varepsilon, \beta_0)$ , where  $\bar{L}_n(\varepsilon, \beta_0)$  is the average of the Lagrangian computed along the  $(n - 1)$ -step approximation of the auxiliary function and  $\bar{\xi}_n$  is a suitable cut-off function. The presence of the cut-off functions allows the auxiliary function to be well defined for all  $\beta_0 \in \mathbb{T}^r$  and thus one can pass to the limit  $n \rightarrow \infty$ . Of course for  $\beta_0$  varying in  $\mathbb{T}^r$  the function  $\bar{b}(\alpha; \varepsilon, \beta_0)$  is no longer a solution of the range equation. Then we show recursively that taking  $\beta_0 = \bar{\beta}_0(\varepsilon)$  as a maximum of  $\bar{L}_\infty(\varepsilon, \beta_0)$ , one has  $\mathcal{M}_n(0; \varepsilon, \bar{\beta}_0(\varepsilon)) = -\partial_{\beta_0}^2 \bar{L}_n(\varepsilon, \bar{\beta}_0(\varepsilon)) \bar{\xi}_n(\varepsilon, \bar{\beta}_0(\varepsilon))$ , implying that  $b(\alpha; \varepsilon, \bar{\beta}_0(\varepsilon)) = \bar{b}(\alpha; \varepsilon, \bar{\beta}_0(\varepsilon))$ . Therefore both the range and the bifurcation equations are satisfied.

We note that, unlike [11], because of the higher dimension we cannot consider directly the linearisation  $\mathcal{M}_n(0; \varepsilon, \beta_0)$  and then construct its second

antiderivatives with respect to  $\beta_0$ . On the contrary, we have to take full advantage of the variational structure of the Hamilton equations, hence compute the Lagrangian along the solution of the range equation and thence exploit the fact that  $\mathcal{M}_n(0; \varepsilon, \beta_0)$  is formally its second derivative. As explained above, it is not obvious that this can be achieved, because there is a priori no control on the set of values of  $\beta_0$  for which all those functions are well defined: to do this we have to introduce the auxiliary function for which everything is well defined and eventually to prove that, by carefully choosing  $\beta_0$ , such a function reduces to the true solution.

Finally we stress again that the reversibility assumption is not used to solve the bifurcation equation, but in order to obtain the invertibility of the operators  $\mathcal{D}_n(\nu)$ .

## 2. Multiscale analysis

We start by introducing the tree formalism; this is an efficient way to take into account the combinatorics in the expression of the formal solution  $b$  described in Sect. 1.2. We shall associate a numerical value with each tree in such a way that the sum over trees of such values equals the components  $b_j$  of  $b$ , with  $j \in \{1, \dots, r\}$ . As we shall see, many notations are very similar to those in [11] apart from the fact that here we are dealing with an  $r$ -dimensional problem and hence we need to take into account the components  $j \in \{1, \dots, r\}$ . In particular some of the results in [11] do not depend on  $r$  and hence hold word by word also in the present case. Some other bounds discussed in the first part of the next section differ from [11] only because of the presence of  $r$ -dependent constants, but they can be proved by reasoning essentially in the same way.

### 2.1. Oriented trees

An *oriented tree*  $\theta$  is a graph (that is a set of points and lines connecting them) with no cycle, such that all the lines are oriented toward a single point (*root*) which has only one incident line  $\ell_\theta$  (*root line*). All the points in a tree except the root are called *nodes*. The orientation of the lines in a tree induces a partial ordering relation ( $\preceq$ ) between the nodes and the lines: we can imagine that each line carries an arrow pointing toward the root. Given two nodes  $v$  and  $w$ , we write  $w \prec v$  every time  $v$  is along the path (of lines) which connects  $w$  to the root. Given a node  $v$ , we denote by  $\pi(v)$  the unique point immediately following  $v$ .

We denote by  $N(\theta)$  and  $L(\theta)$  the sets of nodes and lines in  $\theta$  respectively. Since a line  $\ell \in L(\theta)$  is uniquely identified by the node  $v$  which it leaves, we may write  $\ell = \ell_v$ . We write  $\ell_w \prec \ell_v$  if  $w \prec v$ , and  $w \prec \ell = \ell_v$  if  $w \preceq v$ ; if  $\ell$  and  $\ell'$  are two comparable lines, i.e.  $\ell' \prec \ell$ , we denote by  $\mathcal{P}(\ell, \ell')$  the (unique) path of lines connecting  $\ell'$  to  $\ell$ , with  $\ell$  and  $\ell'$  not included (in particular  $\mathcal{P}(\ell, \ell') = \emptyset$  if  $\ell'$  enters the node  $\ell$  exits).

Given a tree  $\theta$  we call *order* of  $\theta$  the number  $k(\theta) = |N(\theta)| = |L(\theta)|$  (for any finite set  $S$  we denote by  $|S|$  its cardinality). A subset  $T \subset \theta$  is a *subgraph* of  $\theta$  if it is formed by a set of nodes  $N(T) \subseteq N(\theta)$  and a set of lines

$L(T) \subseteq L(\theta)$  in such a way that  $N(T) \cup L(T)$  is connected. If  $T$  is a subgraph of  $\theta$  we call *order* of  $T$  the number  $k(T) = |N(T)|$ . We say that a line enters  $T$  if it connects a node  $v \notin N(T)$  to a node  $w \in N(T)$ , and we say that a line exits  $T$  if it connects a node  $v \in N(T)$  to a node  $w \notin N(T)$  or to the root (which is not included in  $T$  in this case). Of course, if a line  $\ell$  enters or exits  $T$ , then  $\ell \notin L(T)$ .

**2.2. Labels**

In (1.4) we can write

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

where  $|\partial_\beta^s f_\nu(\beta)| \leq s! \Phi_0 \Phi_1^s e^{-\xi|\nu|}$  for suitable positive constants  $\Phi_0$  and  $\Phi_1$ , by the analyticity assumption on  $f$ ; note that  $f_{-\nu}(\beta) = f_\nu(\beta)$  by Hypothesis 2.

With each node  $v \in N(\theta)$  we associate a *mode* label  $\nu_v \in \mathbb{Z}^d$  and we denote by  $s_v$  the number of lines entering  $v$ . With each line  $\ell$  we associate a *momentum*  $\nu_\ell \in \mathbb{Z}_*^d$ , except for the root line which can have either zero momentum or not (i.e.  $\nu_\ell \in \mathbb{Z}^d$ ), and a pair of *component labels*  $(e_\ell, u_\ell) \in \{1, \dots, r\}^2$ . We call *total momentum* of  $\theta$  the momentum associated with  $\ell_\theta$ . Finally, we associate with each line  $\ell$  also a *scale label* such that  $n_\ell = -1$  if  $\nu_\ell = 0$ , while  $n_\ell \in \mathbb{Z}_+$  if  $\nu_\ell \neq 0$ ; note that one can have  $n_\ell = -1$  only if  $\ell$  is the root line of  $\theta$ . We impose the following *conservation law*

$$\nu_\ell = \sum_{\substack{w \in N(\theta) \\ w \prec \ell}} \nu_w. \tag{2.1}$$

A *labelled oriented tree* is an oriented tree with labels associated with its nodes and lines. In the following, for simplicity’s sake, we shall call trees tout court the labelled oriented trees and we shall term *unlabelled tree* the oriented trees without labels. We shall say that two trees are *equivalent* if they can be transformed into each other by continuously deforming the lines in such a way that these do not cross each other and also labels match. This provides an equivalence relation on the set of the trees. From now on we shall call trees such equivalence classes.

**2.3. Clusters and self-energy clusters**

A *cluster*  $T$  on scale  $n$  is a maximal subgraph of a tree  $\theta$  such that all the lines have scales  $n' \leq n$  and there is at least a line with scale  $n$ . The lines entering the cluster  $T$  and the line coming out from it (unique if existing at all) are called the *external* lines of  $T$ .

A *self-energy cluster* is a cluster  $T$  such that (i)  $T$  has only one entering line  $\ell'_T$  and one exiting line  $\ell_T$ , (ii) one has  $\nu_{\ell_T} = \nu_{\ell'_T}$  and hence  $\sum_{v \in N(T)} \nu_v = 0$  by (2.1).

For any self-energy cluster  $T$ , set  $\mathcal{P}_T = \mathcal{P}(\ell_T, \ell'_T)$ . We shall say that a self-energy cluster is on scale  $-1$ , if  $N(T) = \{v\}$  with of course  $\nu_v = 0$  (so that



$\mathcal{P}_T = \emptyset$ ). Given a self-energy cluster  $T$ , for all  $\ell \in \mathcal{P}_T$  one can write

$$\nu_\ell = \nu_\ell^0 + \nu_{\ell'_T}, \quad \nu_\ell^0 := \sum_{\substack{w \in N(T) \\ w \rightarrow \ell}} \nu_w. \tag{2.2}$$

Note that the momenta of the lines along  $\mathcal{P}_T$  are the only labels of  $T$  depending on the labels outside  $T$ . We say that two self-energy clusters  $T_1$  and  $T_2$  have the same *structure* if setting  $\nu_{\ell'_{T_1}} = \nu_{\ell'_{T_2}} = 0$  one has  $T_1 = T_2$ . Of course this provides an equivalence relation on the set of all self-energy clusters; from now on we shall call self-energy clusters tout court such equivalence classes.

**2.4. Renormalised trees, node factors and propagators**

Once we associate a numerical value with each labelled tree, it will be clear that a self-energy cluster represents a contribution to the diagonal of the linearised vector field described in Sect. 1.2. Our resummation procedure implies that the self-energy clusters must appear in the small divisor and not in the vector field. Therefore we shall consider *renormalised trees*, i.e. trees in which no self-energy clusters appear; analogously a *renormalised subgraph* is a subgraph of a tree  $\theta$  which does not contains any self-energy cluster. Denote by  $\Theta_{k,\nu,j}$  the set of renormalised trees with order  $k$ , total momentum  $\nu$  and  $e_{\ell_\theta} = j$ , and by  $\mathfrak{R}_{n,u,e}$  the set of renormalised self-energy clusters on scale  $n$  such that  $u_{\ell_T} = u$  and  $e_{\ell'_T} = e$ .

For any  $\theta \in \Theta_{k,\nu,j}$  and any subgraph  $S \subseteq \theta$  we associate with each node  $v \in N(S)$  a *node factor*

$$\mathcal{F}_v(\beta_0) := \frac{1}{s_v!} \partial_{\beta_{u_{\ell_v}}} \left( \prod_{\substack{w \in N(S) \\ \pi(w)=v}} \partial_{\beta_{e_{\ell_w}}} \right) f_{\nu_v}(\beta_0). \tag{2.3}$$

Introduce a partition of unity as follows. Given a decreasing sequence  $\rho_n, n \in \mathbb{N} \cup \{0\}$ , of positive numbers with  $\rho_{n+1} \leq \rho_n/2$ , let  $\chi : \mathbb{R} \rightarrow \mathbb{R}$  be a  $C^\infty$  function, non-increasing for  $x \geq 0$  and non-decreasing for  $x < 0$ , such that

$$\chi(x) = \begin{cases} 1, & |x| \leq 1/2, \\ 0, & |x| \geq 1, \end{cases}$$

and set  $\chi_n(x) = \chi(x/\rho_n)$  for  $n \geq 0$  and  $\chi_{-1}(x) = 1$ . Set also  $\Psi_n(x) = \chi_{n-1}(x) - \chi_n(x)$  for  $n \geq -1$ ; see Figure 1 in [11].

Next, we introduce the sequences  $\{m_n, p_n\}_{n \geq 0}$ , with  $m_0 = 0$  and, for all  $n \geq 0$ ,  $m_{n+1} = m_n + p_n + 1$ , where  $p_n := \max\{q \in \mathbb{Z}_+ : \alpha_{m_n}(\omega) < 2\alpha_{m_n+q}(\omega)\}$ , with  $\alpha_m(\omega)$  defined in (1.6). The subsequence  $\{\alpha_{m_n}(\omega)\}_{n \geq 0}$  of  $\{\alpha_m(\omega)\}_{m \geq 0}$  is decreasing. A convenient partition of unity is then obtained by choosing  $\rho_n = \alpha_{m_n}(\omega)/8$ , which is the same choice as in [11].

For  $n \geq 0$ , define formally

$$\mathcal{G}_{\varepsilon,u}^{[n]}(x; \varepsilon, \beta_0) := \Psi_n(x) \left[ \left( x^2 \mathbf{1} - \mathcal{M}^{[n-1]}(x; \varepsilon, \beta_0) \right)^{-1} \right]_{\varepsilon,u}, \tag{2.4a}$$

$$\mathcal{M}^{[n-1]}(x; \varepsilon, \beta_0) := \sum_{q=-1}^{n-1} \chi_q(x) M^{[q]}(x; \varepsilon, \beta_0), \tag{2.4b}$$

$$M_{u,e}^{[q]}(x; \varepsilon, \beta_0) := \sum_{T \in \mathfrak{R}_{q,u,e}} \varepsilon^{k(T)} \mathcal{V}_T(x; \varepsilon, \beta_0), \tag{2.4c}$$

$$\mathcal{V}_T(x; \varepsilon, \beta_0) := \left( \prod_{v \in N(T)} \mathcal{F}_v(\beta_0) \right) \left( \prod_{\ell \in L(T)} \mathcal{G}_{e_\ell, u_\ell}^{[n_\ell]}(\omega \cdot \nu_\ell; \varepsilon, \beta_0) \right), \tag{2.4d}$$

where  $x = \omega \cdot \nu_{\ell_T}$  and  $\mathcal{V}_T(x; \varepsilon, \beta_0)$  is called the *renormalised value* of  $T$ . Here and henceforth, the sums and the products over empty sets have to be considered as 0 and 1, respectively. Set  $\mathcal{M} = \{\mathcal{M}^{[n]}(x; \varepsilon, \beta_0)\}_{n \geq -1}$ . We call *self-energies* the  $r \times r$  matrices  $\mathcal{M}^{[n]}(x; \varepsilon, \beta_0)$ .

Then we associate with each line  $\ell \in L(\theta)$  a *propagator*  $\mathcal{G}_\ell$  by setting

$$\mathcal{G}_\ell = \begin{cases} \mathcal{G}_{e_\ell, u_\ell}^{[n_\ell]}(\omega \cdot \nu_\ell; \varepsilon, \beta_0), & n_\ell \geq 0, \\ \delta_{e_\ell, u_\ell}, & n_\ell = -1. \end{cases}$$

where  $\delta_{e,u}$  is the Kronecker symbol. Recall that  $n_\ell = -1$  is possible only if  $\ell$  is the root line.

Note that, in defining the renormalised value in (2.4a) as  $\mathcal{V}_T(\omega \cdot \nu_{\ell_T}; \varepsilon, \beta_0)$ , we have used that, by construction, the propagators of the lines  $\ell \in \mathcal{P}_T$  depend on  $\omega \cdot \nu_{\ell_T}$ , while for the propagators of the lines  $\ell \in L(T) \setminus \mathcal{P}_T$  one has  $\nu_\ell = \nu_\ell^0$ , with the notation (2.2).

### 2.5. Resummed series

For any subgraph  $S$  of any  $\theta \in \Theta_{k,\nu,j}$  define the *renormalised value* of  $S$  as

$$\mathcal{V}(S; \varepsilon, \beta_0) := \left( \prod_{v \in N(S)} \mathcal{F}_v(\beta_0) \right) \left( \prod_{\ell \in L(S)} \mathcal{G}_\ell \right).$$

Set

$$b_{\nu_j}^{[k]}(\varepsilon, \beta_0) := \sum_{\theta \in \Theta_{k,\nu,j}} \mathcal{V}(\theta; \varepsilon, \beta_0), \quad \nu \neq 0, \tag{2.5a}$$

$$G_j^{[k]}(\varepsilon, \beta_0) := \sum_{\theta \in \Theta_{k+1,0,j}} \mathcal{V}(\theta; \varepsilon, \beta_0), \tag{2.5b}$$

and define formally the functions  $b(\alpha; \varepsilon, \beta_0)$  and  $G(\varepsilon, \beta_0)$  with components

$$b_j(\alpha; \varepsilon, \beta_0) := \sum_{k \geq 1} \varepsilon^k \sum_{\nu \in \mathbb{Z}_*^d} e^{i\nu \cdot \alpha} b_{\nu_j}^{[k]}(\varepsilon, \beta_0), \tag{2.6a}$$

$$G_j(\varepsilon, \beta_0) := \sum_{k \geq 0} \varepsilon^{k+1} G_j^{[k]}(\varepsilon, \beta_0). \tag{2.6b}$$

We call (2.6) the *resummed series*. Note that formally  $G(\varepsilon, \beta_0) = [\varepsilon \partial_\beta f(\cdot, \beta_0 + b(\cdot; \varepsilon, \beta_0))]_0$  and  $b = b(\alpha; \varepsilon, \beta_0)$  is the formal solution referred to in Sect. 1.2.

### 2.6. Siegel–Bryuno bounds

For  $\theta \in \Theta_{k,\nu,j}$ , let  $\mathfrak{N}_n(\theta)$  be the number of lines on scale  $\geq n$  in  $\theta$ , and set

$$K(\theta) := \sum_{v \in N(\theta)} |\nu_v|.$$

More generally, for any renormalised subgraph  $T$  of any tree  $\theta$  call  $\mathfrak{N}_n(T)$  the number of lines on scale  $\geq n$  in  $T$ , and set

$$K(T) := \sum_{v \in N(T)} |\nu_v|.$$

The forthcoming two results are identical to Lemmas 4.1 and 4.2 in [11], respectively, since they do not depend on  $r$ .

**Lemma 2.1.** *For any  $\theta \in \Theta_{k,\nu,j}$  such that  $\mathcal{V}(\theta; \varepsilon, \beta_0) \neq 0$  one has  $\mathfrak{N}_n(\theta) \leq 2^{-(m_n-2)}K(\theta)$  for all  $n \geq 0$ .*

The proof is as in [11], Appendix A.

**Lemma 2.2.** *For any  $T \in \mathfrak{R}_{n,u,e}$  such that  $\mathcal{V}_T(x; \varepsilon, \beta_0) \neq 0$ , one has  $K(T) \geq 2^{m_n-1}$  and  $\mathfrak{N}_p(T) \leq 2^{-(m_p-2)}K(T)$  for all  $0 \leq p \leq n$ .*

The proof is as in [11], Appendix B.

### 3. Convergence of the resummed series

We first show that, under the assumption that the propagators satisfy suitable bounds (that we call property 1), the convergence of the resummed series would follow by standard arguments of multiscale analysis. As we shall see, two fundamental ingredients of the proof will be a cancellation mechanism, which strongly relies on the symmetry in Hypothesis 2 (Lemma 3.5), and some remarkable identities which relate the self-energies to the averaged Lagrangian (Lemma 3.9).

Next, to remove the undesired assumption, we proceed as follows. We modify the propagators by replacing the self-energies  $\mathcal{M}$  with new matrices  $\overline{\mathcal{M}}^\varepsilon$ , which still satisfy the symmetry of Lemma 3.5: by construction, this will imply automatically that  $\overline{\mathcal{M}}^\varepsilon$  satisfies property 1. Then we check that on a suitable curve  $\beta_0 = \overline{\beta}_0(\varepsilon)$  one has  $\overline{\mathcal{M}}^\varepsilon = \mathcal{M}$  and hence the range equation is satisfied. Finally, by further exploiting the identities of Lemma 3.9, we show that, on such a curve, also the bifurcation equation is solved.

#### 3.1. Formal analysis

We first assume that the propagators  $\mathcal{G}_\ell$  are bounded proportionally to  $(\omega \cdot \nu_\ell)^{-2}$  and show that, under such an assumption, the convergence of the series is easily checked. Then, we shall prove that the assumption makes sense for a suitable choice of  $\beta_0$ . Let us denote by  $\|\cdot\|$  the  $r$ -dimensional  $L^2$  operator norm.

**Definition 3.1.** We shall say that  $\mathcal{M}$  satisfies *property 1* –  $p$  if for  $-1 \leq n < p$  one has

$$\Psi_{n+1}(x) \left\| x^2 \mathbf{1} - \mathcal{M}^{[n]}(x; \varepsilon, \beta_0) \right\| \geq \Psi_{n+1}(x) x^2 / 2.$$

Denote by  $\Theta_{k,\nu,j}^{\leq p}$  the set of renormalised trees whose lines are on scale  $\leq p$  and set

$$b_j^{\leq p}(\alpha; \varepsilon, \beta_0) := \sum_{k \geq 1} \varepsilon^k \sum_{\nu \in \mathbb{Z}_*^d} e^{i\nu \cdot \alpha} \sum_{\theta \in \Theta_{k,\nu,j}^{\leq p}} \mathcal{V}(\theta; \varepsilon, \beta_0). \tag{3.1}$$

Call  $b^{\leq p}(\alpha; \varepsilon, \beta_0)$  the function with components (3.1).

**Lemma 3.2.** *Assume  $\mathcal{M}$  to satisfy property  $1 - p$ . There are two positive constants  $B_0$  and  $B_1$  such that for all  $(k, \nu, j) \in \mathbb{N} \times \mathbb{Z}^d \times \{1, \dots, r\}$  and for any tree  $\theta \in \Theta_{k,\nu,j}^{\leq p}$  one has  $|\mathcal{V}(\theta; \varepsilon, \beta_0)| \leq B_0 B_1^k e^{-\xi_1 |\nu|}$ , with  $B_0, B_1$  and  $\xi_1 < \xi$ , independent of  $p$ .*

The proof is standard; see for instance [11], Appendix C. Note that the constants  $B_0$  and  $B_1$  increase with  $r$ .

Lemma 3.2 implies immediately the following result.

**Lemma 3.3.** *Assume  $\mathcal{M}$  to satisfy property  $1 - p$ . Then the series (3.1) converge for  $\varepsilon$  small enough. Moreover  $b^{\leq p}(\alpha; \varepsilon, \beta_0)$  is analytic in  $\alpha \in \mathbb{T}_{\xi_2}^d$ , with  $\xi_2 < \xi_1$ , uniformly in  $p$ .*

**Lemma 3.4.** *Assume  $\mathcal{M}$  to satisfy property  $1 - p$ . Then for any  $0 \leq n \leq p$  the self-energies are well defined and one has*

$$\left| \partial_x^j M_{e,u}^{[n]}(x; \varepsilon, \beta_0) \right| \leq |\varepsilon|^2 C_0 e^{-C_1 2^{mn}}, \quad j = 0, 1, 2,$$

for suitable constants  $C_0$  and  $C_1$ , independent of  $p$ .

The proof is as in [11], Appendix E.

By writing

$$\begin{aligned} \mathcal{M}^{[n]}(x; \varepsilon, \beta_0) &= \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) + x \partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) \\ &\quad + x^2 \int_0^1 d\tau (1 - \tau) \partial_x^2 \mathcal{M}^{[n]}(\tau x; \varepsilon, \beta_0), \end{aligned}$$

one checks easily that, if  $\mathcal{M}$  satisfies property  $1 - p$ ,  $\partial_x^j M_{e,u}^{[n]}(\tau x; \varepsilon, \beta_0)$  admits the same bounds as in Lemma 3.4, for  $0 \leq n \leq p$ ,  $j = 0, 1, 2$  and  $\tau \in [0, 1]$ . This implies that

$$\left\| \mathcal{M}^{[n]}(x; \varepsilon, \beta_0) - \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) - x \partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) \right\| \leq C_2 |\varepsilon|^2 x^2$$

for some positive constant  $C_2$  independent of  $p$ .

**Lemma 3.5.** *Assume  $\mathcal{M}$  to satisfy property  $1 - p$ . Then one has  $\partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) = 0$  for all  $-1 \leq n \leq p$ .*

*Proof.* One proves that  $\mathcal{M}^{[n]}(x; \varepsilon, \beta_0) = (\mathcal{M}^{[n]}(-x; \varepsilon, \beta_0))^T = (\mathcal{M}^{[n]}(x; \varepsilon, \beta_0))^\dagger$  by induction as in [9], §4; see also [14]. The key remark is that, if we replace the momentum  $\nu$  of the line  $\ell_T^+$  with  $-\nu$  and then reverse the orientation of the lines in  $\mathcal{P}_T \cup \{\ell_T, \ell_T^+\}$ , then the momenta of the lines  $\ell \in L(T) \setminus \mathcal{P}_T$  do not change, whereas the momenta of the lines  $\ell \in \mathcal{P}_T$  change sign. In particular one has  $\mathcal{M}^{[n]}(-x; \varepsilon, \beta_0) = (\mathcal{M}^{[n]}(x; \varepsilon, \beta_0))^*$ . Moreover Hypothesis 2

yields that the node factors (2.3) are real and this implies inductively that also the propagators are real, so that  $(\mathcal{M}^{[n]}(x; \varepsilon, \beta_0))^* = \mathcal{M}^{[n]}(x; \varepsilon, \beta_0)$ . Therefore  $\mathcal{M}^{[n]}(-x; \varepsilon, \beta_0) = \mathcal{M}^{[n]}(x; \varepsilon, \beta_0)$ .  $\square$

**Remark 3.6.** Lemma 3.5, which is crucial in order to obtain the final result, is the only point where the reversibility condition in Hypothesis 2 is used. For  $r = 1$  we do not need such an assumption, because in that case  $\mathcal{M}^{[n]}(x; \varepsilon, \beta_0)$  is number-valued (instead of matrix-valued) and hence the identity  $\mathcal{M}^{[n]}(x; \varepsilon, \beta_0) = (\mathcal{M}^{[n]}(-x; \varepsilon, \beta_0))^T$  is enough to get the analogous result. Of course the symmetry  $\mathcal{M}^{[n]}(-x; \varepsilon, \beta_0) = \mathcal{M}^{[n]}(x; \varepsilon, \beta_0)$ , for which the parity assumption seems necessary, is much stronger than the desired result  $\partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) = 0$ .

**Remark 3.7.** If one considers a formal expansion in  $\varepsilon$  of  $\partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0)$  one sees that the first non-trivial contribution appears at order  $\varepsilon^3$  and it is given by the sum of three contributions, namely

$$\begin{aligned} & 2 \sum_{\nu_1, \nu_2, \nu_3 \in \mathbb{Z}_*^d}^* \left( \frac{1}{(\omega \cdot \nu_3)^3 (\omega \cdot \nu_1)^2} - \frac{1}{(\omega \cdot \nu_3)^2 (\omega \cdot \nu_1)^3} \right) \partial_\beta^2 f_{\nu_1} \partial_\beta^2 f_{\nu_2} \partial_\beta^2 f_{\nu_3} \\ & - 2 \sum_{\nu_1, \nu_2, \nu_3 \in \mathbb{Z}_*^d}^* \frac{1}{(\omega \cdot \nu_3)^3 (\omega \cdot \nu_2)^2} \partial_\beta^3 f_{\nu_1} \partial_\beta f_{\nu_2} \partial_\beta^2 f_{\nu_3} \\ & + 2 \sum_{\nu_1, \nu_2, \nu_3 \in \mathbb{Z}_*^d}^* \frac{1}{(\omega \cdot \nu_3)^3 (\omega \cdot \nu_2)^2} \partial_\beta^2 f_{\nu_3} \partial_\beta^3 f_{\nu_1} \partial_\beta f_{\nu_2}, \end{aligned}$$

where  $\sum^*$  means that we restrict the sum to  $\nu_i$  on scale  $\leq n$  such that  $\nu_1 + \nu_2 + \nu_3 = 0$ . The parity condition of Hypothesis 2 implies  $\partial_\beta f_\nu = \partial_\beta f_{-\nu}$  and hence exchanging  $\nu_i \rightsquigarrow -\nu_i$  and summing the contributions, due to the change of sign in one denominator in each term we get zero. Note that  $\partial_\beta^2 f_\nu$  and  $\partial_\beta^3 f_\nu \partial_\beta f_\nu$  are  $r \times r$  matrices, and the cancellation of the  $\varepsilon^3$ -order of  $\partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0)$  apparently is to be related to the cancellations

$$[\partial_\beta^3 f_{\nu_1} \partial_\beta f_{\nu_2}, \partial_\beta^2 f_{\nu_3}] = 0, \quad \partial_\beta^2 f_{\nu_1} \partial_\beta^2 f_{\nu_2} \partial_\beta^2 f_{\nu_3} = \partial_\beta^2 f_{\nu_3} \partial_\beta^2 f_{\nu_2} \partial_\beta^2 f_{\nu_1}.$$

Of course one needs a “true cancellation” (i.e. not only up to the  $\varepsilon^3$ -order) but the  $\varepsilon^3$ -order gives a good insight: in particular when  $r = 1$  the parity condition is not needed because numbers obviously commute while this may not be true in the general case. In the non-degenerate case none of this matters because  $\mathcal{M}^{[n]}(0; \varepsilon, \beta_0)$  dominates  $\partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0)x$ ; see [18], formula (3.26). Thus it is possible that in some degenerate cases  $\partial_x \mathcal{M}^{[n]}(0; \varepsilon, \beta_0)$  simply does not vanishes and in turn the propagators cannot be controlled for all diophantine frequencies  $\omega$ . An issue of the same kind appears if one tries different approaches, such as KAM or Nash–Moser.

**Lemma 3.8.** *Assume  $\mathcal{M}$  to satisfy property 1 – p. Then, for all  $\beta_0 \in \mathbb{T}^r$ ,  $b^{\leq p}(\alpha; \varepsilon, \beta_0)$  solves the range equation (1.7a) up to corrections bounded by  $C_3 e^{-\xi' 2^{m_{p+1}}}$ , for some positive constants  $C_3$  and  $\xi'$  independent of  $p$ .*

*Proof.* We want to show that, by shortening  $b^{\leq p} = b^{\leq p}(\alpha; \varepsilon, \beta_0)$ ,

$$(\omega \cdot \nu)^{-2} [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu = [b^{\leq p}]_\nu + O(e^{-\xi' 2^{m_{p+1}}})$$

for some  $\xi' > 0$ . For any  $n \leq p$ , denote by  $\Theta_{k,\nu,j}^{\leq p}(n)$  the set of trees  $\theta \in \Theta_{k,\nu,j}^{\leq p}$  such that the root line is on scale  $n$  and set

$$b_\nu^n := \sum_{k \geq 1} \varepsilon^k \sum_{\theta \in \Theta_{k,\nu,j}^{\leq p}(n)} \mathcal{V}(\theta; \varepsilon, \beta_0).$$

We can write

$$\begin{aligned} [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu &= \sum_{n=0}^\infty \Psi_n(\omega \cdot \nu) [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu \\ &= \sum_{n=0}^p \Psi_n(\omega \cdot \nu) [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu + \sum_{n=p+1}^\infty \Psi_n(\omega \cdot \nu) [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu, \end{aligned} \tag{3.2}$$

with the last sum not vanishing only if  $\Psi_n(\omega \cdot \nu) \neq 0$  for some  $n \geq p + 1$ . In turn this requires  $|\nu| \geq 2^{m_{p+1}}$ , so that, by relying on Lemma 3.2, for such  $\nu$  we can bound

$$c_\nu := (\omega \cdot \nu)^{-2} \left| \sum_{n=p+1}^\infty \Psi_n(\omega \cdot \nu) [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu \right| \leq C e^{-\xi' |\nu|} e^{-\xi' 2^{m_{p+1}}}. \tag{3.3}$$

for some positive constants  $C$  and  $\xi'$ . The first sum in the second line of (3.2) becomes

$$\sum_{n=0}^p \left( (\omega \cdot \nu)^2 - \mathcal{M}^{[n-1]}(\omega \cdot \nu; \varepsilon, \beta_0) \right) \sum_{k \geq 1} \varepsilon^k \sum_{\theta \in \bar{\Theta}_{k,\nu,j}^{\leq p}(n)} \mathcal{V}(\theta; \varepsilon, \beta_0), \tag{3.4}$$

where  $\bar{\Theta}_{k,\nu,j}^{\leq p}(n)$  differs from  $\Theta_{k,\nu,j}^{\leq p}(n)$  as it contains also trees  $\theta$  which have one self-energy cluster with exiting line  $\ell_\theta$ . If we separate the trees containing such self-energy cluster from the others, (3.4) gives

$$\begin{aligned} &\sum_{n=0}^p \left( (\omega \cdot \nu)^2 - \mathcal{M}^{[n-1]}(\omega \cdot \nu; \varepsilon, \beta_0) \right) b_\nu^n + \sum_{m=0}^p \sum_{q=-1}^{m-1} M^{[q]}(\omega \cdot \nu; \varepsilon, \beta_0) \sum_{n=q+1}^p \Psi_n(\omega \cdot \nu) b_\nu^m \\ &= \sum_{n=0}^p \left( (\omega \cdot \nu)^2 - \mathcal{M}^{[n-1]}(\omega \cdot \nu; \varepsilon, \beta_0) \right) b_\nu^n + \sum_{n=0}^p \mathcal{M}^{[n-1]}(\omega \cdot \nu; \varepsilon, \beta_0) b_\nu^n, \end{aligned}$$

so that

$$(\omega \cdot \nu)^{-2} [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p})]_\nu = c_\nu + \sum_{n=0}^p b_\nu^n = c_\nu + [b^{\leq p}]_\nu$$

where  $c_\nu$  vanishes for  $|\nu| < 2^{m_{p+1}}$  and satisfies (3.3) for  $|\nu| \geq 2^{m_{p+1}}$ . □

Let us denote

$$G^{\leq p}(\varepsilon, \beta_0) := [\varepsilon \partial_\beta f(\alpha, \beta_0 + b^{\leq p}(\alpha; \varepsilon, \beta_0))]_0 \tag{3.5}$$

and

$$L^{\leq p}(\varepsilon, \beta_0) := \left[ -\frac{1}{2} (\omega \cdot \partial_\alpha b^{\leq p}(\alpha; \varepsilon, \beta_0))^2 + \varepsilon f(\alpha, \beta_0 + b^{\leq p}(\alpha; \varepsilon, \beta_0)) \right]_0, \tag{3.6}$$

which are the averages of the vector field and of the Lagrangian respectively, both computed along the approximate solution  $\beta_0 + b^{\leq p}(\alpha; \varepsilon, \beta_0)$ .

The variational structure of the Hamilton equations implies that, if we could solve the range equation, then the left hand side of the bifurcation equation (1.7b) would be the derivative with respect of  $\beta_0$  of the average of the Lagrangian computed along the solution. A variational identity of this kind holds, up to some correction, for the approximate solution; indeed we have the following result.

**Lemma 3.9.** *Assume  $\mathcal{M}$  to satisfy property 1 –  $p$ . Then (3.5) and (3.6) are well defined and  $C^\infty$  in both  $\varepsilon$  and  $\beta_0$ . Moreover one has for all  $-1 \leq n \leq p$*

$$\begin{aligned} \mathcal{M}^{[n]}(0; \varepsilon, \beta_0) &= \partial_{\beta_0} G^{\leq n}(\varepsilon, \beta_0) + c_n(\varepsilon, \beta_0), \\ G^{\leq n}(\varepsilon, \beta_0) &= \partial_{\beta_0} L^{\leq n}(\varepsilon, \beta_0) + c'_n(\varepsilon, \beta_0), \end{aligned}$$

for suitable  $c_n(\varepsilon, \beta_0), c'_n(\varepsilon, \beta_0) = O(e^{-\xi' 2^{mn+1}})$ , with  $\xi'$  as in Lemma 3.8.

*Proof.* The first identity can be proved as Lemma 4.8 in [12]. The second one is easier: integrating by parts and using Lemma 3.8 and the fact that  $[b^{\leq p}(\cdot; \varepsilon, \beta_0)]_0 = 0$ , we obtain, shortening again  $b^{\leq p} = b^{\leq p}(\alpha; \varepsilon, \beta_0)$ ,

$$\begin{aligned} \partial_{\beta_0} L^{\leq n} &= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} d\alpha \left( \varepsilon \partial_{\beta} f(\alpha, \beta_0 + b^{\leq n}) (1 + \partial_{\beta_0} b^{\leq n}) - (\omega \cdot \partial_{\alpha} b^{\leq n}) (\omega \cdot \partial_{\alpha} \partial_{\beta_0} b^{\leq n}) \right) \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} d\alpha \left( \varepsilon \partial_{\beta} f(\alpha, \beta_0 + b^{\leq n}) (1 + \partial_{\beta_0} b^{\leq n}) + ((\omega \cdot \partial_{\alpha})^2 b^{\leq n}) \partial_{\beta_0} b^{\leq n} \right) \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} d\alpha \left( \varepsilon f(\alpha, \beta_0 + b^{\leq n}) \right) + c''_n(\varepsilon, \beta_0), \end{aligned}$$

with a suitable  $c''_n(\varepsilon, \beta_0) = O(e^{-\xi' 2^{mn+1}})$ .  $\square$

**Definition 3.10.** We shall say that  $\mathcal{M}$  satisfies *property 1* if for all  $n \geq -1$  one has

$$\Psi_{n+1}(x) \left\| x^2 \mathbf{1} - \mathcal{M}^{[n]}(x; \varepsilon, \beta_0) \right\| \geq \Psi_{n+1}(x) x^2 / 2.$$

**Remark 3.11.** Assuming property 1 means assuming the bounds of property 1 –  $p$  for all  $p \geq 0$ . The reason why we have introduced first property 1 –  $p$ , and not directly property 1, is that property 1 –  $p$  is what is needed to prove by induction that property 1 holds. Indeed, in Sect. 3.2 we shall introduce new matrices  $\overline{\mathcal{M}}^\xi$  and we shall prove that they satisfy property 1. This will be achieved by induction, by assuming that they satisfy property 1 –  $p$  and hence showing that property 1 –  $(p+1)$  follows as well: to this aim we shall need the lemmas above which require property 1 –  $p$  to be satisfied.

If property 1 holds, the approximate solution  $b^{\leq p}(\alpha; \varepsilon, \beta_0)$  is a “scale  $p$  truncation” of the solution to the range equation. Indeed we have the following result.

**Lemma 3.12.** *Assume  $\mathcal{M}$  to satisfy property 1. Then for  $\varepsilon$  small enough the function  $b(\alpha; \varepsilon, \beta_0)$  in (2.6a), with the coefficients given in (2.5), is analytic in  $\alpha \in \mathbb{T}_{\xi_2}^d$ , with  $\xi_2$  as in Lemma 3.3, and solves the Eq. (1.7a).*

The proof follows from Lemmas 3.3 and 3.8, by taking the limit  $p \rightarrow \infty$ . Define formally

$$\mathcal{M}^{[\infty]}(x; \varepsilon, \beta_0) := \lim_{n \rightarrow \infty} \mathcal{M}^{[n]}(x; \varepsilon, \beta_0), \quad L^{[\infty]}(\varepsilon, \beta_0) := \lim_{n \rightarrow \infty} L^{\leq n}(\varepsilon, \beta_0). \tag{3.7}$$

Lemma 3.4 yields that, under the assumption that  $\mathcal{M}$  satisfy property 1, the functions (3.7) are well defined and smooth. Indeed the following result holds.

**Lemma 3.13.** *Assume  $\mathcal{M}$  to satisfy property 1. Then the function  $G(\varepsilon, \beta_0)$  and the self-energies  $\mathcal{M}$ , as well as the limits (3.7), are  $C^\infty$  in both  $\varepsilon$  and  $\beta_0$ .*

**Lemma 3.14.** *Assume  $\mathcal{M}$  to satisfy property 1. Then the implicit function equation  $G(\varepsilon, \beta_0) = 0$  admits a solution  $\beta_0 = \beta_0(\varepsilon)$  such that one has  $\partial_{\beta_0} G(\varepsilon, \beta_0(\varepsilon)) \leq 0$ .*

*Proof.* Property 1 ensures that we can apply Lemma 3.9 with  $n \rightarrow \infty$ , so as to obtain  $\partial_{\beta_0} G(\varepsilon, \beta_0) = \partial_{\beta_0}^2 L^{[\infty]}(\varepsilon, \beta_0)$ . Since  $L^{[\infty]}(\varepsilon, \beta_0)$  is a smooth function on the torus  $\mathbb{T}^r$ , it admits at least one maximum.  $\square$

Lemmas 3.12 and 3.14 give the following result.

**Lemma 3.15.** *Assume  $\mathcal{M}$  to satisfy property 1 and let  $\beta(\varepsilon)$  be as in Lemma 3.14. Then the function  $\beta(t) = \beta_0(\varepsilon) + b(\alpha; \varepsilon, \beta_0(\varepsilon))$  solves (1.5).*

### 3.2. Rigorous analysis: phase locking

We introduce an auxiliary function  $\bar{b}(\alpha; \varepsilon, \beta_0)$  obtained from the tree expansion by replacing the propagators (2.4) in such a way that the new propagators  $\bar{\mathcal{G}}_\ell$  are bounded proportionally to  $(\omega \cdot \nu_\ell)^{-2}$ .

For all  $n \geq 0$ , define the  $C^\infty$  non-increasing functions  $\xi_n$  such that

$$\xi_n(x_1, \dots, x_r) = \begin{cases} 0, & \exists i \in \{1, \dots, r\} \text{ s.t. } x_i \geq \frac{\alpha_{m_{n+1}}^2(\omega)}{2^{11}}, \\ 1, & x_i \leq \frac{\alpha_{m_{n+1}}^2(\omega)}{2^{12}} \quad \forall i = 1, \dots, r, \end{cases}$$

and set  $\xi_{-1}(x_1, \dots, x_r) = 1$ .

Define, for  $n \geq 0$ , the modified propagators as  $\bar{\mathcal{G}}_\ell = \bar{\mathcal{G}}_{e_\ell, u_\ell}^{[n_\ell]}(\omega \cdot \nu_\ell; \varepsilon, \beta_0)$ , with

$$\bar{\mathcal{G}}^{[n]}(x; \varepsilon, \beta_0) = \Psi_n(x) \left( x^2 \mathbf{1} - \bar{\mathcal{M}}^{[n-1]}(x; \varepsilon, \beta_0) \xi_{n-1}(\bar{\lambda}^{[n-1]}(\varepsilon, \beta_0)) \right)^{-1},$$

where  $\bar{\mathcal{M}}^{[-1]}(x; \varepsilon, \beta_0) = \partial_{\beta_0}^2 \bar{L}^{[-1]}(\varepsilon, \beta_0)$ , with  $\bar{L}^{[-1]}(\varepsilon, \beta_0) = \varepsilon f_0(\beta_0)$ , while for  $n \geq 0$  one proceeds recursively as follows:

1. set

$$\bar{b}_j^{[n]}(\alpha; \varepsilon, \beta_0) := \sum_{k \geq 1} \varepsilon^k \sum_{\nu \in \mathbb{Z}_*^d} e^{i\nu \cdot \alpha} \sum_{\theta \in \Theta_{k, \nu, j}^{\leq n}} \bar{\Psi}(\theta; \varepsilon, \beta_0), \tag{3.8}$$



where

$$\overline{\mathcal{V}}(\theta; \varepsilon, \beta_0) := \left( \prod_{v \in N(\theta)} \mathcal{F}_v(\beta_0) \right) \left( \prod_{\ell \in L(\theta)} \overline{\mathcal{G}}_\ell \right),$$

and call  $\overline{b}^{[n]}(\alpha; \varepsilon, \beta_0)$  the function with components (3.8);

2. define

$$\overline{L}^{[n]}(\varepsilon, \beta_0) := \left[ -\frac{1}{2}(\omega \cdot \partial_\alpha \overline{b}^{[n]}(\alpha; \varepsilon, \beta_0))^2 + \varepsilon f(\alpha, \beta_0 + \overline{b}^{[n]}(\alpha; \varepsilon, \beta_0)) \right]_0;$$

3. consider any matrices  $\overline{c}_n(\varepsilon, \beta_0)$  and  $\overline{\mathfrak{R}}^{[n]}(x; \varepsilon, \beta_0)$ , satisfying the constraints

$$\|\overline{c}_n(\varepsilon, \beta_0)\| \leq K|\varepsilon|^2 e^{-\xi' 2^{m_{n+1}}}, \tag{3.9a}$$

$$\left\| \overline{\mathfrak{R}}^{[n]}(x; \varepsilon, \beta_0) - \overline{\mathfrak{R}}^{[n-1]}(x; \varepsilon, \beta_0) \right\| \leq K|\varepsilon|^2 x^2 e^{-\xi' 2^{m_n}}, \tag{3.9b}$$

for some positive constant  $K$ , with  $\overline{\mathfrak{R}}^{[-1]}(x; \varepsilon, \beta_0) = 0$  and with  $\xi'$  as in Lemma 3.8;

4. call  $\overline{\lambda}^{[n]}(\varepsilon, \beta_0) = (\overline{\lambda}_1^{[n]}(\varepsilon, \beta_0), \dots, \overline{\lambda}_r^{[n]}(\varepsilon, \beta_0))$  the eigenvalues of  $\partial_{\beta_0}^2 \overline{L}^{[n]}(\varepsilon, \beta_0)$ ;

5. define

$$\overline{\mathcal{M}}^{[n]}(x; \varepsilon, \beta_0) = \partial_{\beta_0}^2 \overline{L}^{[n]}(\varepsilon, \beta_0) + \overline{c}_n(\varepsilon, \beta_0) + \overline{\mathfrak{R}}^{[n]}(x; \varepsilon, \beta_0). \tag{3.10}$$

We write  $\overline{\mathcal{M}}^\xi = \{\overline{\mathcal{M}}^{[n]}(x; \varepsilon, \beta_0) \xi_n(\overline{\lambda}^{[n]}(\varepsilon, \beta_0))\}_{n \geq -1}$ . The following result can be proved by reasoning as in the proof of Lemma 3.4.

**Lemma 3.16.** *Assume  $\overline{\mathcal{M}}^\xi$  to satisfy property 1 – p. Then for any  $0 \leq n \leq p$  one has*

$$\max_{i_1 + \dots + i_r \leq 3} \left| \partial_{\beta_{01}}^{i_1} \dots \partial_{\beta_{0r}}^{i_r} \left( \overline{L}^{[n]}(\varepsilon, \beta_0) - \overline{L}^{[n-1]}(\varepsilon, \beta_0) \right) \right| \leq |\varepsilon|^2 K_0 e^{-K_1 2^{m_n}},$$

for suitable  $p$ -independent constants  $K_0$  and  $K_1$ .

Lemma 3.16 implies that, under the same assumptions, one has

$$\left\| \overline{\mathcal{M}}^{[n]}(x; \varepsilon, \beta_0) - \overline{\mathcal{M}}^{[n-1]}(x; \varepsilon, \beta_0) \right\| \leq |\varepsilon|^2 K_2 e^{-K_1 2^{m_n}},$$

for a suitable  $p$ -independent constant  $K_2$ .

**Lemma 3.17.**  *$\overline{\mathcal{M}}^\xi$  satisfies property 1.*

*Proof.* We shall prove that  $\overline{\mathcal{M}}^\xi$  satisfies property 1 – p for all  $p \geq 0$ , by induction on  $p$ . Property 1-0 is trivially satisfied for  $\varepsilon$  small enough. Assume  $\overline{\mathcal{M}}^\xi$  to satisfies property 1 – p. Then (3.9) implies that, for all  $n \leq p$  and  $x$  such that  $\Psi_n(x) \neq 0$ , one has

$$\begin{aligned} & \left\| x^2 \mathbf{1} - \partial_{\beta_0}^2 \overline{L}^{[n-1]}(\varepsilon, \beta_0) \xi_{n-1}(\overline{\lambda}^{[n-1]}(\varepsilon, \beta_0)) - \overline{c}_{n-1}(\varepsilon, \beta_0) \xi_{n-1}(\overline{\lambda}^{[n-1]}(\varepsilon, \beta_0)) \right. \\ & \quad \left. - \overline{\mathfrak{R}}^{[n-1]}(x, \varepsilon, \beta_0) \xi_{n-1}(\overline{\lambda}^{[n-1]}(\varepsilon, \beta_0)) \right\| \end{aligned}$$

$$\begin{aligned} &\geq \left\| x^2 \mathbf{1} - \partial_{\beta_0}^2 \bar{L}^{[n-1]}(\varepsilon, \beta_0) \xi_{n-1}(\bar{\lambda}^{[n-1]}(\varepsilon, \beta_0)) \right\| - \|\bar{c}_{n-1}(\varepsilon, \beta_0)\| - \left\| \bar{\mathfrak{R}}^{[n-1]}(x, \varepsilon, \beta_0) \right\| \\ &\geq x^2 - \max_{1 \leq i \leq r} \bar{\lambda}_i^{[n-1]}(\varepsilon, \beta_0) \xi_{n-1}(\bar{\lambda}^{[n-1]}(\varepsilon, \beta_0)) + O(e^{-\varepsilon' 2^{m_{n+1}}}) + O(\varepsilon^2 x^2) \geq \frac{x^2}{2} \end{aligned}$$

and hence  $\bar{\mathcal{M}}^\varepsilon$  satisfies property 1-( $p + 1$ ). □

Define

$$\bar{b}(\alpha; \varepsilon, \beta_0) := \lim_{n \rightarrow \infty} \bar{b}^{[n]}(\alpha; \varepsilon, \beta_0), \tag{3.11a}$$

$$\bar{L}^{[\infty]}(\varepsilon, \beta_0) := \lim_{n \rightarrow \infty} \bar{L}^{[n]}(\varepsilon, \beta_0). \tag{3.11b}$$

By Lemma 3.17, the limits in (3.11) are well defined functions, which are  $C^\infty$  in both  $\varepsilon$  and  $\beta_0$  and  $2\pi$ -periodic in each component of  $\beta_0$ . Moreover  $\bar{b}(\alpha; \varepsilon, \beta_0)$  is analytic in  $\alpha \in \mathbb{T}_{\xi''}^d$ , for some  $\xi'' < \xi'$ .

Set  $\bar{G}(\varepsilon, \beta_0) := \partial_{\beta_0} \bar{L}^{[\infty]}(\varepsilon, \beta_0)$ . The following result is the analogos of Lemma 3.14

**Lemma 3.18.** *For  $\varepsilon$  small enough there is at least one value  $\bar{\beta}_0(\varepsilon) \in \mathbb{T}^r$  such that  $\bar{G}(\varepsilon, \bar{\beta}_0(\varepsilon)) = 0$  and  $\partial_{\beta_0}^2 \bar{L}^{[\infty]}(\varepsilon, \bar{\beta}_0(\varepsilon)) \leq 0$ .*

**Lemma 3.19.** *Let  $\bar{\beta}_0(\varepsilon)$  be as in Lemma 3.18. Then for  $\varepsilon$  small enough and all  $n \geq 0$  one has  $\xi_n(\lambda^{[n]}(\varepsilon, \beta_0)) \equiv 1$  for all  $\beta_0$  such that  $|\beta_0 - \bar{\beta}_0(\varepsilon)| \leq \alpha_{m_{n+1}}^2(\omega)$ .*

*Proof.* Fix  $\beta_0 = \bar{\beta}_0(\varepsilon)$  so that  $\partial_{\beta_0}^2 \bar{L}^{[\infty]}(\varepsilon, \bar{\beta}_0(\varepsilon)) \leq 0$  by Lemma 3.18. One has

$$\partial_{\beta_0}^2 \bar{L}^{[n]}(\varepsilon, \beta_0) = \left( \partial_{\beta_0}^2 \bar{L}^{[n]}(\varepsilon, \beta_0) - \partial_{\beta_0}^2 \bar{L}^{[\infty]}(\varepsilon, \beta_0) \right) + \partial_{\beta_0}^2 \bar{L}^{[\infty]}(\varepsilon, \beta_0),$$

where, by Lemma 3.16,

$$\left\| \partial_{\beta_0}^2 \bar{L}^{[n]}(\varepsilon, \beta_0) - \partial_{\beta_0}^2 \bar{L}^{[\infty]}(\varepsilon, \beta_0) \right\| \leq K_3 |\varepsilon| e^{-K_1 2^{m_{n+1}}} \leq \frac{\alpha_{m_{n+1}}^2(\omega)}{2^{14}},$$

for a suitable positive constant  $K_3$ . Moreover

$$\left\| \bar{L}^{[n]}(\varepsilon, \beta_0) - \bar{L}^{[n]}(\varepsilon, \beta_0(\varepsilon)) \right\| \leq K_4 |\varepsilon|^2 |\beta_0 - \bar{\beta}_0(\varepsilon)|,$$

for a suitable positive constant  $K_4$ . Therefore, for  $|\beta_0 - \bar{\beta}_0(\varepsilon)| \leq \alpha_{m_{n+1}}^2(\omega)$ , one has

$$\bar{\lambda}_j^{[n]}(\varepsilon, \beta_0) \leq \bar{\lambda}_j^{[\infty]}(\varepsilon, \bar{\beta}_0(\varepsilon)) + \frac{\alpha_{m_{n+1}}^2(\omega)}{2^{13}} \leq \frac{\alpha_{m_{n+1}}^2(\omega)}{2^{13}}, \quad j = 1, \dots, r,$$

so that the assertion follows. □

**Lemma 3.20.** *One can choose the matrices  $\bar{c}_n(\varepsilon, \beta_0)$  and  $\bar{\mathfrak{R}}^{[n]}(x; \varepsilon, \beta_0)$  in (3.10) so that the function  $\beta(t) = \bar{\beta}_0(\varepsilon) + \bar{b}(\alpha; \varepsilon, \bar{\beta}_0(\varepsilon))$  solves (1.5).*

*Proof.* We want to prove by induction that we can choose the matrices  $\bar{c}_n(\varepsilon, \beta_0)$  and  $\bar{\mathfrak{R}}^{[n]}(x; \varepsilon, \beta_0)$ , for  $n \geq 0$ , so as to have

$$\bar{\mathcal{M}}^{[n]}(x; \varepsilon, \bar{\beta}_0(\varepsilon)) = \mathcal{M}^{[n]}(x; \varepsilon, \bar{\beta}_0(\varepsilon)) \tag{3.12}$$

for all  $n \geq -1$ . This will imply that, when setting  $\beta_0 = \bar{\beta}_0(\varepsilon)$ ,  $\mathcal{M}$  satisfies property 1 as well as  $\overline{\mathcal{M}}$ . Hence we obtain  $b(\alpha; \varepsilon, \bar{\beta}_0(\varepsilon)) = \bar{b}(\alpha; \varepsilon, \bar{\beta}_0(\varepsilon))$  and we can apply Lemma 3.15 with  $\beta_0(\varepsilon) = \bar{\beta}_0(\varepsilon)$  to deduce that  $\bar{\beta}_0(\varepsilon) + \bar{b}(\alpha; \varepsilon, \bar{\beta}_0(\varepsilon))$  solves the range equation (1.7a).

For  $n = -1$  the identity (3.12) trivially holds. Assume that (3.12) is satisfied up to  $n < p$  for a suitable choice of the matrices  $\bar{c}_n(\varepsilon, \beta_0)$  and  $\overline{\mathfrak{R}}^{[n]}(x; \varepsilon, \beta_0)$ , with  $0 \leq n < p$ . Then, for  $|\beta_0 - \bar{\beta}_0(\varepsilon)| \leq \alpha_{m_{n+1}}^2(\omega)$ ,  $\mathcal{M}$  satisfies property 1 –  $p$  and hence  $\mathcal{M}^{[p]}(x; \varepsilon, \beta_0)$  is well defined. Moreover for such  $\beta_0$  one has  $\bar{L}^{[p]}(\varepsilon, \beta_0) = L^{\leq p}(\varepsilon, \beta_0)$  and hence

$$\partial_{\beta_0}^2 \bar{L}^{[p]}(\varepsilon, \beta_0) = \mathcal{M}^{[p]}(0; \varepsilon, \beta_0) - d_p(\varepsilon, \beta_0),$$

with the matrix  $d_p(\varepsilon, \beta_0) = \partial_{\beta_0} c'_p(\varepsilon, \beta_0) + c_p(\varepsilon, \beta_0)$  such that  $\|d_p(\varepsilon, \beta_0)\| \leq K_5 |\varepsilon|^2 e^{-\xi' 2^{m_{p+1}}}$ , for a suitable positive constant  $K_5$ . Then (3.10) gives

$$\overline{\mathcal{M}}^{[p]}(x; \varepsilon, \beta_0) = \mathcal{M}^{[p]}(0; \varepsilon, \beta_0) - d_p(\varepsilon, \beta_0) + \bar{c}_p(\varepsilon, \beta_0) + \overline{\mathfrak{R}}^{[p]}(x; \varepsilon, \beta_0),$$

and one can take the matrices  $\bar{c}_p(\varepsilon, \beta_0)$  and  $\overline{\mathfrak{R}}^{[p]}(x; \varepsilon, \beta_0)$  so that

$$\bar{c}_p(\varepsilon, \beta_0) = d_p(\varepsilon, \beta_0), \quad \overline{\mathfrak{R}}^{[p]}(x; \varepsilon, \beta_0) = \mathcal{M}^{[p]}(x; \varepsilon, \beta_0) - \mathcal{M}^{[p]}(0; \varepsilon, \beta_0),$$

for  $|\beta_0 - \bar{\beta}_0(\varepsilon)| \leq \alpha_{m_{n+1}}^2(\omega)$ . Thanks to Lemma 3.5, this yields that both inequalities (3.9) are satisfied. Therefore (3.12) follows for  $n = p$ .

On the other hand one has

$$G(\varepsilon, \bar{\beta}_0(\varepsilon)) = \lim_{n \rightarrow \infty} \partial_{\beta_0} L^{\leq n}(\varepsilon, \bar{\beta}_0(\varepsilon)) = \partial_{\beta_0} \bar{L}^{[\infty]}(\varepsilon, \bar{\beta}_0(\varepsilon)) = \bar{G}(\varepsilon, \bar{\beta}_0(\varepsilon)) = 0$$

and hence the bifurcation equation (1.7b) is solved as well. □

## References

- [1] Ambrosetti, A., Zelati, V.C., Ekeland, I.: Symmetry breaking in Hamiltonian systems. *J. Differ. Equ.* **67**(2), 165–184 (1987)
- [2] Bernstein, D., Katok, A.: Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians. *Invent. Math.* **88**(2), 225–241 (1987)
- [3] Berti, M.: Nonlinear oscillations of Hamiltonian PDEs, *Progress in Nonlinear Differential Equations and their Applications*, 74, vol. 74. Birkhuser Boston Inc., Boston (2007)
- [4] Bricmont, J., Gawędzki, K., Kupiainen, A.: KAM theorem and quantum field theory. *Commun. Math. Phys.* **201**(3), 699–727 (1999)
- [5] Cheng, Ch-Q: Birkhoff–Kolmogorov–Arnold–Moser tori in convex Hamiltonian systems. *Commun. Math. Phys.* **177**(3), 529–559 (1996)

- [6] Cheng, Ch-Q: Lower-dimensional invariant tori in the regions of instability for nearly integrable Hamiltonian systems. *Commun. Math. Phys.* **203**(2), 385–419 (1999)
- [7] Cheng, Ch-Q, Wang, Sh: The surviving of lower-dimensional tori from a resonant torus of Hamiltonian systems. *J. Differ. Equ.* **155**(2), 311–326 (1999)
- [8] Cong, F., Küpper, T., Li, Y., You, J.: KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems. *J. Nonlinear Sci.* **10**(1), 49–68 (2000)
- [9] Corsi, L., Feola, R., Gentile, G.: Lower-dimensional invariant tori for perturbations of a class of non-convex Hamiltonian functions. *J. Stat. Phys.* **150**(1), 156–180 (2013)
- [10] Corsi, L., Gentile, G.: Melnikov theory to all orders and Puiseux series for subharmonic solutions. *J. Math. Phys.* **49**(11), 112701 (2008)
- [11] Corsi, L., Gentile, G.: Oscillator synchronisation under arbitrary quasi-periodic forcing. *Commun. Math. Phys.* **316**(2), 489–529 (2012)
- [12] Corsi, L., Gentile, G.: Resonant motions in the presence of degeneracies for quasi-periodically perturbed systems. *Ergod. Theory Dynam. Syst.* **35**(4), 1079–1140 (2015)
- [13] Feola, R.: Resonant tori and one-dimensional systems with quasi-periodic forcing, Graduation thesis in Mathematics, Rome (2012)
- [14] Gallavotti, G., Gentile, G.: Degenerate elliptic resonances. *Commun. Math. Phys.* **257**(2), 319–362 (2005)
- [15] Gallavotti, G., Gentile, G., Giuliani, A.: Fractional Lindstedt series. *J. Math. Phys.* **47**(1), 012702 (2006)
- [16] Gentile, G.: Degenerate lower-dimensional tori under the Bryuno condition. *Ergod. Theory Dyn. Syst.* **27**(2), 427–457 (2007)
- [17] Gentile, G.: Quasiperiodic motions in dynamical systems: review of a renormalization group approach. *J. Math. Phys.* **51**(1), 015207 (2010)
- [18] Gallavotti, G., Gentile, G.: Hyperbolic low-dimensional invariant tori and summations of divergent series. *Commun. Math. Phys.* **227**(3), 421–460 (2002)
- [19] Jorba, À., de la Llave, R., Zou, M.: Lindstedt series for lower-dimensional tori. In: *Hamiltonian Systems with Three or More Degrees of Freedom* (S’Agaró, 1995), vol. 533, pp. 151–167, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Kluwer Academic Publishers, Dordrecht (1999)
- [20] Kappeler, Th., Pöschel, J.: KdV & KAM. In: *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*, vol. 45. Springer, Berlin (2003)
- [21] Plotnikov, P.I., Kuznetsov, I.V., Kolmogorov’s theorem for low-dimension invariant tori of Hamiltonian systems. *Dokl. Akad. Nauk* **439**(3), 311–314 (2011); English translation in *Dokl. Math.* **84** (2011), no. 1, 498–501

- [22] Plotnikov, P.I., Kuznetsov, I.V.: Kolmogorov's theorem for low-dimension invariant tori of Hamiltonian systems. preprint (2015). [arXiv:1506.00383](https://arxiv.org/abs/1506.00383)
- [23] Rabinowitz, P.H.: Minimax methods in critical point theory with applications to differential equations. In: CBMS Regional Conference Series in Mathematics, vol. 65. Providence, RI (1986)
- [24] Treshchëv, D.V.: A mechanism for the destruction of resonance tori in Hamiltonian systems. *Mat. Sb.* **180**(10), 1325–1346, 1439 (1989); English translation in *Math. USSR-Sb.* **68** (1991), no. 1, 181–203
- [25] You, J.: A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems. *Commun. Math. Phys.* **192**(1), 145–168 (1998)

Livia Corsi  
School of Mathematics  
Georgia Institute of Technology  
686 Cherry St. NW  
Atlanta GA 30332  
USA  
e-mail: [lcorsi6@math.gatech.edu](mailto:lcorsi6@math.gatech.edu)

Guido Gentile  
Dipartimento di Matematica e Fisica  
Università Roma Tre  
Rome 00146  
Italy  
e-mail: [gentile@mat.uniroma3.it](mailto:gentile@mat.uniroma3.it)

Received: 31 March 2016.

Accepted: 24 November 2016.