INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **19** (2006) 2727–2753 doi:10.1088/0951-7715/19/12/001

Renormalization of multidimensional Hamiltonian flows

Kostya Khanin[1], **João Lopes Dias**[2] **and Jens Marklof**[3]

[1] Department of Mathematics, University of Toronto, 100 St George Street, Toronto, Ontario M5S 3G3, Canada

[2] Departamento de Matemática, ISEG, Universidade Técnica de Lisboa, Rua do Quelhas 6, 1200-781 Lisboa, Portugal

[3] School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

E-mail: khanin@math.toronto.edu, jldias@iseg.utl.pt and j.marklof@bristol.ac.uk

Received 8 February 2006, in final form 3 October 2006
Published 27 October 2006
Online at stacks.iop.org/Non/19/2727

Recommended by R de la Llave

Abstract

We construct a renormalization operator acting on the space of analytic Hamiltonians defined on $T^*\mathbb{T}^d$, $d \geqslant 2$, based on the multidimensional continued fractions algorithm developed by the authors. We show convergence of orbits of the operator around integrable Hamiltonians satisfying a non-degeneracy condition. This in turn yields a new proof of a KAM-type theorem on the stability of diophantine invariant tori.

Mathematics Subject Classification: 37E20, 37J40, 11J70

1. Introduction

The connection between KAM and renormalization theories has been realized for quite some time. The renormalization approach to KAM has several important advantages. First of all, it provides a unified setting which allows us to deal with both the cases of smooth KAM-type invariant tori and non-smooth critical tori. Secondly, the proofs based on renormalizations are conceptually very simple and give a different perspective on the problem of small divisors. For the continuous-time situation, several KAM results for small-divisor problems in quasiperiodic motion have been obtained by studying the stability of trivial fixed sets of renormalization operators (cf e.g. [3, 7, 10, 12, 13]). There was however a relevant restriction when dealing with multiple frequencies. Because renormalization methods rely fundamentally on the continued fractions expansion of the frequency vector, the lack of a multidimensional version of continued fractions was the reason for failing to replicate KAM in its full generality. This limitation was recently overcome in [6] by adapting Lagarias' algorithm [11] and deriving estimates for multidimensional continued fractions (MCFs) expansions of diophantine vectors.

We present here a further application of the multidimensional renormalization method following [6] (for vector fields on the torus) and [9] (for skew-product flows over translations on the torus), illustrating once again the connection between KAM and renormalization methods tackling quasiperiodic motion problems. Moreover, we hope that our work could lead to a better understanding of the behaviour of renormalization around critical fixed points. The only rigorous result in this direction is a computer-assisted proof of the existence of such critical fixed point in the golden-mean $d = 2$ case [8].

Our present renormalization scheme is similar in spirit to Koch's [7]. One of the differences is that the (analytic) Hamiltonians considered in [7] are close to the integrable (degenerate) Hamiltonian $\mathbb{R}^d \ni \boldsymbol{y} \mapsto \boldsymbol{\omega}\cdot\boldsymbol{y}$. So, due to the degeneracy condition there are unstable directions for the trivial fixed point of renormalization, and thus the KAM domain will correspond to the stable manifold. In our approach we deal with an extra quadratic term in the integrable case which implies convergence under renormalization on a ball. Moreover, the frequency vector $\boldsymbol{\omega} \in \mathbb{R}^d$ in [7] is assumed to be of a special kind (known as Koch type, cf [12]) corresponding to a zero Lebesgue measure set. In our work the result on the stability of invariant tori is valid for any diophantine vector, a full measure set. It is still a fundamental open problem to determine the largest set of frequencies for which the stability of KAM tori holds. We also expect that our methods can be adapted in order to deal with Hamiltonians of class C^k.

Let $B \subset \mathbb{R}^d$, $d \geqslant 2$, be an open set containing the origin, and let H^0 be a real-analytic Hamiltonian function

$$H^0(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{\omega}\cdot\boldsymbol{y} + \tfrac{1}{2}\,{}^{\top}\boldsymbol{y}Q\boldsymbol{y}, \qquad (\boldsymbol{x},\boldsymbol{y}) \in \mathbb{T}^d \times B, \tag{1.1}$$

with $\boldsymbol{\omega} \in \mathbb{R}^d$ and a real symmetric $d \times d$ matrix Q. H^0 is said to be non-degenerate if $\det Q \neq 0$. We say $\boldsymbol{\omega} \in \mathbb{R}^d$ is *diophantine* if there are constants $\beta > 0$ and $C > 0$ such that

$$\|\boldsymbol{k}\|^{d-1+\beta}|\boldsymbol{k}\cdot\boldsymbol{\omega}| > C, \qquad \boldsymbol{k} \in \mathbb{Z}^d - \{\boldsymbol{0}\}. \tag{1.2}$$

In this paper we prove the following theorem.

Theorem 1.1. *Suppose H^0 is non-degenerate and $\boldsymbol{\omega}$ is diophantine. If H is a real analytic Hamiltonian on $\mathbb{T}^d \times B$ sufficiently close to H^0, then the Hamiltonian flow of H leaves invariant a Lagrangian d-dim torus where it is analytically conjugated to the linear flow $\phi_t(\boldsymbol{x}) = \boldsymbol{x} + t\boldsymbol{\omega}$ on $\mathbb{T}^d$, $t \geqslant 0$. The conjugacy depends analytically on H.*

Sketch of the proof

Our proof of theorem 1.1 is related to the one in [6] done in the context of vector fields on $\mathbb{T}^d$. Hamiltonian vector fields involve more complicated analysis since there is extra dynamics on a vertical direction (action) and we need to preserve the symplectic nature of the problem. Our goal is to find an analytic embedding $\mathbb{T}^d \to \mathbb{T}^d \times B$ that conjugates the Hamiltonian flow to the linear flow on the torus given by $\boldsymbol{\omega}$.

We do not work directly with vector fields, instead we renormalize Hamiltonian functions $H(\boldsymbol{x},\boldsymbol{y}) = H^0(\boldsymbol{x},\boldsymbol{y}) + F(\boldsymbol{x},\boldsymbol{y})$ where $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{T}^d \times B$ and F is a sufficiently small analytic perturbation. Using a rescaling of time we may assume that $\boldsymbol{\omega} = \begin{pmatrix}\boldsymbol{\alpha}\\ 1\end{pmatrix}$ for some diophantine $\boldsymbol{\alpha} \in \mathbb{R}^{d-1}$. The perturbation F is decomposed in a Taylor–Fourier series $F(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k},\boldsymbol{\nu}} F_{\boldsymbol{k},\boldsymbol{\nu}} y_1^{\nu_1}\cdots y_d^{\nu_d} \mathrm{e}^{2\pi \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$ where the sum is taken over $\boldsymbol{k} \in \mathbb{Z}^d$ and $\nu_i \in \mathbb{N} \cup \{0\}$. By the analyticity of F, its modes decay exponentially as $\|\boldsymbol{k}\| \to +\infty$ for fixed $\boldsymbol{\nu}$.

Renormalization is an iterative scheme that at each step produces a new Hamiltonian. Suppose that after the $(n-1)$th step the Hamiltonian is of the form

$$H_{n-1}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{\omega}^{(n-1)}\cdot\boldsymbol{y} + \tfrac{1}{2}\,{}^{\top}\boldsymbol{y}Q_{n-1}\boldsymbol{y} + F_{n-1}(\boldsymbol{x},\boldsymbol{y}), \tag{1.3}$$

where $\boldsymbol{\omega}^{(n-1)} = \binom{\alpha^{(n-1)}}{1}$, $\boldsymbol{\alpha}^{(n-1)}$ is given by the continued fraction algorithm (see section 2) and Q_{n-1} is a symmetric matrix with non-zero determinant. Moreover, we assume that F_{n-1} only contains Taylor–Fourier resonant modes (said to be in I_{n-1}^{+}), i.e. satisfying $|\boldsymbol{\omega}^{(n-1)} \cdot \boldsymbol{k}| \leqslant \sigma_{n-1}\|\boldsymbol{k}\|$ or $\|\boldsymbol{\nu}\| \geqslant \tau_{n-1}\|\boldsymbol{k}\|$ for some $\sigma_{n-1}, \tau_{n-1} > 0$. So, the nth step is defined by the following operations:

(1) Apply a linear operator corresponding to an affine symplectic transformation given by $(\boldsymbol{x}, \boldsymbol{y}) \mapsto ({T^{(n)}}^{-1}\boldsymbol{x},\ {}^{\top}T^{(n)}\boldsymbol{y} + \boldsymbol{b}_n)$ for some fixed vector $\boldsymbol{b}_n$.
(2) Rescale the action in order to 'zoom in' around the invariant torus.
(3) Rescale time (energy) to ensure that the frequency vector is in the form $\boldsymbol{\omega}^{(n)} = \binom{\alpha^{(n)}}{1}$.
(4) Eliminate the constant mode of the Hamiltonian.
(5) Eliminate all the modes outside the resonant cone I_n^{+} (thus avoiding dealing with small divisors) by a close to the identity symplectomorphism.

The first transformation above has a conjugate action $\boldsymbol{k} \mapsto {}^{\top}{T^{(n)}}^{-1}\boldsymbol{k}$. It follows from the hyperbolicity of $T^{(n)}$ that this transformation contracts I_{n-1}^{+} if σ_{n-1} and τ_{n-1}^{-1} are small enough. This significantly improves the analyticity domain in the $\boldsymbol{x}$ direction which implies the decrease in the estimates for the corresponding modes. As a result, all modes with $\boldsymbol{k} \neq 0$ become smaller.

Besides the (trivial) case $(\boldsymbol{k}, \boldsymbol{\nu}) = (0, 0)$ which is dealt with by operation (4), we control the size of the remaining $\boldsymbol{k} = 0$ modes in different ways. The case $S := \sum_i \nu_i = 1$ (corresponding to the linear term in the action $\boldsymbol{y}$) is eliminated by a proper choice of the affine parameter $\boldsymbol{b}_n$ depending on Q_{n-1} and the perturbation. That is, $\boldsymbol{b}_n$ is used to eliminate an unstable direction related to frequency vectors. The quadratic term in the action ($S = 2$) is included in the new symmetric matrix Q_n which has again a non-zero determinant and becomes smaller due to the action rescaling. Finally, we show that the action rescaling is also responsible for the decrease in the higher terms $S \geqslant 3$.

The overall consequence of the iterative scheme just described is that it converges to a limit set of Hamiltonians of the type $\boldsymbol{y} \mapsto \boldsymbol{v} \cdot \boldsymbol{y}$. That is, the 'limit' is a degenerate linear function of the action and from that we show the existence of an $\boldsymbol{\omega}$-invariant torus for the initial Hamiltonian. To prove convergence we need to find proper choices of σ_n and τ_n as well as of stopping times t_n, which turn out to be possible for diophantine $\boldsymbol{\omega}$. Roughly, too small values of σ_{n-1} and τ_{n-1}^{-1} make it harder to eliminate modes as they are 'too' resonant. On the other hand, large values imply that $T^{(n)}$ does not contract I_{n-1}^{+}. Similarly, large $t_n - t_{n-1}$ improve the hyperbolicity of the matrices $T^{(n)}$ but worsen the estimates on their norms and consequently enlarge the perturbation.

In section 2 we review the MCF algorithm contained in [6] and state estimates needed for the following sections. In section 3 we define the renormalization operator and iterate it to show convergence to a trivial limit set. We are then able to prove theorem 1.1 in section 4. In section 5 we present a proof of theorem 3.6 (similar to [1, 7]) that finds a symplectomorphism capable of eliminating the non-resonant modes of a Hamiltonian.

2. Multidimensional continued fractions

For completeness we review here the ideas contained in [6].

2.1. *Flow on homogeneous space*

Denote by $G = \mathrm{SL}(d, \mathbb{R})$, $\Gamma = \mathrm{SL}(d, \mathbb{Z})$ and take a fundamental domain $\mathcal{F} \subset G$ of the homogeneous space $\Gamma\backslash G$ (the space of d-dimensional non-degenerate unimodular lattices).

On $\mathcal{F}$ consider the flow:

$$\Phi^t\colon \mathcal{F} \to \mathcal{F}, \qquad M \mapsto P(t) M E^t, \tag{2.1}$$

where

$$E^t = \mathrm{diag}(\mathrm{e}^{-t}, \cdots, \mathrm{e}^{-t}, \mathrm{e}^{(d-1)t}) \in G$$

and $P(t)$ is the unique family in Γ that keeps $\Phi^t(M)$ in $\mathcal{F}$ for every $t \geqslant 0$.

Given $\boldsymbol{\omega} = \binom{\boldsymbol{\alpha}}{1} \in \mathbb{R}^d$, we are interested in the orbit under Φ^t of the matrix

$$M_\omega = \begin{pmatrix} I & \boldsymbol{\alpha} \\ \mathbf{0} & 1 \end{pmatrix}. \tag{2.2}$$

For this, consider a sequence of times

$$t_0 = 0 < t_1 < t_2 < \cdots \to +\infty \tag{2.3}$$

such that the matrices $P(t)$ in (2.1) satisfy

$$P^{(n)} := P(t_n) \neq P(t_{n-1}). \tag{2.4}$$

The sequence of matrices $P^{(n)} \in \mathrm{SL}(d, \mathbb{Z})$ are the rational approximates of $\boldsymbol{\omega}$, called multidimensional continued fractions expansion. In addition we define the *transfer matrices*

$$T^{(n)} = P^{(n)} {P^{(n-1)}}^{-1}. \tag{2.5}$$

The flow of M_ω taken at the time sequence is thus the sequence of matrices

$$M^{(n)} = \Phi^{t_n}(M_\omega) = P^{(n)} M_\omega E^{t_n}. \tag{2.6}$$

Using some properties of the flow, the above can be decomposed (see [6]) into

$$M^{(n)} = \begin{pmatrix} I & \boldsymbol{\alpha}^{(n)} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} A^{(n)} & \mathbf{0} \\ {}^\top\boldsymbol{\beta}^{(n)} & \gamma^{(n)} \end{pmatrix} \tag{2.7}$$

with $\gamma^{(n)}$ being the dth component of the vector $\mathrm{e}^{(d-1)t_n} P^{(n)} \boldsymbol{\omega}$, $A^{(n)}$ is a $(d-1) \times (d-1)$ real matrix and $\boldsymbol{\alpha}^{(n)}, \boldsymbol{\beta}^{(n)} \in \mathbb{R}^{d-1}$.

Define $\boldsymbol{\omega}^{(n)} = \begin{pmatrix} \boldsymbol{\alpha}^{(n)} \\ 1 \end{pmatrix}$, $\boldsymbol{\omega}^{(0)} = \boldsymbol{\omega}$ and, for $n \in \mathbb{N}$,

$$\boldsymbol{\omega}^{(n)} = {\gamma^{(n)}}^{-1} M^{(n)} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \lambda_n P^{(n)} \boldsymbol{\omega} = \eta_n T^{(n)} \boldsymbol{\omega}^{(n-1)}, \tag{2.8}$$

where

$$\lambda_n = \frac{1}{\gamma^{(n)}} \mathrm{e}^{(d-1)t_n} \qquad \text{and} \qquad \eta_n = \frac{\lambda_n}{\lambda_{n-1}}. \tag{2.9}$$

Consider now the cone

$$K^{(n)} = \{\boldsymbol{\xi} \in \mathbb{R}^d \colon |\boldsymbol{\xi} \cdot \boldsymbol{\omega}^{(n)}| \leqslant \sigma_n \|\boldsymbol{\xi}\|\} \tag{2.10}$$

for a given $\sigma_n > 0$. We are using the norm $\|\xi\| = \sum_{i=1}^d |\xi_i|$.

Let $\|\cdot\|$ denote the usual matrix norm

$$\|M\| := \sup_{\boldsymbol{x} \neq \mathbf{0}} \frac{\|M\boldsymbol{x}\|}{\|\boldsymbol{x}\|}. \tag{2.11}$$

Notice that any $A \in \mathrm{SL}(d, \mathbb{Z})$ has $\|A\| \geqslant 1$, as is the case of the norm of $T^{(n)}$, its inverse and transpose.

Lemma 2.1 ([6]). *If $\boldsymbol{\xi} \in K^{(n-1)}$, then there is $c_d > 0$ such that for all $n \in \mathbb{N}$*

$$\| {}^\top {T^{(n)}}^{-1} \boldsymbol{\xi}\| \leqslant c_d \left(\sigma_{n-1} \|{T^{(n)}}^{-1}\| + \mathrm{e}^{-\delta t_n} \|M^{(n-1)}\| \, \|{M^{(n)}}^{-1}\| \right) \|\boldsymbol{\xi}\|, \tag{2.12}$$

where $\delta t_n = t_n - t_{n-1}$.

2.2. Norm estimates for diophantine vectors

It is a well-known fact that the sets $DC(\beta)$ of diophantine vectors with exponent $\beta > 0$ are of full Lebesgue measure [2]. On the other hand, the set $DC(0)$ has zero Lebesgue measure.

Proposition 2.2 ([6]). *Let $\omega \in DC(\beta)$, $\beta \geqslant 0$. There are constants $c_1, c_2, c_3, c_4, c_5, c_6, c_7 > 0$ such that, for all $n \in \mathbb{N} \cup \{0\}$,*

$$\|M^{(n)}\| \leqslant c_1 \exp[(d-1)\theta t_n], \tag{2.13}$$

$$\|M^{(n)^{-1}}\| \leqslant c_2 \exp(\theta t_n), \tag{2.14}$$

$$\|P^{(n)}\| \leqslant c_3 \exp[(d\,\theta + 1 - \theta) t_n], \tag{2.15}$$

$$\|P^{(n)^{-1}}\| \leqslant c_4 \exp[(d - 1 + \theta) t_n], \tag{2.16}$$

$$\|T^{(n)}\| \leqslant c_5 \exp[(1-\theta)\delta t_n + d\,\theta\, t_n], \tag{2.17}$$

$$\|T^{(n)^{-1}}\| \leqslant c_6 \exp[(d-1)(1-\theta)\delta t_n + d\,\theta\, t_n] \tag{2.18}$$

and

$$c_7 \exp\left[-\theta\left(\frac{d^2}{1-\theta} - (d-1)\right)t_n\right] \leqslant |\gamma^{(n)}| \leqslant c_1 \exp[(d-1)\theta t_n], \tag{2.19}$$

where $\delta t_n = t_n - t_{n-1}$ and $\theta = \beta/(d+\beta)$.

Proposition 2.3. *Let $\omega \in DC(\beta)$, $\beta \geqslant 0$. If $\boldsymbol{\xi} \in K^{(n-1)}$, then there is $c_d > 0$ for all $n \in \mathbb{N}$*

$$\| {}^{\top}T^{(n)^{-1}} \boldsymbol{\xi}\| \leqslant c_d \mathrm{e}^{-(1-\theta)\delta t_n + d\theta t_{n-1}} \left(c_6 \sigma_{n-1} \mathrm{e}^{d\delta t_n} + c_1 c_2\right) \|\boldsymbol{\xi}\|, \tag{2.20}$$

with $\theta = \beta/(d+\beta)$.

Proof. The estimate follows from applying proposition 2.2 to lemma 2.1. □

3. Renormalization of Hamiltonian flows

3.1. Preliminaries

Consider the symplectic manifold $T^*\mathbb{T}^d$ with respect to the canonical symplectic form $\sum_{i=1}^d dy_i \wedge dx_i$. As the cotangent bundle of $\mathbb{T}^d$ is trivial, $T^*\mathbb{T}^d \simeq \mathbb{T}^d \times \mathbb{R}^d$, we identify functions on $T^*\mathbb{T}^d$ with functions on $\mathbb{T}^d \times \mathbb{R}^d$. By lifting to the universal cover, we consider functions from $\mathbb{R}^{2d}$ into $\mathbb{R}$ and extend them to the complex domain.

Let Ω be a neighbourhood of $\mathbb{R}^d \times \{0\}$ in $\mathbb{C}^{2d}$. A Hamiltonian is a complex analytic function $H: \Omega \to \mathbb{C}$, $\mathbb{Z}^d$-periodic on the first coordinate, written in the form of a Taylor–Fourier series

$$H(\boldsymbol{x}, \boldsymbol{y}) = \sum_{(\boldsymbol{k},\boldsymbol{\nu}) \in I} H_{\boldsymbol{k},\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k}\cdot\boldsymbol{x}}, \qquad (\boldsymbol{x}, \boldsymbol{y}) \in \Omega, \tag{3.1}$$

where $I = \mathbb{Z}^d \times (\mathbb{N} \cup \{0\})^d$, $H_{\boldsymbol{k},\boldsymbol{\nu}} \in \mathbb{C}$ and $\boldsymbol{y}^{\boldsymbol{\nu}} = y_1^{\nu_1} \cdots y_d^{\nu_d}$.

Let the positive real numbers ρ and r be given in order to determine the domain

$$\mathcal{D}_{\rho,r} = D_\rho \times B_r, \tag{3.2}$$

where

$$\begin{aligned} D_\rho &= \{\boldsymbol{x} \in \mathbb{C}^d : \|\mathrm{Im}\boldsymbol{x}\| < \rho/2\pi\} \text{ and} \\ B_r &= \{\boldsymbol{y} \in \mathbb{C}^d : \|\boldsymbol{y}\| < r\}, \end{aligned} \tag{3.3}$$

for the norm $\|\boldsymbol{u}\| = \sum_{i=1}^d |u_i|$ on $\mathbb{C}^d$. Moreover, we will be using the norm of matrices given by $\|Q\| = \max_{j=1,\ldots,d} \sum_{i=1}^d |Q_{i,j}|$, where $Q_{i,j}$ are the entries of a $d \times d$ matrix Q.

Consider the Banach space $\mathcal{A}_{\rho,r}$ of Hamiltonians defined on $\Omega = \mathcal{D}_{\rho,r}$, which extend continuously to the boundary and with finite norm

$$\|H\|_{\rho,r} = \sum_{(\boldsymbol{k},\boldsymbol{\nu})\in I} |H_{\boldsymbol{k},\nu}|\, r^{\|\boldsymbol{\nu}\|}\, \mathrm{e}^{\rho\|\boldsymbol{k}\|}. \tag{3.4}$$

Similarly, take a norm on the product space $\mathcal{A}^{2d}_{\rho,r} = \mathcal{A}_{\rho,r} \times \cdots \times \mathcal{A}_{\rho,r}$ given by $\|(H_1, \cdots, H_{2d})\|_{\rho,r} = \sum_{i=1}^{2d} \|H_i\|_{\rho,r}$. Using this we define the Banach space $\mathcal{A}'_{\rho,r}$ of Hamiltonians $H \in \mathcal{A}_{\rho,r}$ with finite norm

$$\|H\|'_{\rho,r} = \|H\|_{\rho,r} + \|\nabla H\|_{\rho,r}.$$

A property that will be used several times in this paper is the Cauchy estimate: for any $\delta > 0$ we have

$$\begin{aligned} &\|\partial_i H\|_{\rho,r} \leqslant \tfrac{2\pi}{\delta}\|H\|_{\rho+\delta,r}, && H \in \mathcal{A}_{\rho+\delta,r}, \quad 1 \leqslant i \leqslant d,\\ &\|\partial_j H\|_{\rho,r} \leqslant \tfrac{1}{\delta}\|H\|_{\rho,r+\delta}, && H \in \mathcal{A}_{\rho,r+\delta}, \quad d+1 \leqslant j \leqslant 2d, \end{aligned} \tag{3.5}$$

where ∂_k denotes the partial derivative with respect to the kth argument. In particular

$$\|H\|'_{\rho,r} \leqslant \left(1 + \frac{2\pi + 1}{\delta}\right) \|H\|_{\rho+\delta,r+\delta}. \tag{3.6}$$

The constant Fourier modes will be written by the projection

$$\mathbb{E}F(\boldsymbol{y}) = \int_{\mathbb{T}^d} F(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{x} = \sum_{\boldsymbol{\nu}} F_{0,\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}}, \qquad \mathbb{E}_{\nu}F(\boldsymbol{y}) = F_{0,\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}}. \tag{3.7}$$

The space where $\mathbb{E}F$ lies is denoted by $\mathbb{E}\mathcal{A}_r$ and the natural induced norm is $\|\cdot\|_r$. Similarly, we define $\mathbb{E}\mathcal{A}'_r$ with norm $\|\cdot\|'_r$.

In the following we will use the notation $A \ll B$ to mean that there is a constant $C > 0$ such that $A \leqslant CB$.

Remark 3.1. We will be dealing with maps between Banach spaces over $\mathbb{C}$ with a notion of analyticity stated as follows (cf e.g. [5]): a map F defined on a domain is analytic if it is locally bounded and Gâteux differentiable. If it is analytic on a domain, it is continuous and Fréchet differentiable. Moreover, we have a convergence theorem which is going to be used later on. Let $\{F_k\}$ be a sequence of functions analytic and uniformly locally bounded on a domain D. If $\lim_{k\to+\infty} F_k = F$ on D, then F is analytic on D.

3.2. *Change of basis and rescaling*

The following transformations leave invariant the dynamics of the flow generated by a Hamiltonian, producing an equivalent system. They consist of

- an affine symplectic transformation of the phase space,

$$L_n\colon (\boldsymbol{x},\boldsymbol{y}) \mapsto ({T^{(n)}}^{-1}\boldsymbol{x},\ {}^{\top}T^{(n)}\boldsymbol{y} + \boldsymbol{b}_n), \qquad (\boldsymbol{x},\boldsymbol{y}) \in \mathbb{C}^{2d}, \tag{3.8}$$

  for some $\boldsymbol{b}_n \in \mathbb{C}^d$,
- a linear time (energy) change,

$$H \mapsto \eta_n H \tag{3.9}$$

  where η_n is defined in (2.9),
- a linear action rescaling,

$$H \mapsto \frac{1}{\mu_n} H(\cdot, \mu_n \cdot) \tag{3.10}$$

  with a choice of $\mu_n > 0$ to be specified later on,

- and the (trivial) elimination of the constant term

$$H \mapsto (\mathbb{I} - \mathbb{E}_0)H. \tag{3.11}$$

Notice that $\mathbb{E}H \circ R_z = \mathbb{E}H$ and

$$R_z \circ L_n = L_n \circ R_{T^{(n)}z} \tag{3.12}$$

with

$$R_z\colon (\boldsymbol{x}, \boldsymbol{y}) \mapsto (\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y}), \qquad z \in \mathbb{C}^d. \tag{3.13}$$

For $n \in \mathbb{N}$, $\rho_{n-1} > 0$ and $r > 0$, we are going to apply the transformations (3.8)–(3.11) to Hamiltonians of the form

$$H(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\omega}^{(n-1)} \cdot \boldsymbol{y} + \tfrac{1}{2}\, {}^{\top}\boldsymbol{y} Q_{n-1} \boldsymbol{y} + F(\boldsymbol{x}, \boldsymbol{y}), \qquad (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{D}_{\rho_{n-1}, r}, \tag{3.14}$$

where Q_{n-1} is a $d \times d$ symmetric matrix and $F \in \mathcal{A}_{\rho_{n-1}, r}$. We thus get new Hamiltonians which are images under the map

$$\mathcal{L}_n(H) = (\mathbb{I} - \mathbb{E}_0)\frac{\eta_n}{\mu_n} H \circ L_n(\cdot, \mu_n \cdot).$$

In order to simplify notations, we write

$$\Phi_n(\boldsymbol{y}) = \mu_n\, {}^{\top}T^{(n)}\boldsymbol{y} + \boldsymbol{b}_n. \tag{3.15}$$

So, for any $(\boldsymbol{x}, \boldsymbol{y}) \in L_n^{-1}\mathcal{D}_{\rho_{n-1}, r}$,

$$\begin{aligned} &\mathcal{L}_n(H)(\boldsymbol{x}, \boldsymbol{y}) \\ &\quad = (\mathbb{I} - \mathbb{E}_0)\frac{\eta_n}{\mu_n}\left[\boldsymbol{\omega}^{(n-1)} \cdot \Phi_n(\boldsymbol{y}) + \frac{1}{2}\, {}^{\top}\Phi_n(\boldsymbol{y}) Q_{n-1} \Phi_n(\boldsymbol{y}) + F \circ L_n(\boldsymbol{x}, \mu_n \boldsymbol{y})\right]. \end{aligned} \tag{3.16}$$

By the decomposition $F = (\mathbb{I} - \mathbb{E})F + F_0$ and using the Taylor expansion of F_0:

$$F_0 \circ \Phi_n(\boldsymbol{y}) = F_0(\boldsymbol{b}_n) + \mu_n\, {}^{\top}\nabla F_0(\boldsymbol{b}_n)\, {}^{\top}T^{(n)}\boldsymbol{y} + \frac{\mu_n^2}{2}\, {}^{\top}\boldsymbol{y} T^{(n)} D^2 F_0(\boldsymbol{b}_n)\, {}^{\top}T^{(n)}\boldsymbol{y} + \Upsilon_n(\boldsymbol{y}), \tag{3.17}$$

with $\Upsilon_n(\boldsymbol{y}) = \mathcal{O}(\|\boldsymbol{y}\|^3)$, we get

$$\begin{aligned} \mathcal{L}_n(H)(\boldsymbol{x}, \boldsymbol{y}) = {} & \boldsymbol{\omega}^{(n)} \cdot \boldsymbol{y} + \eta_n \left[{}^{\top}\boldsymbol{b}_n Q_{n-1} + {}^{\top}\nabla F_0(\boldsymbol{b}_n)\right] {}^{\top}T^{(n)}\boldsymbol{y} \\ & + \frac{\eta_n \mu_n}{2}\, {}^{\top}\boldsymbol{y} T^{(n)} \left[Q_{n-1} + D^2 F_0(\boldsymbol{b}_n)\right] {}^{\top}T^{(n)}\boldsymbol{y} \\ & + \frac{\eta_n}{\mu_n}\Upsilon_n(\boldsymbol{y}) + \frac{\eta_n}{\mu_n}(\mathbb{I} - \mathbb{E})F \circ L_n(\boldsymbol{x}, \mu_n \boldsymbol{y}). \end{aligned} \tag{3.18}$$

In order to ‘normalize’ the (Fourier constant) linear term in $\boldsymbol{y}$ of $\mathbb{E}\mathcal{L}_n(H)$ by making it equal to $\omega^{(n)} \cdot \boldsymbol{y}$, we choose $\boldsymbol{b}_n$ inside the domain of ∇F_0 such that

$$Q_{n-1}\boldsymbol{b}_n + \nabla F_0(\boldsymbol{b}_n) = 0. \tag{3.19}$$

The quadratic term is dealt with by considering a new symmetric $d \times d$ matrix Q_n being

$$Q_n = \eta_n \mu_n T^{(n)} \left[Q_{n-1} + D^2 F_0(\boldsymbol{b}_n)\right] {}^{\top}T^{(n)}. \tag{3.20}$$

We can finally write

$$\mathcal{L}_n(H)(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\omega}^{(n)} \cdot \boldsymbol{y} + \tfrac{1}{2}\, {}^{\top}\boldsymbol{y} Q_n \boldsymbol{y} + \widehat{\mathcal{L}}_n(F_0)(\boldsymbol{y}) + \widetilde{\mathcal{L}}_n(F - F_0)(\boldsymbol{x}, \boldsymbol{y}), \tag{3.21}$$

where we have introduced the operator

$$\widehat{\mathcal{L}}_n\colon F_0 \mapsto \frac{\eta_n}{\mu_n}\Upsilon_n \tag{3.22}$$

for the cubic and higher terms in $\boldsymbol{y}$ and

$$\widetilde{\mathcal{L}}_n\colon (\mathbb{I} - \mathbb{E})F \mapsto \frac{\eta_n}{\mu_n}(\mathbb{I} - \mathbb{E})F \circ L_n(\cdot, \mu_n \cdot) \tag{3.23}$$

for the non-constant Fourier modes. The above operators are defined in $\mathbb{E}\mathcal{A}_r$ and $(\mathbb{I} - \mathbb{E})\mathcal{A}_{\rho_{n-1}, r}$.

For a given $\gamma > 0$, denote by Δ_γ the set of all H as in (3.14) such that $\|F_0\|_{\rho_{n-1}, r} < \gamma$.

Lemma 3.2. *If* $\det(Q_{n-1}) \neq 0$ *and*

$$\gamma_n = \frac{r^2}{16\|Q_{n-1}^{-1}\|}, \tag{3.24}$$

there is $b_n \in C^1(\Delta_{\gamma_n}, \mathbb{C}^d)$ *such that, for all* $H \in \Delta_{\gamma_n}$, $\boldsymbol{b}_n = b_n(H)$ *satisfies* (3.19) *and*

$$\|b_n(H)\| < (2/r)\|Q_{n-1}^{-1}\|\,\|F_0\|_r < \frac{r}{8}. \tag{3.25}$$

Moreover, $\det(Q_n) \neq 0$ *where* Q_n *is given by* (3.20) *and*

$$\|Q_n^{-1}\| \leqslant \frac{\|T^{(n)^{-1}}\|\,\|{}^{\top}T^{(n)^{-1}}\|}{\mu_n|\eta_n|(\|Q_{n-1}^{-1}\|^{-1} - \frac{16}{r^2}\|F_0\|_r)}. \tag{3.26}$$

In the case F_0 *is real-analytic and* Q_{n-1} *is real,* $b_n(H) \in \mathbb{R}^d$ *and* Q_n *is also real.*

Proof. Consider the differentiable function $\mathcal{F}(H, \boldsymbol{b}) = \boldsymbol{b} + Q_{n-1}^{-1}\nabla F_0(\boldsymbol{b})$ defined on $\Delta_{\gamma_n} \times B_{r/2}$. Notice that $\mathcal{F}(H_{n-1}^0, 0) = 0$. Moreover, the derivative of $\mathcal{F}$ with respect to the second argument,

$$D_2\mathcal{F}(H, \boldsymbol{b}) = I + Q_{n-1}^{-1}D^2F_0(\boldsymbol{b}), \qquad (H, \boldsymbol{b}) \in \Delta_{\gamma_n} \times B_{r/2},$$

admits a bounded inverse because

$$\begin{aligned}\|D^2F_0\|_{r/2} &= \max_{d+1\leqslant j\leqslant 2d} \|\partial_j \nabla F_0\|_{r/2} \\ &\leqslant (4/r)\|\nabla F_0\|_{3r/4} \\ &\leqslant (16/r^2)\|F_0\|_r \\ &< \|Q_{n-1}^{-1}\|^{-1}\end{aligned} \tag{3.27}$$

by the Cauchy estimate. Thus, the implicit function theorem implies the existence of a C^1 function $b_n\colon H \mapsto b_n(H)$ in a neighbourhood of H_{n-1}^0 such that

$$\mathcal{F}(H, b_n(H)) = b_n(H) + Q_{n-1}^{-1}\nabla F_0(b_n(H)) = 0,$$

i.e. a solution of (3.19). Notice that for any $H \in \Delta_{\gamma_n}$ the operator $\mathrm{Id} - \mathcal{F}(H, \cdot)$ is a contraction with a unique fixed point $b_n(H)$. Hence the domain of the C^1 function $H \mapsto b(H)$ is extendable to Δ_{γ_n} and thus (3.25). Assuming F_0 to be real-analytic and Q_{n-1} with real entries, the same argument is still valid when considering $B_{r/2} \cap \mathbb{R}^d$. So, $b(H)$ is real and Q_n is a real symmetric matrix.

From (3.27),

$$\|Q_{n-1}^{-1}D^2F_0(b_n(H))\| < 1, \qquad H \in \Delta_{\gamma_n}.$$

Hence, $A = Q_{n-1}[I + Q_{n-1}^{-1}D^2F_0(b_n(H))]$ is invertible. Moreover,

$$\|A^{-1}\| \leqslant 1/(\|Q_{n-1}^{-1}\|^{-1} - \|D^2F_0\|_{r/2}). \tag{3.28}$$

Now, $Q_n^{-1} = (\eta_n\mu_n)^{-1}\,{}^{\top}T^{(n)^{-1}}A^{-1}T^{(n)^{-1}}$, thus (3.26). □

Lemma 3.3. *If* $r < r'$ *and*

$$\mu_n < \frac{r}{4r'\|{}^{\top}T^{(n)}\|}, \tag{3.29}$$

then $\widehat{\mathcal{L}}_n\colon \mathbb{E}\mathcal{A}_r \cap \Delta_{\gamma_n} \to \mathbb{E}\mathcal{A}'_{r'}$ *and*

$$\|\widehat{\mathcal{L}}_n\| \leqslant \mu_n^2|\eta_n|\left(1 + \frac{1}{2r'}\right)\frac{(4r'\|{}^{\top}T^{(n)}\|)^3}{r^2(r - 4r'\mu_n\|{}^{\top}T^{(n)}\|)}. \tag{3.30}$$

Proof. Let $H \in \Delta_{\gamma_n}$, $R = \frac{r}{4r'\mu_n \| {}^\top T^{(n)} \|} > 1$, $\boldsymbol{y} \in B_{r'}$ and the map

$$
\begin{aligned}
&f\colon \{z \in \mathbb{C}\colon |z| \leqslant R\} \to \mathbb{C}^d \\
&z \mapsto F_0(z\mu_n\, {}^\top T^{(n)}\boldsymbol{y} + b_n(H)).
\end{aligned} \tag{3.31}
$$

Hence Υ_n as in (3.17) can be written as

$$
f(1) - f(0) - Df(0) - \frac{1}{2}D^2 f(0) = \frac{1}{2\pi \mathrm{i}} \oint_{|z|=R} \frac{f(z)}{z^3(z-1)} \mathrm{d}z.
$$

Therefore,

$$
\begin{aligned}
\|\Upsilon_n\|'_{r'} &= \frac{1}{2\pi} \left\| \oint_{|z|=R} \frac{f(z)}{z^3(z-1)} \mathrm{d}z \right\|'_{r'} \\
&\leqslant \frac{1}{R^2(R-1)} \sup_{|z|=R} \|F_0(z\mu_n\, {}^\top T^{(n)} \cdot + b_n(H))\|'_{r'}.
\end{aligned}
$$

Since $\|\boldsymbol{y}\| < r'$, in view of (3.25),

$$
\sup_{|z|=R} \|z\mu_n\, {}^\top T^{(n)}\boldsymbol{y} + b_n(H)\| \leqslant R\mu_n \|{}^\top T^{(n)}\|\, r' + \|b_n(H)\| < r/2
$$

and

$$
\begin{aligned}
\sup_{|z|=R} \|F_0(z\mu_n\, {}^\top T^{(n)} \cdot + b_n(H))\|'_{r'} &\leqslant \|F_0\|_{r/2} + R\mu_n \|{}^\top T^{(n)}\|\, \|\nabla F_0\|_{r/2} \\
&\leqslant \|F_0\|_{r/2} + \frac{1}{2r'} \|F_0\|_r \leqslant \left(1 + \frac{1}{2r'}\right) \|F_0\|_r.
\end{aligned} \tag{3.32}
$$

Thus, $\|\Upsilon_n\|'_{r'} \leqslant (1 + 1/2r')[R^2(R-1)]^{-1} \|F_0\|_r$ and

$$
\|\widehat{\mathcal{L}}_n(F_0)\|'_{r'} = \frac{|\eta_n|}{\mu_n} \|\Upsilon_n\|'_{r'} \leqslant \frac{|\eta_n|}{\mu_n} \left(1 + \frac{1}{2r'}\right) \frac{(4r'|\mu_n|\, \|{}^\top T^{(n)}\|)^3}{r^2(r - 4r'|\mu_n|\, \|{}^\top T^{(n)}\|)} \|F_0\|_r.
$$

□

3.3. *Far from resonance modes*

Given $\sigma_n, \tau_n > 0$, we call *far from resonance* modes with respect to $\boldsymbol{\omega}^{(n)}$ the Taylor–Fourier modes with indices in

$$
I_n^- = \left\{ (\boldsymbol{k}, \boldsymbol{\nu}) \in I\colon |\boldsymbol{\omega}^{(n)} \cdot \boldsymbol{k}| > \sigma_n \|\boldsymbol{k}\|,\ \|\boldsymbol{\nu}\| < \tau_n \|\boldsymbol{k}\| \right\}. \tag{3.33}
$$

The *resonant modes* are the ones in $I_n^+ = I - I_n^-$. We also have the projections $\mathbb{I}_n^+$ and $\mathbb{I}_n^-$ over the spaces of Hamiltonians by restricting the Taylor–Fourier modes to I_n^+ and I_n^-, respectively. The identity operator is $\mathbb{I} = \mathbb{I}_n^+ + \mathbb{I}_n^-$.

Moreover, take

$$
A_n = \sup_{\boldsymbol{k} \neq \boldsymbol{0},\, |\omega^{(n)} \cdot \boldsymbol{k}| \leqslant \sigma_n \|\boldsymbol{k}\|} \frac{\| {}^\top T^{(n+1)^{-1}} \boldsymbol{k}\|}{\|\boldsymbol{k}\|}. \tag{3.34}
$$

3.4. *Analyticity improvement*

The next lemma means that every Hamiltonian in $\mathbb{I}_{n-1}^+ \mathcal{A}_{\rho_{n-1},r} \cap \Delta_{\gamma_n}$, i.e. a function on $\mathcal{D}_{\rho_{n-1},r}$ into $\mathbb{C}$, is mapped by $\mathcal{L}_n$ into $\mathcal{A}'_{\rho_n',r'}$. The analyticity strip width is improved whenever A_{n-1} is small enough. Lemma 3.5 will 'convert' this improvement into a norm reduction.

Lemma 3.4. *If* $\delta > 0$, $r < r'$,

$$\rho'_n \leqslant \frac{\rho_{n-1}}{A_{n-1}} - \delta \qquad \textit{and} \qquad \tau_n \geqslant \frac{2}{\log 2}(\rho'_n + \delta)\| {}^{\top}T^{(n)^{-1}}\|, \tag{3.35}$$

then $\widetilde{\mathcal{L}}_n$ *as a map from* $(\mathbb{I}^+_{n-1} - \mathbb{E})\mathcal{A}_{\rho_{n-1},r} \cap \Delta_{\gamma_n}$ *to* $(\mathbb{I} - \mathbb{E})\mathcal{A}'_{\rho'_n,r'}$ *is continuous with*

$$\|\widetilde{\mathcal{L}}_n\| \leqslant \left(1 + \frac{2\pi}{\delta} + \frac{r}{2r'^2 \log 2}\right)\frac{|\eta_n|}{\mu_n}. \tag{3.36}$$

Proof. Let $F \in (\mathbb{I}^+_{n-1} - \mathbb{E})\mathcal{A}_{\rho_{n-1},r} \cap \Delta_{\gamma_n}$,

$$E = \{(0,\boldsymbol{\nu}) : \boldsymbol{\nu} \in (\mathbb{N} \times \{0\})^d\} \qquad \text{and} \qquad J_n = \{\boldsymbol{k} \in \mathbb{Z}^d : |\boldsymbol{k} \cdot \boldsymbol{\omega}^{(n)}| \leqslant \sigma_n \|\boldsymbol{k}\|\}. \tag{3.37}$$

Using lemma 3.2 and (3.29) we have

$$\psi_n = \mu_n \| {}^{\top}T^{(n)}\| \, r' + \|b_n(H)\| \leqslant \frac{r}{4} + \frac{2}{r}\|Q^{-1}_{n-1}\| \, \|F_0\|_r < \frac{r}{2}. \tag{3.38}$$

We want to find an upper bound on

$$\begin{aligned} &\|F \circ L_n(\cdot, \mu_n \cdot)\|'_{\rho'_n,r'} \\ &\quad \leqslant \sum_{I^+_{n-1}-E} \left(1 + 2\pi \| {}^{\top}T^{(n)^{-1}}\boldsymbol{k}\| + \mu_n \| {}^{\top}T^{(n)}\| \, \|\boldsymbol{\nu}\|/r'\right)|F_{\boldsymbol{k},\nu}|\psi_n^{\|\boldsymbol{\nu}\|} \mathrm{e}^{\rho'_n \| {}^{\top}T^{(n)^{-1}}\boldsymbol{k}\|} \\ &\quad \leqslant \sum_{I^+_{n-1}-E} \left(1 + \frac{2\pi}{\delta}\mathrm{e}^{\delta \| {}^{\top}T^{(n)^{-1}}\boldsymbol{k}\|} + \frac{r}{4r'^2\xi_n}\mathrm{e}^{\xi_n\|\boldsymbol{\nu}\|}\right)|F_{\boldsymbol{k},\nu}|\psi_n^{\|\boldsymbol{\nu}\|}\mathrm{e}^{\rho'_n \| {}^{\top}T^{(n)^{-1}}\boldsymbol{k}\|}, \end{aligned} \tag{3.39}$$

where we have used the inequality $\zeta\,\mathrm{e}^{-\delta\zeta} \leqslant \delta^{-1}$ with $\zeta \geqslant 0$ and again a choice of μ_n verifying (3.29). Here $\xi_n = \frac{1}{2}\log(r/\psi_n) > \frac{1}{2}\log 2$.

Consider separately the two cases corresponding to the definition of the resonance cone I^+_{n-1}. We deal first with the modes corresponding to $\boldsymbol{k} \in J_{n-1} - \{0\}$. By (3.34) and (3.35) each one of these modes in (3.39) is bounded from above by

$$\left(1 + \frac{2\pi}{\delta} + \frac{r}{2r'^2 \log 2}\right) r^{\|\boldsymbol{\nu}\|}\mathrm{e}^{\rho_{n-1}\|\boldsymbol{k}\|}. \tag{3.40}$$

Now, consider $\|\boldsymbol{\nu}\| \geqslant \tau_n\|\boldsymbol{k}\|$ with $\boldsymbol{k} \neq 0$, so that

$$\| {}^{\top}T^{(n)^{-1}}\boldsymbol{k}\| \leqslant \tau_n^{-1}\| {}^{\top}T^{(n)^{-1}}\| \, \|\boldsymbol{\nu}\|. \tag{3.41}$$

These modes in (3.39) are estimated by

$$\left(1 + \frac{2\pi}{\delta} + \frac{r}{4r'^2\xi_n}\mathrm{e}^{\xi_n\|\boldsymbol{\nu}\|}\right)\left(r\,\mathrm{e}^{-2\xi_n + (\rho'_n+\delta)\| {}^{\top}T^{(n)^{-1}}\|/\tau_n}\right)^{\|\boldsymbol{\nu}\|} \leqslant \left(1 + \frac{2\pi}{\delta} + \frac{r}{2r'^2\log 2}\right) r^{\|\boldsymbol{\nu}\|}, \tag{3.42}$$

where we have used (3.35).

Finally, we get

$$\|F \circ L_n(\cdot, \mu_n \cdot)\|'_{\rho'_n,r'} \leqslant \left(1 + \frac{2\pi}{\delta} + \frac{r}{2r'^2 \log 2}\right)\|F\|_{\rho_{n-1},r},$$

and (3.36) follows from (3.23). □

Let $0 < \rho_n'' \leqslant \rho_n'$ and the inclusion

$$\mathcal{I}_n : \mathcal{A}'_{\rho_n',r'} \to \mathcal{A}'_{\rho_n'',r'}, \qquad H \mapsto H|\mathcal{D}_{\rho_n'',r'}. \tag{3.43}$$

The norm of the $\boldsymbol{k} \neq 0$ modes can be improved by the application of $\mathcal{I}_n$.

Lemma 3.5. *If $\phi_n \geqslant 1$ and*

$$0 < \rho_n'' \leqslant \rho_n' - \log(\phi_n), \tag{3.44}$$

then

$$\|\mathcal{I}_n(\mathbb{I} - \mathbb{E})\| \leqslant \phi_n^{-1}. \tag{3.45}$$

The proof is immediate and will be omitted.

3.5. Elimination of far from resonance modes

The theorem below states the existence of a symplectomorphism isotopic to the identity that cancels the far from resonance modes of a Hamiltonian close to the quadratic integrable Hamiltonian

$$H_n^0 : \boldsymbol{y} \mapsto \boldsymbol{\omega}^{(n)} \cdot \boldsymbol{y} + \tfrac{1}{2}\,{}^\top\boldsymbol{y} Q_n \boldsymbol{y}. \tag{3.46}$$

Given $\rho_n, \nu > 0$, denote by $\mathcal{V}_\varepsilon$ the open ball in $\mathcal{A}'_{\rho_n+\nu,r'}$ centred at H_n^0 with radius $\varepsilon > 0$. We define also

$$\varepsilon_n = \frac{\sigma_n^2 (\min\left\{1, \frac{\nu}{2\pi}, r' - r\right\})^2}{12(4\|\boldsymbol{\omega}^{(n)}\| + d\sigma_n) r' (2\pi + 1)^2 (1 + 2\pi + \frac{\tau_n + 1}{r'})^2} \tag{3.47}$$

and

$$\varphi_n = 1 + \sqrt{3r' \frac{4\|\boldsymbol{\omega}^{(n)}\| + d\sigma_n}{\varepsilon_n}}. \tag{3.48}$$

Theorem 3.6. *Let $r < r'$ and $\sigma_n > 2r'\|Q_n\|$. Then there exist analytic maps $\mathfrak{G} : \mathcal{V}_{\varepsilon_n} \to \mathcal{A}^{2d}_{\rho_n,r}$ where $\mathfrak{G}(H)$ is a symplectomorphism and $\mathcal{U} : \mathcal{V}_{\varepsilon_n} \to \mathbb{I}_n^+ \mathcal{A}_{\rho_n,r}$ given by $\mathcal{U}(H) = H \circ \mathfrak{G}(H)$, such that $\mathbb{I}_n^- \mathcal{U}(H) = 0$ and*

$$\begin{aligned} \|\mathfrak{G}(H) - \mathrm{Id}\|'_{\rho_n,r} &\leqslant \frac{1}{\varepsilon_n} \|\mathbb{I}_n^- H\|_{\rho_n,r}, \\ \|\mathcal{U}(H) - H_n^0\|_{\rho_n,r} &\leqslant \varphi_n \|H - H_n^0\|'_{\rho_n+\nu,r'}. \end{aligned} \tag{3.49}$$

Moreover, if H is real-analytic, then $\mathfrak{G}(H)$ is real-analytic.

A proof of this theorem is included in section 5.

Lemma 3.7. *In the conditions of theorem 3.6, if $\boldsymbol{x} \in \mathbb{R}^d$ and $H \in \mathcal{V}_{\epsilon_n}$, then*

$$\mathfrak{G}(H \circ R_{\boldsymbol{x}}) = R_{\boldsymbol{x}}^{-1} \circ \mathfrak{G}(H) \circ R_{\boldsymbol{x}} \tag{3.50}$$

on $\mathcal{D}_{\rho_n,r}$.

Proof. If $g = \mathfrak{G}(H)$ is a solution of $\mathbb{I}_n^- H \circ g = 0$ in $\mathcal{D}_{\rho_n,r}$, then $\widetilde{g} = R_{\boldsymbol{x}}^{-1} \circ \mathfrak{G}(H) \circ R_{\boldsymbol{x}}$ solves the same equation for $\widetilde{H} = H \circ R_{\boldsymbol{x}}$, i.e. $\mathbb{I}^- \widetilde{H} \circ \widetilde{g} = 0$ in $\mathcal{D}_{\rho_n,r}$. □

3.6. Convergence of renormalization

For a resonance set I_n^+ and $\mu_n > 0$, the nth *step renormalization* operator is defined to be

$$\mathcal{R}_n = \mathcal{U}_n \circ \mathcal{I}_n \circ \mathcal{L}_n \circ \mathcal{R}_{n-1} \qquad \text{and} \qquad \mathcal{R}_0 = \mathcal{U}_0,$$

where $\mathcal{U}_n$ is as in theorem 3.6 at the step n. Notice that if

$$H^+(\boldsymbol{y}) = \boldsymbol{\omega} \cdot \boldsymbol{y} + \tfrac{1}{2}\,{}^\top\boldsymbol{y} Q \boldsymbol{y} + \boldsymbol{v} \cdot \boldsymbol{y},$$

then

$$\mathcal{R}_n(H^+) = \boldsymbol{\omega}^{(n)} \cdot \boldsymbol{y} + \tfrac{1}{2}\lambda_n \chi_n\, {}^\top\boldsymbol{y} P^{(n)} Q\, {}^\top P^{(n)} \boldsymbol{y}$$

for every $\boldsymbol{v} \in \mathbb{C}^d$, where

$$\chi_n = \prod_{i=1}^{n} \mu_i. \tag{3.51}$$

This means that the renormalizations eliminate the direction corresponding to linear terms in $\boldsymbol{y}$. From the previous sections the map $\mathcal{R}_n$ on its domain of validity is analytic by construction. In addition, whenever a Hamiltonian H is real-analytic, the same is true for $\mathcal{R}_n(H)$.

Let $r' > r > 0$, $\rho_0 > 0$ and fix a sequence $\sigma_n < 1$, $n \in \mathbb{N}$ and $\sigma_0 > 2r'\|Q\|$. To complete the specification of the resonant modes and of ε_n in theorem 3.6, take $\tau_0 = 1$ and

$$\tau_n = \frac{2\rho_0 \|\,{}^\top T^{(n)-1}\|}{B_{n-1} \log 2} \tag{3.52}$$

according to lemma 3.4, with

$$B_n = \prod_{i=0}^{n} A_i. \tag{3.53}$$

Notice that the A_ns depend on σ_n.

Consider also the constants ν and δ as they appear in theorem 3.6 and lemma 3.4, respectively.

We now define the non-increasing sequence $\Theta_0 = 1$,

$$\begin{aligned}\Theta_n = \min \Bigg\{ &\Theta_{n-1}, \frac{\sigma_n^2}{(4r'\|Q\|)^2} \prod_{i=1}^{n} \frac{2^6 \zeta_i}{|\eta_i|\, \|T^{(i)}\|^2 \|\,{}^\top T^{(i)}\|^2}, \\ &\frac{\varepsilon_n^3}{\prod_{i=1}^{n} \|T^{(i)-1}\|^3}, \prod_{i=1}^{n} \frac{\min\{|\eta_i|^{-3}, |\eta_i|^2\}}{2^{24}\zeta_i^3 \|T^{(i)-1}\|^2 \|\,{}^\top T^{(i)-1}\|^6} \Bigg\} \leqslant 1,\end{aligned} \tag{3.54}$$

with

$$\zeta_n = \left(1 + \frac{1}{2r'}\right)\left(\frac{r'}{r}\right)^3 \varphi_n \|\,{}^\top T^{(n)}\|^3 > 1.$$

In order to use the results obtained earlier connected with the building blocks of the renormalization operator and to get convergence of the renormalization (in the theorem below), we choose

$$\rho_n = \frac{1}{B_{n-1}} \left[\rho_0 - \sum_{i=0}^{n-1} B_i \log(\phi_{i+1}) - (\delta + \nu) \sum_{i=0}^{n-1} B_i \right], \tag{3.55}$$

where

$$\begin{aligned} \phi_n &= \max\left\{1, 2\left(1 + \tfrac{2\pi}{\delta} + \tfrac{r}{2r'^2 \log 2}\right) \tfrac{\varphi_n |\eta_n| \Theta_{n-1}}{\mu_n \Theta_n}\right\} \geqslant 1, \\ \mu_n &= \left(\tfrac{\Theta_n}{2^8 \zeta_n \max\{1, |\eta_n|\} \Theta_{n-1}}\right)^{1/2} \leqslant 1. \end{aligned} \tag{3.56}$$

Recall that ϕ_n is our choice for lemma 3.5. Moreover, our choice of μ_n implies that

$$\mu_n \leqslant \frac{1}{2^4 \zeta_n^{1/2}} \leqslant \frac{1}{2^4}\left(\frac{r}{r' \| {}^\top T^{(n)} \|}\right)^{3/2} \leqslant \frac{r}{8r' \| {}^\top T^{(n)} \|}, \tag{3.57}$$

so lemma 3.3 holds.

To have ρ_n positive for all n we need to study the following function of $\boldsymbol{\omega} \in \mathbb{R}^d$ associated to the choice of σ_n:

$$\mathcal{B}(\boldsymbol{\omega}) = \sum_{i=0}^{+\infty} B_i \log(\phi_{i+1}) + (\delta + \nu) \sum_{i=0}^{+\infty} B_i. \tag{3.58}$$

It is simple to see that $\mathcal{B}$ depends on the multidimensional continued fraction expansion of $\boldsymbol{\omega}$ through the matrices $T^{(n)}$ and the scalars η_n. The remaining dependences are on fixed constants and on Q, but these turn out to be irrelevant as we will be uniquely interested in the convergence of the series in (3.58). In this sense, we can look at $\mathcal{B}$ as only depending on the arithmetics of $\boldsymbol{\omega}$. As we will see in the following part of this section, for diophantine vectors $\boldsymbol{\omega}$ we can find a sequence σ_n for which $\mathcal{B}(\boldsymbol{\omega})$ converges.

Notice that if $\mathcal{B}(\boldsymbol{\omega})$ converges, then $B_n \to 0$ as $n \to +\infty$. Also, $\tau_n \gg B_{n-1}^{-1} \to \infty$ by (3.52) and $\varepsilon_n \ll \tau_n^{-2} \to 0$ by (3.47). Hence, $\Theta_n \ll \varepsilon_n^3 \to 0$ by the third term in $\min\{\cdots\}$ of (3.54).

We denote

$$H_n = \mathcal{R}_n(H)$$

and associate the sequence H_n^0 of quadratic integrable Hamiltonians given by (3.46), where Q_n is defined by (3.20).

Theorem 3.8. *Suppose that* $\det(Q) \neq 0$*,*

$$\mathcal{B}(\boldsymbol{\omega}) < +\infty, \tag{3.59}$$

and $\rho > \mathcal{B}(\boldsymbol{\omega}) + \nu$*. There exists* $c, K > 0$ *such that if* $H \in \mathcal{A}_{\rho,r'}$ *and* $\|H - H^0\|_{\rho,r'} < c$*, then* H *is in the domain of* $\mathcal{R}_n$ *and*

$$\|H_n - H_n^0\|_{\rho_n,r} \leqslant K\Theta_n \|H - H^0\|_{\rho,r'}, \qquad n \in \mathbb{N} \cup \{0\}. \tag{3.60}$$

Proof. Let $\rho_0 = \rho - \nu > \mathcal{B}(\boldsymbol{\omega})$. Hence, by the definition of ρ_n, there is $R > 0$ satisfying $\rho_n > R B_{n-1}^{-1}$ for all $n \in \mathbb{N}$.

If $c \leqslant \varepsilon_0$ we use theorem 3.6 to get $\mathcal{R}_0(H) \in \mathbb{I}_0^+ \mathcal{A}_{\rho_0,r}$ with

$$\|H_0 - H^0\|_{\rho_0,r} \leqslant K\Theta_0 \|H - H^0\|_{\rho,r'}$$

for some $K > 0$. Take $Q_0 = Q$.

Now, for $n \in \mathbb{N}$ assume that $H_{n-1} \in \mathbb{I}_{n-1}^+ \mathcal{A}_{\rho_{n-1},r}$. Suppose that

$$\begin{aligned}
\|H_{n-1} - H_{n-1}^0\|_{\rho_{n-1},r} &\leqslant K\Theta_{n-1} \|H - H^0\|_{\rho,r'}, \\
\|Q_{n-1}\| &\leqslant \|Q\| \prod_{i=1}^{n-1} \frac{3}{2} \mu_i |\eta_i| \, \|T^{(i)}\| \, \| {}^\top T^{(i)} \|, \\
\|Q_{n-1}^{-1}\| &\leqslant \|Q^{-1}\| \prod_{i=1}^{n-1} 2\mu_i^{-1} |\eta_i|^{-1} \|T^{(i)^{-1}}\| \, \| {}^\top T^{(i)^{-1}} \|.
\end{aligned} \tag{3.61}$$

So, for c small enough, using the last term in (3.54) we get

$$\|Q_{n-1}^{-1}\| \ll \frac{\Theta_{n-1}^{1/2}}{\Theta_{n-1}} \prod_{i=1}^{n-1} 2^5 \zeta_i^{1/2} \|T^{(i)^{-1}}\| \, \| {}^{\top}T^{(i)^{-1}}\| \max\left\{\frac{1}{|\eta_i|^{1/2}}, \frac{1}{|\eta_i|}\right\} \leqslant \frac{r^2}{32cK\Theta_{n-1}}. \tag{3.62}$$

Thus, lemma 3.2 is valid and as a consequence $\|b_n(H_{n-1})\| < r/8$.

After performing the operators $\mathcal{L}_n$ and $\mathcal{I}_n$, we want to estimate the norm of the resulting Hamiltonians. The constant and non-constant Fourier modes are dealt with separately in

$$\mathcal{I}_n\mathcal{L}_n(H) = H_n^0 + \widehat{\mathcal{L}}_n(\mathbb{E}H_{n-1}) + \mathcal{I}_n\widetilde{\mathcal{L}}_n(\mathbb{I}-\mathbb{E})(H_{n-1}). \tag{3.63}$$

For the former we use lemma 3.3 and for the latter lemmas 3.4 and 3.5. So,

$$\begin{aligned} \|\widehat{\mathcal{L}}_n(\mathbb{E}H_{n-1})\|'_{r'} &\leqslant 2^7 K \left(1+\tfrac{1}{2r'}\right)\left(\tfrac{r'}{r}\right)^3 \mu_n^2 |\eta_n| \, \| {}^{\top}T^{(n)}\|^3 \Theta_{n-1} \|H-H^0\|_{\rho,r'} \\ &\leqslant \tfrac{K}{2\varphi_n} \Theta_n \|H-H^0\|_{\rho,r'}. \end{aligned} \tag{3.64}$$

Furthermore, ϕ_n yields

$$\begin{aligned} \|\mathcal{I}_n\widetilde{\mathcal{L}}_n(\mathbb{I}-\mathbb{E})(H_{n-1})\|'_{\rho_n'',r'} &\leqslant K\left(1+\tfrac{2\pi}{\delta}+\tfrac{r}{2r'^2\log 2}\right)\mu_n^{-1}\phi_n^{-1}|\eta_n|\Theta_{n-1}\|H-H^0\|_{\rho,r'} \\ &\leqslant \tfrac{K}{2\varphi_n}\Theta_n\|H-H^0\|_{\rho,r'}. \end{aligned} \tag{3.65}$$

Moreover, assuming c to be small enough, we estimate (3.20) using (3.27), $\|Q_{n-1}\|^{-1} \leqslant \|Q_{n-1}^{-1}\|$, (3.62) and the second inequality in (3.61) to obtain

$$\begin{aligned} \|Q_n\| &\leqslant \mu_n|\eta_n| \, \|T^{(n)}\| \, \| {}^{\top}T^{(n)}\| \|Q_{n-1}\|(1+16r^{-2}cK\Theta_{n-1}\|Q_{n-1}\|^{-1}) \\ &\leqslant \|Q\| \prod_{i=1}^{n} \frac{3}{2}\mu_i|\eta_i| \, \|T^{(i)}\| \, \| {}^{\top}T^{(i)}\| \leqslant \frac{\sigma_n}{4r'}, \end{aligned} \tag{3.66}$$

where the last inequality comes from the second term in (3.54). By (3.26) and again (3.62),

$$\begin{aligned} \|Q_n^{-1}\| &\leqslant \frac{\|T^{(n)^{-1}}\| \, \| {}^{\top}T^{(n)^{-1}}\| \|Q_{n-1}^{-1}\|}{\mu_n|\eta_n|(1-16r^{-2}cK\Theta_{n-1}\|Q_{n-1}^{-1}\|)} \\ &\leqslant \|Q^{-1}\| \prod_{i=1}^{n} 2\mu_i^{-1}|\eta_i|^{-1}\|T^{(i)^{-1}}\| \, \| {}^{\top}T^{(i)^{-1}}\|. \end{aligned} \tag{3.67}$$

The Hamiltonian $\mathcal{I}_n\mathcal{L}_n(H_{n-1})$ is inside the domain of $\mathcal{U}_n$ since for c small enough $\varphi_n^{-1}cK\Theta_n < \varepsilon_n$ and $\|Q_n\| < \sigma_n/(2r')$. The result follows from (3.49). □

Remark 3.9. The above can be generalized for a small analyticity radius ρ by considering a sufficiently large N and applying the above theorem to $\widetilde{H} = \mathcal{U}_N\mathcal{L}_N\cdots\mathcal{U}_1\mathcal{L}_1\mathcal{U}_0(H)$, where H is close enough to H^0. We recover the large strip case since ρ_N is of the order of B_{N-1}^{-1}. It remains to check that $\rho_N > \mathcal{B}(\omega^{(N)}) + \nu$. This follows from the fact that $\mathcal{B}(\omega^{(N)}) = B_{N-1}^{-1}[\mathcal{B}(\omega) - \mathcal{B}_N(\omega)]$ where $\mathcal{B}_N(\omega)$ is the sum of the first N terms of $\mathcal{B}(\omega)$ so that $\mathcal{B}_N(\omega) \to \mathcal{B}(\omega)$ as $N \to +\infty$.

Lemma 3.10. *If* $\omega = \begin{pmatrix}\alpha \\ 1\end{pmatrix} \in \mathbb{R}^d$ *is diophantine, then (3.59) is verified.*

Proof. To show (3.59) it is only necessary to check that we can find sequences σ_n and t_n for which the series $\sum B_n|\log|\eta_{n+1}||$, $\sum B_n \log\|T^{(n+1)}\|$, $\sum B_n \log\|T^{(n+1)^{-1}}\|$, $\sum B_n \log\|\omega^{(n+1)}\|$, $\sum B_n|\log\sigma_{n+1}|$ and $\sum B_n|\log\Theta_{n+1}|$ converge.

Let us set, for each $n \in \mathbb{N}$,

$$t_n = (1+\xi)^n \qquad \text{and} \qquad \sigma_n = \mathrm{e}^{-a\delta t_n},$$

where positive constants ξ and a will be chosen in the following and $\delta t_n = t_n - t_{n+1}$. We shall assume that ξ is large enough that

$$\xi\left(d - \frac{a}{1+\xi}\right) \geqslant \log\frac{c_1 c_2}{c_6}. \tag{3.68}$$

So, $\sigma_{n-1}\exp(d\delta t_n) \geqslant c_1c_2/c_6$ as in proposition 2.3. Hence,

$$A_{n-1} \ll \mathrm{e}^{-\alpha\delta t_n}, \tag{3.69}$$

with

$$\alpha = \frac{d(\xi-\beta)}{\xi(d+\beta)} - \left(d - \frac{a}{1+\xi}\right) \tag{3.70}$$

which is positive if $a > d(1+\xi)[1-1/(d+\beta)]$. Thus,

$$B_n = \prod_{i=0}^{n} A_i \ll C^n \mathrm{e}^{-\alpha t_{n+1}}, \tag{3.71}$$

where C is some positive constant. Clearly, $\sum B_n < \infty$.

From (2.8) we have $\|\omega^{(n)}\| \leqslant \|M^{(n)}\|\,|\gamma^{(n)}|^{-1}$. Thus, using (2.13) and (2.19) we have

$$\|\omega^{(n)}\| \ll \exp\left[\frac{d\beta(1+\xi)}{\xi}\delta t_n\right]. \tag{3.72}$$

Now, using (2.9) and the bounds (2.19), (2.17) and (2.18) we get

$$\begin{aligned}
\|T^{(n)}\| &\ll \exp\left[\frac{d(1+\xi)(\beta+1)}{\xi(d+\beta)}\delta t_n\right],\\
\|T^{(n)^{-1}}\| &\ll \exp\left[\frac{d(1+\xi)(d-1+\beta)}{\xi(d+\beta)}\delta t_n\right],\\
|\eta_n| &\ll \exp\left[\frac{d(1+\xi)}{\xi}\left(\frac{d-1}{d+\beta}+\beta\right)\delta t_n\right].
\end{aligned} \tag{3.73}$$

Finally,

$$\begin{aligned}
\log\prod_{i=1}^{n}\|T^{(i)}\|,\ \log\prod_{i=1}^{n}\|T^{(i)^{-1}}\|,\ |\log\prod_{i=1}^{n}|\eta_i|| &\ll t_n,\\
\log\prod_{i=1}^{n}\|\omega^{(i)}\|,\ |\log\prod_{i=1}^{n}\sigma_i|,\ |\log\prod_{i=1}^{n}B_{i-1}| &\ll t_n,
\end{aligned} \tag{3.74}$$

so that $|\log\Theta_n| \ll t_n$.

Since B_n decays exponentially with t_{n+1} and $\log\phi_{n+1}$ grows at most linearly, the series (3.59) converges. □

4. Construction of the invariant torus

Here we will always assume to be in the conditions of section 3.6. We use theorem 3.8 to determine the existence of an ω-invariant torus for the flow of analytic Hamiltonians H close enough to H^0 (theorem 1.1). This follows from the construction of an analytic conjugacy between the linear flow on $\mathbb{T}^d$ of rotation vector ω and an orbit of H.

Let the set Δ be given by

$$\Delta = \{H \in \mathcal{A}_{\rho,r'} \colon \|H - H^0\|_{\rho,r'} < c\}, \tag{4.1}$$

which is contained in the domain of $\mathcal{R}_n$ for all $n \in \mathbb{N} \cup \{0\}$. Given $H \in \Delta$, $H_n \in \mathbb{I}_n^+ \mathcal{A}_{\rho_n, r}$. It is simple to check that

$$\begin{aligned} H_n = \frac{\lambda_n}{\chi_n}[(\mathbb{I} - \mathbb{E}_0)(H \circ g_0 \circ L_1^{\mu_1} \circ g_1 \circ \cdots \circ L_n^{\mu_n})] \circ g_n \\ = \frac{\lambda_n}{\chi_n} \{(\mathbb{I} - \mathbb{E}_0) H \circ g_0 \circ [\mathcal{P}_1(H) \circ g_1 \circ \mathcal{P}_1(H)^{-1}] \circ \\ \cdots \circ [\mathcal{P}_{n-1}(H) \circ g_{n-1} \circ \mathcal{P}_{n-1}(H)^{-1}] \circ \mathcal{P}_n(H)\} \circ g_n. \end{aligned} \tag{4.2}$$

Here, $g_k = \mathfrak{G}_k(\mathcal{L}_k(H_{k-1}))$ is given by theorem 3.6 at the kth step and

$$L_k^{\mu_k} : (\boldsymbol{x}, \boldsymbol{y}) \mapsto (T^{(k)^{-1}} \boldsymbol{x}, \Phi_k(H_{k-1})(\boldsymbol{y})), \tag{4.3}$$

where $\Phi_k(H_{k-1})(\boldsymbol{y}) = \mu_k {}^\top T^{(k)} \boldsymbol{y} + b_k(H_{k-1})$. In addition, we have the conformally symplectic map

$$\mathcal{P}_n(H) = L_1^{\mu_1} \cdots L_n^{\mu_n} : (\boldsymbol{x}, \boldsymbol{y}) \mapsto \left(P^{(n)^{-1}} \boldsymbol{x}, \Phi_1(H) \cdots \Phi_n(H_{n-1})(\boldsymbol{y}) \right), \qquad n \geqslant 1, \tag{4.4}$$

and we set $\mathcal{P}_0(H) = \mathrm{Id}$. Notice that

$$\Phi_1(H) \cdots \Phi_n(H_{n-1})(\boldsymbol{y}) = \chi_n {}^\top P^{(n)} \boldsymbol{y} + v_n(H), \tag{4.5}$$

with

$$v_n(H) = b_1(H) + \sum_{i=2}^{n} \chi_{i-1} {}^\top P^{(i-1)} b_i(H_{i-1}).$$

For $n \geqslant 1$ define

$$\begin{aligned} a_n(H) = \lim_{m \to +\infty} \Phi_n(H_{n-1}) \cdots \Phi_m(H_{m-1})(0) \\ = b_n(H_{n-1}) + \sum_{i=n+1}^{+\infty} \mu_n \cdots \mu_{i-1} {}^\top T^{(n)} \cdots {}^\top T^{(i-1)} b_i(H_{i-1}) \end{aligned} \tag{4.6}$$

if it converges. If that is the case,

$$a(H) = a_1(H) = \lim_{n \to +\infty} v_n(H) \tag{4.7}$$

and

$$a(H) - v_n(H) = \chi_n {}^\top P^{(n)} a_{n+1}(H). \tag{4.8}$$

Lemma 4.1. *The maps $a_n : \Delta \to B_{r/2}$ are well defined and analytic, taking any real-analytic H into $\mathbb{R}^d$.*

Proof. From lemma 3.2 we obtain $\|b_k(H_{k-1})\| < r/8$ for any $k \in \mathbb{N}$. Thus, by (3.57),

$$\mu_n \cdots \mu_{i-1} \| {}^\top T^{(n)} \cdots {}^\top T^{(i-1)} b_i(H_{i-1}) \| \leqslant \frac{r}{8} \left(\frac{r}{8r'} \right)^{i-n}, \tag{4.9}$$

where $1 \leqslant n \leqslant i - 1$. Hence, (4.6) converges and each $a_n(H)$ is well defined in $\mathbb{C}^d$. In case H is real, $a_n(H) \in \mathbb{R}^d$. The maps $H \mapsto a_n(H)$ are analytic since the convergence is uniform. Moreover, (4.6) can be estimated using (4.9),

$$\|a_n(H)\| \leqslant \frac{r}{8} + \frac{r}{8} \frac{\frac{r}{8r'}}{1 - \frac{r}{8r'}} < \frac{r}{2}.$$

□

Lemma 4.2. *There is an open ball B centred at H^0 in Δ such that, if $H \in B$, we can find sequences $R_n, r_n > 0$ satisfying: $R_{-1} = \rho$, $r_{-1} = r'$,*

$$R_n + 2\pi K \Theta_n^{2/3} \|H - H^0\|_{\rho,r'} \leqslant R_{n-1} \leqslant \frac{\rho_{n-1}}{\|P^{(n-1)}\|}, \tag{4.10}$$

$$r_n + K \Theta_n^{2/3} \|H - H^0\|_{\rho,r'} \leqslant r_{n-1} \leqslant \frac{\chi_{n-1} r}{2\| {}^{\top}P^{(n-1)^{-1}}\|}, \tag{4.11}$$

$n \geqslant 0$, and

$$\lim_{n\to+\infty} R_n^{-1} \Theta_n^{2/3} = 0. \tag{4.12}$$

Proof. Let $\rho_* = \min \rho_n$. Since χ_n is decreasing and $\|P^{(n)}\| \leqslant \prod_{i=1}^n \|T^{(n)}\|$ (similar relations hold for the transpose and inverse matrices), it is enough to check (using the last term in (3.54)) that

$$\Theta_n^{2/3} \ll \min \left\{ \lambda^n \rho_* \prod_{i=1}^n \|T^{(i)}\|^{-1}, \chi_n \prod_{i=1}^n \| {}^{\top}T^{(i)^{-1}}\|^{-1} \right\}$$

for some $0 < \lambda < 1$ by taking $R_n = c_1 \lambda^{-n} \Theta_n^{2/3}$ and $r_n = c_2 \Theta_n^{2/3}$ with small constants $c_1, c_2 > 0$. Thus, the inequalities (4.10) and (4.11) hold whenever we take a sufficiently small bound on $\|H - H^0\|_{\rho,r}$. The limit (4.12) is now immediate. □

Let the vertical translation

$$V_{\boldsymbol{z}} \colon (\boldsymbol{x}, \boldsymbol{y}) \mapsto (\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}), \tag{4.13}$$

for any $\boldsymbol{z} \in \mathbb{C}^d$. For a given $H \in \Delta$, define the norm $\|X\|_n = \|X \circ V_{a(H)}\|_{R_n,r_n}$ whenever $X \circ V_{a(H)} \in \mathcal{A}^{2d}_{R_n,r_n}$.

Now, consider the isotopic to the identity analytic symplectomorphism

$$W_n(H) = \mathcal{P}_n(H) \circ \mathfrak{G}_n(\mathcal{L}_n(H_{n-1})) \circ \mathcal{P}_n(H)^{-1} \tag{4.14}$$

on $\mathcal{P}_n(H)\mathcal{D}_{\rho_n,r}$ with $n \geqslant 0$ and $H \in \Delta$. In particular, $W_n(H^0) = \mathrm{Id}$. Notice that for H real-analytic, $W_n(H)$ is real-analytic.

Lemma 4.3. *W_n is an analytic map on B such that, if $H \in B$,*

$$W_n(H) \colon V_{a(H)}(\mathcal{D}_{\rho_n,r_n}) \to V_{a(H)}(\mathcal{D}_{\rho_{n-1},r_{n-1}})$$

and there is $K' > 0$ verifying

$$\|W_n(H) - \mathrm{Id}\|_n \leqslant K' \Theta_n^{2/3} \|H - H^0\|_{\rho,r'}. \tag{4.15}$$

Proof. For $H \in \Delta$ and $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{D}_{R_n,r_n}$,

$\|\mathrm{Im}\, P^{(n)}\boldsymbol{x}\| < \|P^{(n)}\| R_n/2\pi \leqslant \rho_n/2\pi$,

$$\begin{aligned} \|\Phi_n^{-1}(H_{n-1}) \cdots \Phi_1^{-1}(H)(\boldsymbol{y} + a(H))\| &= \|\chi_n^{-1}\, {}^{\top}P^{(n)^{-1}}(\boldsymbol{y} + a(H) - v_n(H))\| \\ &\leqslant \chi_n^{-1} \| {}^{\top}P^{(n)^{-1}}\| r_n + \|a_{n+1}(H)\| < r. \end{aligned} \tag{4.16}$$

Therefore, $\mathcal{P}_n(H)^{-1} \circ V_{a(H)}(\mathcal{D}_{R_n,r_n}) \subset \mathcal{D}_{\rho_n,r}$. Moreover, using (3.49),

$$\begin{aligned} \|W_n(H) - \mathrm{Id}\|_n &= \|\widehat{\mathcal{P}}_n(H) \circ [\mathfrak{G}_n(\mathcal{I}_n\mathcal{L}_n(H_{n-1})) - \mathrm{Id}] \circ \mathcal{P}_n(H)^{-1} \circ V_{a(H)}\|_{R_n,r_n} \\ &\leqslant \varepsilon_n^{-1} \|\widehat{\mathcal{P}}_n(H)\| \, \|\mathcal{I}_n\mathcal{L}_n(H_{n-1}) - H_n^0\|_{\rho_n,r'} \\ &\leqslant K' \Theta_n^{2/3} \|H - H^0\|_{\rho,r'}, \end{aligned} \tag{4.17}$$

where $\widehat{\mathcal{P}}_n(H)$ corresponds to the linear part $(\boldsymbol{x}, \boldsymbol{y}) \mapsto (P^{(n)^{-1}}\boldsymbol{x}, \chi_n\, {}^{\top}P^{(n)}\boldsymbol{y})$ of $\mathcal{P}_n(H)$ which has norm bounded by $\|\widehat{\mathcal{P}}_n(H)\| \leqslant \|P^{(n)^{-1}}\| + \chi_n \| {}^{\top}P^{(n)}\|$.

Now, for $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{D}_{R_n, r_n}$ and $H \in B$,

$$\begin{aligned}
&\|\pi_1 \mathrm{Im} W_n(H) \circ V_{a(H)}(\boldsymbol{x}, \boldsymbol{y})\| \leqslant \|\mathrm{Im}(\pi_1 W_n(H) \circ V_{a(H)}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{x})\| + \|\mathrm{Im}\boldsymbol{x}\| \\
&\quad < \|W_n(H) - \mathrm{Id}\|_n + R_n/2\pi < R_{n-1}/2\pi, \\
&\|\pi_2 W_n(H) \circ V_{a(H)}(\boldsymbol{x}, \boldsymbol{y}) - a(H)\| \leqslant \|\pi_2 W_n(H) \circ V_{a(H)}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{y} - a(H)\| + \|\boldsymbol{y}\| \\
&\quad < \|W_n(H) - \mathrm{Id}\|_n + r_n < r_{n-1}.
\end{aligned}$$

So, $W_n(H)\colon V_{a(H)}(\mathcal{D}_{R_n, r_n}) \to V_{a(H)}(\mathcal{D}_{R_{n-1}, r_{n-1}})$. □

Define the analytic map Γ_n on B satisfying $\Gamma_n(H)\colon V_{a(H)}(\mathcal{D}_{R_n, r_n}) \to V_{a(H)}(\mathcal{D}_{\rho, r'})$,

$$\Gamma_n(H) = W_0(H) \circ \cdots \circ W_n(H), \tag{4.18}$$

with $H \in B$. We then rewrite (4.2) as

$$H \circ \Gamma_n(H) = \frac{\chi_n}{\lambda_n} H_n \circ \mathcal{P}_n(H)^{-1} + E(H), \tag{4.19}$$

where $E(H)$ represents a constant (irrelevant) term. Since each $W_n(H)$ is symplectic, thus $\Gamma_n(H)$ is symplectic and $H \circ \Gamma_n(H)$ is canonically equivalent to the Hamiltonian H_n. In particular, if $H_n = H_n^0$ for some n, there is an $\boldsymbol{\omega}$-invariant torus in the phase space of H_n. We are interested in the general case, $H_n - H_n^0 \to 0$ as $n \to +\infty$.

Lemma 4.4. *There is $c > 0$ such that for $H \in B$*

$$\|\Gamma_n(H) - \Gamma_{n-1}(H)\|_n \leqslant c\Theta_n^{2/3} \|H - H^0\|_{\rho, r'}.$$

Proof. For each $k = 0, \cdots, n-1$, consider the transformations

$$\begin{aligned}
G_k(z, H) &= (W_k(H) - \mathrm{Id}) \circ (\mathrm{Id} + G_{k+1}(z, H)) + G_{k+1}(z, H), \\
G_n(z, H) &= z(W_n(H) - \mathrm{Id}),
\end{aligned}$$

with $(z, H) \in \{z \in \mathbb{C}\colon |z| < 1 + d_n\} \times B$, where we have $c' > 0$ such that

$$d_n = \frac{c'}{\Theta_n^{2/3} \|H - H^0\|_{\rho, r'}} - 1 > 0.$$

If $\|G_{k+1}(z, H)\|_n \leqslant (R_k - R_n)/2\pi$, then G_k is well defined as an analytic map and

$$\|G_k(z, H)\|_n \leqslant \|W_k(H) - \mathrm{Id}\|_k + \|G_{k+1}(z, H)\|_n.$$

An inductive scheme shows that

$$\begin{aligned}
&\|G_n(z, H)\|_n \leqslant (R_{n-1} - R_n)/2\pi, \\
&\|G_k(z, H)\|_n \leqslant \sum_{i=k}^{n-1} \|W_i(H) - \mathrm{Id}\|_i + |z| \, \|W_n(H) - \mathrm{Id}\|_n \\
&\qquad \leqslant (R_{k-1} - R_n)/2\pi.
\end{aligned}$$

By Cauchy's formula

$$\begin{aligned}
\|\Gamma_n(H) - \Gamma_{n-1}(H)\|_n &= \|G_0(1, H) - G_0(0, H)\|_n \\
&= \left\| \frac{1}{2\pi i} \oint_{|z| = 1 + d_n/2} \frac{G_0(z, H)}{z(z-1)} \mathrm{d}z \right\|_n
\end{aligned}$$

and

$$\begin{aligned}\|\Gamma_n(H) - \Gamma_{n-1}(H)\|_n &\leqslant \frac{2}{d_n} \sup_{|z|=1+d_n/2} \|G_0(z, H)\|_n \\ &\ll \Theta_n^{2/3} \|H - H^0\|_{\rho,r'}.\end{aligned}$$

□

Consider the Banach space $C^1_{\text{per}}(\mathbb{R}^d, \mathbb{C}^{2d})$ of C^1 functions $\mathbb{Z}^d$-periodic, endowed with the norm

$$\|f\|_{C^1} = \max_{k\leqslant 1} \max_{\boldsymbol{x}\in\mathbb{R}^d} \|D^k f(\boldsymbol{x})\|.$$

Our goal is to find parametrizations of invariant tori of the type $\boldsymbol{\theta} \mapsto (\boldsymbol{\theta}, a(H)) + f(\boldsymbol{\theta})$.

Lemma 4.5. *There exist $C > 0$, an open ball $B' \subset B$ centred at H^0 and an analytic map Υ on B' such that, for every $H \in B'$, $\Upsilon(H) = \lim_{n\to+\infty} \Gamma_n(H)|_{\{\boldsymbol{y}=a(H)\}}$ is an embedding $\mathbb{R}^d \to \mathbb{C}^{2d}$, $\Upsilon(H) - (\mathrm{Id}, a(H)) \in C^1_{\text{per}}(\mathbb{R}^d, \mathbb{C}^{2d})$ and*

$$\|\Upsilon(H) - (\mathrm{Id}, a(H))\|_{C^1} \leqslant C\|H - H^0\|_{\rho,r'}. \tag{4.20}$$

If $H \in B'$ is real-analytic, then $\Upsilon(H): \mathbb{R}^d \to \mathbb{R}^{2d}$.

Proof. For each $H \in B$, by the first inequality in (3.5),

$$\begin{aligned}&\|[\Gamma_n(H) - \Gamma_{n-1}(H)](\cdot, a(H))\|_{C^1} \\ &\quad\leqslant \max_{k\leqslant 1} \sup_{\boldsymbol{x}\in D_{\rho_n/2}} \|D^k[\Gamma_n(H)(\boldsymbol{x}, a(H)) - \Gamma_{n-1}(H)(\boldsymbol{x}, a(H))]\| \\ &\quad\leqslant \frac{4\pi}{R_n}\|\Gamma_n(H) - \Gamma_{n-1}(H)\|_n,\end{aligned} \tag{4.21}$$

which is estimated using (4.12). Hence, $\Gamma_n(H)(\cdot, a(H)) - (\mathrm{Id}, a(H))$ converges in the Banach space $C^1_{\text{per}}(\mathbb{R}^d, \mathbb{C}^{2d})$, and (4.20) holds. The convergence of Γ_n is uniform in B; thus Υ is analytic. If H is sufficiently close to H^0, $\Upsilon(H)$ is in fact an injective immersion (embedding) as the space of embeddings is closed for the C^1 norm and $\Upsilon(H)$ is close to $(\mathrm{Id}, a(H))$. Finally, for H real-analytic we have $\Upsilon(H)(\mathbb{R}^d) \subset \mathbb{R}^{2d}$ in view of the similar property for each $W_n(H)$. □

The Hamiltonian vector field of a Hamiltonian H is $X_H = \mathbb{J}\nabla H$, where $\mathbb{J}: (x, y) \mapsto (y, -x)$. The next lemma shows the invariance of the torus defined by $\Upsilon(H)$ which corresponds to the linear vector field $\dot{\boldsymbol{\theta}} = \boldsymbol{\omega}$.

Lemma 4.6. *For $H \in B'$, we have on $\mathbb{R}^d$*

$$X_H \circ \Upsilon(H) = D(\Upsilon(H))\,\boldsymbol{\omega}. \tag{4.22}$$

Proof. Since $\Gamma_n(H)$ is a symplectomorphism, we have for $\boldsymbol{x} \in \mathbb{R}^d$,

$$\begin{aligned}Y_n(\boldsymbol{x}) = X_H \circ \Gamma_n(H) \circ V_{a(H)}(\boldsymbol{x}, 0) - D(\Gamma_n(H)) \circ V_{a(H)}(\boldsymbol{x}, 0)\, X_{H^0}(\boldsymbol{x}, 0) \\ = [D(\Gamma_n(H)) \circ V_{a(H)}\, X_{H\circ\Gamma_n(H)\circ V_{a(H)}-H^0}](\boldsymbol{x}, 0).\end{aligned} \tag{4.23}$$

Hence,

$$\|Y_n(\boldsymbol{x})\| \leqslant \|D(\Gamma_n(H))(\boldsymbol{x}, a(H))\|\, \|\nabla[H \circ \Gamma_n(H) \circ V_{a(H)} - H^0](\boldsymbol{x}, 0)\|. \tag{4.24}$$

In order to estimate the above we first recall (4.19) to show that

$$\nabla[H \circ \Gamma_n(H) \circ V_{a(H)} - H^0](\boldsymbol{x}, 0) = \frac{\chi_n}{\lambda_n} \nabla[(H_n - H_n^0) \circ \mathcal{P}_n(H)^{-1} \circ V_{a(H)}](\boldsymbol{x}, 0)$$
$$+ \frac{1}{\lambda_n \chi_n} {}^{\top}[a(H) - v_n(H)] P^{(n)^{-1}} Q_n {}^{\top}P^{(n)^{-1}}. \tag{4.25}$$

Notice that by induction we get

$$\frac{1}{\lambda_n \chi_n} P^{(n)^{-1}} Q_n {}^{\top}P^{(n)^{-1}} = Q + \sum_{i=0}^{n-1} \frac{1}{\lambda_i \chi_i} P^{(i)^{-1}} D^2 F_0^{(i)}(b_{i+1}(H_i)) {}^{\top}P^{(i)^{-1}}. \tag{4.26}$$

Since $\sum_{i=1}^{n-1} (\chi_i |\lambda_i|)^{-1} \|P^{(i)^{-1}}\| \, \|{}^{\top}P^{(i)^{-1}}\| \Theta_i \ll 1$ and by (4.6) and (4.8)

$$\|a(H) - v_n(H)\| \leqslant \chi_n \|{}^{\top}P^{(n)}\| \, \|a_{n+1}(H)\| \ll \Theta_n^{2/3}, \tag{4.27}$$

the last term in (4.25) is estimated from above by $\Theta_n^{2/3}$. Moreover, the first term in the rhs of (4.25) is bounded times a constant by

$$\frac{1}{|\lambda_n|} \|{}^{\top}P^{(n)^{-1}}\| \, \|H_n - H_n^0\|_{\rho_n, r} \ll \Theta_n^{2/3}. \tag{4.28}$$

Finally, from the convergence of Γ_n and

$$\|D\Gamma_n(H)(\boldsymbol{x}, a(H))\| \ll \frac{1}{R_n} \|\Gamma_n(H)\|_n \ll \frac{1}{R_n}, \tag{4.29}$$

we find that $\|Y_n(\boldsymbol{x})\|$ converges uniformly to 0 as $n \to +\infty$ because of (4.12). □

Lemma 4.7. *If $H \in B'$ and $\boldsymbol{x} \in \mathbb{R}^d$, then*

$$\Upsilon(H \circ R_{\boldsymbol{x}}) = R_{\boldsymbol{x}}^{-1} \circ \Upsilon(H) \circ \widehat{R}_{\boldsymbol{x}}, \tag{4.30}$$

where $\widehat{R}_{\boldsymbol{x}} : z \mapsto z + \boldsymbol{x}$ is a translation on $\mathbb{C}^d$.

Proof. For each $n \in \mathbb{N}$, (3.12) implies that $\mathcal{P}_n(H \circ R_z) = \mathcal{P}_n(H)$ and we know that $\mathcal{P}_n(H) \circ R_{P^{(n)}z}^{-1} = R_z^{-1} \circ \mathcal{P}_n(H), z \in \mathbb{C}^d$. So, from lemma 3.7,

$$W_n(H \circ R_{\boldsymbol{x}}) = \mathcal{P}_n(H) \circ \mathfrak{G}_n(\mathcal{L}_n \mathcal{R}_{n-1}(H \circ R_{\boldsymbol{x}})) \circ \mathcal{P}_n(H)^{-1}$$
$$= R_{\boldsymbol{x}}^{-1} \circ W_n(H) \circ R_{\boldsymbol{x}}. \tag{4.31}$$

Thus, $\Gamma_n(H \circ R_{\boldsymbol{x}}) = R_{\boldsymbol{x}}^{-1} \circ \Gamma_n(H) \circ R_{\boldsymbol{x}}$ and (4.30) follows using the convergence of Γ_n. □

The flow generated by X_H is denoted by ϕ_H^t taken at time $t \geqslant 0$. Hence,

$$\phi_{H^0}^t|_{\mathbb{T}^d \times \{0\}} = \widehat{R}_{\omega t}.$$

We prove below the existence of an invariant torus $\mathcal{T}$ for H close to H^0, i.e. an analytic conjugacy between $\phi_H^t|_{\mathcal{T}}$ and $\widehat{R}_{\omega t}$.

Theorem 4.8. *Let $D \subset \mathbb{R}^d$ be an open ball about the origin. If $H \in C^\omega(\mathbb{T}^d \times D)$ is sufficiently close to H^0, then there exists a C^ω embedding $\gamma : \mathbb{T}^d \to \mathbb{T}^d \times D$ such that*

$$\phi_H^t \circ \gamma = \gamma \circ \widehat{R}_{\omega t}, \qquad t \geqslant 0, \tag{4.32}$$

and $\mathcal{T} = \gamma(\mathbb{T}^d) \simeq \mathbb{T}^d$ is a submanifold homotopic to $\{\boldsymbol{y} = 0\}$. Furthermore, the map $H \mapsto \gamma$ is analytic.

Proof. The lift $\widetilde{H}$ to $\mathbb{R}^d \times D$ of H is assumed to have a unique analytic extension to $\mathcal{D}_{\rho,r'}$. Consider the real-analytic Hamiltonian $G = \widetilde{H} \in \mathcal{A}_{\rho,r'}$. Suppose that G is close enough to H^0 such that $G \in B'$ and $G \circ R_{\boldsymbol{z}} \in B'$ for $\eta > 0$ and $\boldsymbol{z} \in D_\eta$. Then, $\gamma = \Upsilon(G)|_{[0,1)^d}$, which is C^1 and homotopic to $(\mathrm{Id}, a(G))$, verifies (4.32). This follows from (4.22) and the equivalent equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}(\phi_H^t \circ \gamma) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}(\gamma \circ \widehat{R}_{\omega t}),$$

which we integrate for initial condition $\phi_H^0 = \widehat{R}_0 = \mathrm{Id}$.

We now want to extend analytically γ to a complex neighbourhood of its domain. Take $\widetilde{\gamma}(z) = R_z \circ \Upsilon(G \circ R_z)(0)$, $z \in D_\eta$. The maps $z \mapsto G \circ R_z$ and $H \mapsto \Upsilon(H)$ are analytic and $C^1_{\mathrm{per}}(\mathbb{R}^d, \mathbb{C}^{2d}) \ni g \mapsto g(0)$ is bounded. As $\widetilde{\gamma} \colon D_\eta \to \mathbb{C}^{2d}$ involves their composition, it is analytic and $\mathbb{Z}^d$-periodic. From (4.30), for any $\boldsymbol{x} \in \mathbb{R}^d$, we have

$$\widetilde{\gamma}(\boldsymbol{x}) = \Upsilon(G) \circ \widehat{R}_{\boldsymbol{x}}(0) = \Upsilon(G)(\boldsymbol{x}) = \gamma(\boldsymbol{x}).$$

Finally, since Υ is analytic, the same is true for the map $H \mapsto \gamma$. □

As a quasiperiodic invariant torus $\mathcal{T}$ is always Lagrangian (cf [4]), we have now concluded the proof of theorem 1.1.

5. Elimination of modes

Here we present a proof of theorem 3.6. It is similar to the related methods appearing in, e.g. [7, 1]. As we have fixed n, we will not include it in our notations.

Let $R = (R_1, R_2)$ and $R' = (R_1', R_2')$ be such that $R > R' > 0$ componentwise. We will be interested in the set $\mathcal{G}_{R'}$ of analytic symplectomorphisms $g \colon \mathcal{D}_{R'} \to \mathcal{D}_R$ satisfying $g - \mathrm{Id} \in \mathcal{A}_{R'}^{2d}$ and

$$\|g - \mathrm{Id}\|_{R'} < \delta = \min\{(R_1 - R_1')/2\pi, R_2 - R_2'\}.$$

We use the notation $\{\cdot,\cdot\}$ for the usual Poisson bracket associated to $\mathbb{J} \colon (\boldsymbol{x}, \boldsymbol{y}) \mapsto (\boldsymbol{y}, -\boldsymbol{x})$. In the following $R - \delta$ stands for $R - \delta(1, 1)$ and $\pi_2 \colon (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{y}$ is the projection on the second component. The lemma below constructs a symplectomorphism g generated by a function G and gives several related estimates to be used later.

Lemma 5.1. *Let* $0 < \xi \leqslant \frac{1}{2}$. *If* $G \in \mathcal{A}_{R'}'$ *and* $\|G\|_{R'}' < \xi\delta/(2\pi + 1)$, *then there is a unique analytic symplectomorphism* $g \colon \mathcal{D}_{R'-2\delta} \to \mathbb{C}^{2d}$ *such that* $\|g - \mathrm{Id}\|_{R'-2\delta} < \xi\delta$ *and*

$$g = \mathrm{Id} + \mathbb{J}\nabla G \circ \widehat{g}, \tag{5.1}$$

where $\widehat{g}(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \pi_2 g(\boldsymbol{x}, \boldsymbol{y}))$, $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{D}_{R'-2\delta}$. *Moreover, for any* $H \in \mathcal{A}_{R'}$

$$\begin{aligned} &\|H \circ g\|_{R'-2\delta} \leqslant \|H\|_{R'}, \\ &\|H \circ g - H\|_{R'-2\delta} \leqslant 2\xi \|H\|_{R'}, \\ &\|H \circ g - H - \{H, G\}\|_{R'-2\delta} \leqslant 2\xi^2 \|H\|_{R'}, \end{aligned} \tag{5.2}$$

and the maps $G \mapsto g$ *and* $G \mapsto H \circ g$ *are analytic.*

Proof. Define the map $T \colon g \mapsto \mathrm{Id} + \mathbb{J}\nabla G \circ \widehat{g}$ on the open ball B in $\mathcal{A}_{R'-2\delta}^{2d}$ centred at the identity and with radius $\xi\delta$. It is simple to check that $T(B) \subset B$; in particular, a fixed point $T(g) = g \in B$ is symplectic. We now show that T is a contraction on B and thus its unique fixed point is the map we are looking for. In fact, whenever $g \in B$ we obtain

$$\begin{aligned} \|DT(g)\| &\leqslant \|D\nabla G \circ \widehat{g}\|_{R'-2\delta} \leqslant \|D\nabla G\|_{R'-\delta} \\ &\leqslant \frac{2\pi + 1}{\delta}\|\nabla G\|_{R'} \leqslant \frac{2\pi + 1}{\delta}\|G\|_{R'}' < \xi. \end{aligned} \tag{5.3}$$

For the estimates in (5.2) (the first is now immediate) we introduce the differentiable function

$$f\colon \{z \in \mathbb{C}\colon |z| < \zeta\} \to \mathcal{A}_{R'}$$
$$z \mapsto H \circ (\mathrm{Id} + z\mathbb{J}\nabla G(\mathrm{Id} + z(\widehat{g} - \mathrm{Id}))), \tag{5.4}$$

where $\zeta = 1/\xi \geqslant 2$. Cauchy's integral formula yields that

$$\begin{aligned}\|H \circ g - H\|_{R'-2\delta} &= \|f(1) - f(0)\|_{R'-2\delta} \\ &\leqslant \frac{1}{2\pi}\oint_{|z|=\zeta} \frac{\|f(z)\|_{R'-2\delta}}{|z(z-1)|}\mathrm{d}z \\ &\leqslant \frac{1}{\zeta - 1}\sup_{|z|=\zeta}\|f(z)\|_{R'-2\delta} \leqslant 2\xi\|H\|_{R'}\end{aligned} \tag{5.5}$$

and

$$\begin{aligned}\|H \circ g - H - \{H, G\}\|_{R'-2\delta} &= \|f(1) - f(0) - f'(0)\|_{R'-2\delta} \\ &\leqslant \frac{1}{2\pi}\oint_{|z|=\zeta} \frac{\|f(z)\|_{R'-2\delta}}{|z^2(z-1)|}\mathrm{d}z \\ &\leqslant \frac{1}{\zeta(\zeta - 1)}\sup_{|z|=\zeta}\|f(z)\|_{R'-2\delta} \leqslant 2\xi^2\|H\|_{R'}.\end{aligned} \tag{5.6}$$

By the implicit function theorem the maps $G \mapsto g$ and $G \mapsto H \circ g$ are analytic. □

Lemma 5.2. *Let $\sigma > 2R_2\|Q\|$, $\varepsilon' > 0$ and $H \in \mathcal{A}'_R$ such that*

$$\|H - H^0\|_R < \varepsilon' \leqslant \frac{\sigma\delta}{(2\pi + 1)[1 + 2\pi + (\tau + 1)/R_2]}. \tag{5.7}$$

Then there is $G \in \mathbb{I}^-\mathcal{A}'_{R'}$ such that

$$\mathbb{I}^-(H + \{H, G\}) = 0 \qquad \textit{and} \qquad \|G\|'_{R'} \leqslant \frac{\delta}{(2\pi + 1)\varepsilon'}\|\mathbb{I}^- H\|_{R'}. \tag{5.8}$$

Moreover, the map $H \mapsto G$ is analytic.

Proof. Consider the linear operator associated to H:

$$\mathcal{F}\colon \mathbb{I}^-\mathcal{A}'_{R'} \to \mathbb{I}^-\mathcal{A}_{R'}, \qquad K \mapsto \mathbb{I}^-\{H, K\}. \tag{5.9}$$

It is well defined since

$$\begin{aligned}\|\mathbb{I}^-\{H, K\}\|_{R'} &\leqslant \|\nabla H\|_{R'}\|\nabla K\|_{R'} \\ &\leqslant \|H\|'_{R'}\|K\|'_{R'}.\end{aligned}$$

We will show that $\mathcal{F}^{-1}\colon \mathbb{I}^-\mathcal{A}_{R'} \to \mathbb{I}^-\mathcal{A}'_{R'}$ is bounded and

$$\|\mathcal{F}^{-1}\| < \frac{1}{\frac{\pi R_2\sigma}{(2\pi+1)R_2+\tau+1} - 2\frac{2\pi+1}{\delta}\varepsilon'} \leqslant \frac{\delta}{(2\pi + 1)\varepsilon'}. \tag{5.10}$$

A solution of (5.8) is simply given by $G = \mathcal{F}^{-1}(-\mathbb{I}^- H)$. Therefore, $\|G\|'_{R'} \leqslant \|\mathcal{F}^{-1}\|\,\|\mathbb{I}^- H\|_{R'}$.

We start by decomposing any Hamiltonian $H = H^0 + F$ as

$$H(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{k}} H_{\boldsymbol{k}}(\boldsymbol{y})\,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \qquad \text{with} \quad H_{\boldsymbol{k}}(\boldsymbol{y}) = \sum_{\nu} H_{\boldsymbol{k},\nu}\boldsymbol{y}^{\nu}.$$

Write $D_0 = \nabla_2 H^0 \cdot \nabla_1$, with ∇_1 and ∇_2 standing for the derivatives with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$. The definition of $\mathcal{F}$ in (5.9) yields

$$\mathcal{F}(K) = \mathbb{I}^-(\widehat{F} - D_0)\, K = -\left(\mathbb{I} - \mathbb{I}^- \widehat{F} D_0^{-1}\right) D_0 K,$$

where $\widehat{F}(K) = \{F, K\}$. If the inverse of $\mathcal{F}$ exists it is given by

$$\mathcal{F}^{-1} = -D_0^{-1}\left(\mathbb{I} - \mathbb{I}^- \widehat{F} D_0^{-1}\right)^{-1}. \tag{5.11}$$

The map $D_0^{-1}: \mathbb{I}^-\mathcal{A}_{R'} \to \mathbb{I}^-\mathcal{A}'_{R'}$ is linear and given by

$$D_0^{-1}\, W(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d - \{0\}} \frac{W_{\boldsymbol{k}}(\boldsymbol{y})}{2\pi \mathrm{i}(\boldsymbol{k} \cdot \nabla_2 H^0(\boldsymbol{y}))} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{k} \cdot \boldsymbol{x}}, \qquad W \in \mathbb{I}^-\mathcal{A}_{R'}.$$

For each $\boldsymbol{k} \in I^-$, using (3.33) and $\|Q\| < \sigma/(2R_2)$ thus $|\boldsymbol{k} \cdot Q\boldsymbol{y}/\boldsymbol{k} \cdot \boldsymbol{\omega}| < 1/2$,

$$\frac{W_{\boldsymbol{k}}(\boldsymbol{y})}{\boldsymbol{k} \cdot \boldsymbol{\omega}\left(1 + \frac{\boldsymbol{k} \cdot Q\boldsymbol{y}}{\boldsymbol{k} \cdot \boldsymbol{\omega}}\right)} = \frac{W_{\boldsymbol{k}}(\boldsymbol{y})}{\boldsymbol{k} \cdot \boldsymbol{\omega}} \sum_{n \geqslant 0} \left(-\frac{\boldsymbol{k} \cdot Q\boldsymbol{y}}{\boldsymbol{k} \cdot \boldsymbol{\omega}}\right)^n, \tag{5.12}$$

we get the estimate

$$\begin{aligned}
\left\| \frac{W_{\boldsymbol{k}}}{\boldsymbol{k} \cdot \nabla_2 H^0} \right\|_{R_2} &\leqslant \sum_{n \geqslant 0} \sum_{\nu} \frac{|W_{\boldsymbol{k},\nu}|\, R_2^{\|\nu\|} \|Q\|^n R_2^n}{\sigma^{n+1} \|\boldsymbol{k}\|} \\
&< \sum_{n \geqslant 0} \sum_{\nu} \frac{|W_{\boldsymbol{k},\nu}|\, R_2^{\|\nu\|}}{\sigma \|\boldsymbol{k}\|} \left(\frac{1}{2}\right)^n \\
&= \frac{2}{\sigma \|\boldsymbol{k}\|} \|W_{\boldsymbol{k}}\|_{R_2}.
\end{aligned} \tag{5.13}$$

Similarly, we find the bound

$$\left\| \frac{\nabla_2 W_{\boldsymbol{k}}}{\boldsymbol{k} \cdot \nabla_2 H^0} \right\|_{R_2} \leqslant \sum_{\nu} \frac{2\|\nu\|\, |W_{\boldsymbol{k},\nu}| R_2^{\|\nu\|-1}}{\sigma \|\boldsymbol{k}\|} < \frac{2\tau}{\sigma R_2} \|W_{\boldsymbol{k}}\|_{R_2}. \tag{5.14}$$

Finally,

$$\left\| \frac{W_{\boldsymbol{k}} Q\boldsymbol{k}}{(\boldsymbol{k} \cdot \nabla_2 H^0)^2} \right\|_{R_2} < \frac{2}{\sigma R_2 \|\boldsymbol{k}\|} \|W_{\boldsymbol{k}}\|_{R_2}. \tag{5.15}$$

It is now immediate to see that

$$\|D_0^{-1} W\|_{R'} \leqslant \frac{2}{2\pi\sigma} \|W\|_{R'} \qquad \text{and} \qquad \|\nabla_1 (D_0^{-1} W)\|_{R'} \leqslant \frac{2}{\sigma} \|W\|_{R'}.$$

Moreover,

$$\nabla_2 \left(\frac{W_{\boldsymbol{k}}(\boldsymbol{y})}{\boldsymbol{k} \cdot \nabla_2 H^0(\boldsymbol{y})}\right) = \frac{\nabla_2 W_{\boldsymbol{k}}(\boldsymbol{y})}{\boldsymbol{k} \cdot \nabla_2 H^0(\boldsymbol{y})} - \frac{W_{\boldsymbol{k}}(\boldsymbol{y}) Q\boldsymbol{k}}{(\boldsymbol{k} \cdot \nabla_2 H^0(\boldsymbol{y}))^2},$$

which implies

$$\|\nabla_2 (D_0^{-1} W)\|_{R'} < \frac{\tau + 1}{\pi \sigma R_2} \|W\|_{R'}. \tag{5.16}$$

Hence,

$$\|D_0^{-1}\| < \frac{2}{\sigma}\left(1 + \frac{1}{2\pi} + \frac{\tau + 1}{2\pi R_2}\right).$$

As $\widehat{F}: \mathbb{I}^-\mathcal{A}'_{R'} \to \mathcal{A}_{R'}$ with $\|\widehat{F}\| \leqslant 2\, \|\nabla F\|_{R'} \leqslant 2 \frac{2\pi+1}{\delta} \|F\|_R$ (by Cauchy's estimate),

$$\|\mathbb{I}^- \widehat{F}\, D_0^{-1}\| < \frac{4}{\sigma}\left(1 + \frac{1}{2\pi} + \frac{\tau + 1}{2\pi R_2}\right) \|\nabla F\|_{R'} < 1$$

and

$$\left\| \left(\mathbb{I} - \mathbb{I}^{-} \widehat{F} D_0^{-1} \right)^{-1} \right\| < \left[1 - \frac{4}{\sigma} \left(1 + \frac{1}{2\pi} + \frac{\tau + 1}{2\pi R_2} \right) \|\nabla F\|_{R'} \right]^{-1}.$$

Thus $\mathcal{F}^{-1}$ exists given by (5.11) and the estimate (5.10) on its norm follows immediately. □

Consider the pairs $R = (\rho_n + \nu, r')$ and $R' = (\rho_n, r)$, $\sigma > 2r'\|Q\|$ and $H_0 = H$ as given in theorem 3.6. We are going to iterate the procedure indicated in the previous lemmas. Let a sequence of Hamiltonians be given by

$$H_k = H_{k-1} \circ g_k, \qquad k \in \mathbb{N},$$

where G_k and g_k are determined for H_{k-1} by lemmas 5.2 and 5.1, respectively. In addition, denote by

$$g^{(k)} = g_1 \circ \cdots \circ g_k \tag{5.17}$$

the composition of all symplectomorphisms up to the kth-step so that $H_k = H \circ g^{(k)}$. In order to determine the right domains of H_k, G_k and g_k, define the sequences

$$R_k = R_{k-1} - 4\delta_k = R - 4 \sum_{i=1}^{k} \delta_i, \tag{5.18}$$

with $R_0 = R$ and

$$\delta_k = \frac{1}{2^{k+2}} \min \left\{ 1, \frac{\nu}{2\pi}, r' - r \right\} \leqslant \frac{1}{2^k}. \tag{5.19}$$

So, $\lim_{n \to +\infty} R_k \geqslant R'$ componentwise. From now on, assume that

$$\varepsilon' = \min \left\{ \frac{1}{2} \|H^0\|_R, \frac{\sigma \delta_1}{(2\pi + 1)(1 + 2\pi + \frac{\tau+1}{r'})} \right\}. \tag{5.20}$$

Lemma 5.3. *If for every $k \in \mathbb{N}$, $\|\mathbb{I}^{-} H_{k-1}\|_{R_{k-1}} \leqslant \varepsilon'/2$ and*

$$\|G_k\|'_{R_{k-1}-\delta_k} < \frac{\delta_k}{(2\pi + 1)\varepsilon'} \|\mathbb{I}^{-} H_{k-1}\|_{R_{k-1}},$$

then $g_k(\mathcal{D}_{R_k}) \subset \mathcal{D}_{R_{k-1}}$ and

$$\|g^{(k)} - \mathrm{Id}\|_{R_k} \leqslant \sum_{i=1}^{k} \frac{\delta_i}{\varepsilon'} \|\mathbb{I}^{-} H_{i-1}\|_{R_{i-1}},$$

$$\|g^{(k)} - g^{(k-1)}\|_{R_k} \leqslant \frac{1}{\varepsilon'} \|\mathbb{I}^{-} H_{k-1}\|_{R_{k-1}}. \tag{5.21}$$

Proof. Recall lemma 5.1 for $\xi = \|\mathbb{I}^{-} H_{k-1}\|_{R_{k-1}}/\varepsilon'$ and check that

$$\|g_k - \mathrm{Id}\|_{R_k} \leqslant \|g_k - \mathrm{Id}\|_{R_{k-1}-3\delta_k} < \frac{\delta_k}{\varepsilon'} \|\mathbb{I}^{-} H_{k-1}\|_{R_{k-1}}$$

and $R_k + \delta_k < R_{k-1}$ componentwise. Now,

$$g^{(k)} - \mathrm{Id} = \sum_{i=1}^{k-1} (g_i - \mathrm{Id}) \circ g_{i+1} \circ \cdots \circ g_k + g_k - \mathrm{Id}. \tag{5.22}$$

Thus,

$$\|g^{(k)} - \mathrm{Id}\|_{R_k} \leqslant \sum_{i=1}^{k} \|g_i - \mathrm{Id}\|_{R_i} \leqslant \sum_{i=1}^{k} \frac{\delta_i}{\varepsilon'} \|\mathbb{I}^- H_{i-1}\|_{R_{i-1}}. \tag{5.23}$$

Furthermore, as

$$g^{(k)} - g^{(k-1)} = (g^{(k-1)} - \mathrm{Id}) \circ g_k - (g^{(k-1)} - \mathrm{Id}) + (g_k - \mathrm{Id}), \tag{5.24}$$

we get

$$\begin{aligned} \|g^{(k)} - g^{(k-1)}\|_{R_k} &\leqslant (\|Dg^{(k-1)} - I\|_{R_k} + 1)\, \|g_k - \mathrm{Id}\|_{R_k} \\ &\leqslant \frac{\delta_k}{\varepsilon'} \|\mathbb{I}^- H_{k-1}\|_{R_{k-1}} \left(\frac{2\pi + 1}{4\delta_k} \sum_{i=1}^{k-1} \frac{\delta_i}{\varepsilon'} \|\mathbb{I}^- H_{i-1}\|_{R_{i-1}} + 1 \right) \\ &\leqslant \frac{1}{\varepsilon'} \|\mathbb{I}^- H_{k-1}\|_{R_{k-1}}. \end{aligned} \tag{5.25}$$

□

Notice that since $\varepsilon' \leqslant \frac{1}{2}\|H^0\|_R$, we have

$$\varepsilon' \leqslant \|H^0\|_R - \varepsilon' \leqslant \|H\|_R \leqslant \|H^0\|_R + \varepsilon' \tag{5.26}$$

and also

$$\tfrac{1}{2}\|H^0\|_R \leqslant \|H\|_R \leqslant \tfrac{3}{2}\|H^0\|_R. \tag{5.27}$$

Lemma 5.4. *For any $k \in \mathbb{N}$, if $\|\mathbb{I}^- H\|_R \leqslant \varepsilon'^2/(8\|H\|_R)$, then*

$$\|\mathbb{I}^- H_k\|_{R_k} \leqslant \left(\frac{4\|H\|_R}{\varepsilon'^2} \right)^{2^k - 1} \|\mathbb{I}^- H\|_R^{2^k} \leqslant \frac{\varepsilon'}{2}, \tag{5.28}$$

$$\|H_k - H_{k-1}\|_{R_k} \leqslant \frac{4\|H\|_R}{\varepsilon'} \|\mathbb{I}^- H_{k-1}\|_{R_{k-1}}, \tag{5.29}$$

$$\|H_k\|_{R_k} \leqslant 2\|H\|_R. \tag{5.30}$$

Proof. We will prove the above inequalities by induction. The generating Hamiltonian G_1 given by lemma 5.2 and the symplectomorphism g_1 by lemma 5.1 satisfy $\|G_1\|'_{R_0-\delta_1} \leqslant \delta_1\|\mathbb{I}^- H\|/[(2\pi + 1)\varepsilon']$, $\|g_1 - \mathrm{Id}\|_{R_0-3\delta_1} < \|\mathbb{I}^- H\|_R \delta_1/\varepsilon'$ and $\mathbb{I}^- H_1 = \mathbb{I}^- H \circ g_1 - \mathbb{I}^-(H + \{H, G_1\})$. Hence,

$$\|\mathbb{I}^- H_1\|_{R_1} \leqslant \|H \circ g_1 - H - \{H, G_1\}\|_{R_1} \leqslant 2 \left(\frac{\|\mathbb{I}^- H\|_R}{\varepsilon'} \right)^2 \|H\|_R \tag{5.31}$$

and

$$\|H_1 - H\|_{R_1} \leqslant \|\nabla H\|_{R_1} \|g_1 - \mathrm{Id}\|_{R_1} \leqslant \frac{2\pi + 1}{4\varepsilon'} \|\mathbb{I}^- H\|_R \|H\|_R \leqslant \frac{2}{\varepsilon'} \|\mathbb{I}^- H\|_R \|H\|_R. \tag{5.32}$$

Thus, equations (5.28) and (5.29) are valid for $k = 1$ and so is (5.30) because $\|H_1\|_{R_1} \leqslant \|H_1 - H\|_{R_1} + \|H\|_R$.

Now, assume that the inequalities are true for k. Under these conditions, lemma 5.2 guarantees the existence of G_{k+1} so that

$$\|G_{k+1}\|'_{R_{k+1}} \leqslant \frac{\delta_{k+1}}{(2\pi + 1)\varepsilon'} \|\mathbb{I}^- H_k\|_{R_k} \tag{5.33}$$

and lemma 5.1 yields g_{k+1}. Therefore, $\mathbb{I}^- H_{k+1} = \mathbb{I}^- H_k \circ g_{k+1} - \mathbb{I}^-(H_k + \{H_k, G_{k+1}\})$ and

$$\begin{aligned}
\|\mathbb{I}^- H_{k+1}\|_{R_{k+1}} &\leqslant \|H_k \circ g_{k+1} - H_k - \{H_k, G_{k+1}\}\|_{R_{k+1}} \\
&\leqslant 2\left(\frac{\|\mathbb{I}^- H_k\|_{R_k}}{\varepsilon'}\right)^2 \|H_k\|_{R_k} \\
&\leqslant \left(\frac{4\|H\|_R}{\varepsilon'^2}\right)^{2^{k+1}-1} \|\mathbb{I}^- H\|_R^{2^{k+1}}.
\end{aligned} \tag{5.34}$$

Similarly,

$$\begin{aligned}
\|H_{k+1} - H_k\|_{R_{k+1}} &\leqslant \|\nabla H_k\|_{R_{k+1}} \|g_{k+1} - \mathrm{Id}\|_{R_{k+1}} \\
&\leqslant \frac{2\pi + 1}{4\delta_{k+1}\varepsilon'} \|\mathbb{I}^- H_k\|_{R_k} \delta_{k+1} \|H_k\|_{R_k} \\
&\leqslant \frac{4}{\varepsilon'} \|\mathbb{I}^- H_k\|_{R_k} \|H\|_R.
\end{aligned} \tag{5.35}$$

Finally, making use of the above inequality,

$$\begin{aligned}
\|H_{k+1}\|_{R_{k+1}} &\leqslant \|H\|_R + \sum_{i=1}^{k+1} \|H_i - H_{i-1}\|_{R_{k+1}} \\
&\leqslant \|H\|_R + \frac{4\|H\|_R}{\varepsilon'} \sum_{i=1}^{k+1} \|\mathbb{I}^- H_{i-1}\|_{R_{i-1}} \\
&\leqslant \|H\|_R + \|H\|_R \sum_{i=1}^{k+1} \left(\frac{4\|H\|_R \|\mathbb{I}^- H\|_R}{\varepsilon'^2}\right)^{2^{i-1}} \\
&\leqslant \left(1 + \sum_{i=1}^{k+1} \frac{1}{2^{2^{i-1}}}\right) \|H\|_R < 2\|H\|_R.
\end{aligned} \tag{5.36}$$

□

Theorem 3.6 will now be a consequence of the result below noticing that $\|H^0\|_R \leqslant R_2\|\boldsymbol{\omega}\| + (dR_2^2/2)\|Q\| \leqslant R_2(\|\boldsymbol{\omega}\| + d\sigma/4)$.

Theorem 5.5. *If*

$$\|H - H^0\|_R < \varepsilon = \frac{\varepsilon'^2}{12\|H^0\|_R} \leqslant \frac{\varepsilon'^2}{8\|H\|_R}, \tag{5.37}$$

then there exists $g = \lim_{k\to+\infty} g^{(k)} \in \mathcal{G}_{R'}$ *such that* $\mathbb{I}^- H \circ g = 0$ *on* $\mathcal{D}_{R'}$. *Furthermore, the maps* $\mathfrak{G}\colon H \mapsto g$ *and* $\mathcal{U}\colon H \mapsto H \circ g$ *are analytic and*

$$\|g - \mathrm{Id}\|_{R'} \leqslant \frac{1}{\varepsilon}\|\mathbb{I}^- H\|_R, \tag{5.38}$$

$$\|H \circ g - H^0\|_{R'} \leqslant \left(1 + \sqrt{\frac{12\|H^0\|_R}{\varepsilon}}\right) \|H - H^0\|_R. \tag{5.39}$$

Proof. Lemmas 5.3 and 5.4 imply that the sequence $g^{(k)}$ converges to a map $g\colon \mathcal{D}_{R'} \to \mathcal{D}_R$ which is analytic and symplectic, and $H_\infty = \lim_{k\to+\infty} H_k = H \circ g$. Moreover, $\mathbb{I}^- H \circ g = \mathbb{I}^- H_\infty = 0$. Since the convergence is uniform, the maps $H \mapsto g$ and $H \mapsto H \circ g$ are analytic.

Notice that

$$\sum_{i=1}^{+\infty}\left(\frac{4\|H\|_R\|\mathbb{I}^-H\|_R}{\varepsilon'^2}\right)^{2^{i-1}} \leqslant \frac{4\|H\|_R\|\mathbb{I}^-H\|_R}{\varepsilon'^2} + \sum_{i=1}^{+\infty}\left(\frac{4\|H\|_R\|\mathbb{I}^-H\|_R}{\varepsilon'^2}\right)^{2i}$$
$$\leqslant \left(1+\frac{16\|H\|_R}{3\varepsilon'^2}\|\mathbb{I}^-H\|_R\right)\frac{4\|H\|_R}{\varepsilon'^2}\|\mathbb{I}^-H\|_R$$
$$\leqslant \frac{20\|H\|_R}{3\varepsilon'^2}\|\mathbb{I}^-H\|_R \leqslant \frac{1}{\varepsilon}\|\mathbb{I}^-H\|_R. \tag{5.40}$$

The inequality in (5.38) follows by taking the limit $k \to +\infty$ in (5.21). That is,

$$\|g-\mathrm{Id}\|_{R'} \leqslant \sum_{i=1}^{+\infty}\frac{\delta_i}{\varepsilon'}\|\mathbb{I}^-H_{i-1}\|_{R_{i-1}} \leqslant \frac{1}{\varepsilon}\|\mathbb{I}^-H\|_R. \tag{5.41}$$

Now,

$$\|H_\infty - H^0\|_{R_k} \leqslant \|H-H^0\|_R + \sum_{i=1}^{+\infty}\|H_i - H_{i-1}\|_{R_i}$$
$$\leqslant \left(1+\sqrt{\frac{12\|H^0\|_R}{\varepsilon}}\right)\|H-H^0\|_R,$$

where we have used lemma 5.4 and the fact that $\|\mathbb{I}^-H\|_R \leqslant \|H-H^0\|_R$. □

Acknowledgments

The authors would like to express their gratitude to H Koch for useful discussions and comments. JLD was supported by FCT through Program POCTI/FEDER and JM by an EPSRC Advanced Research Fellowship.

References

[1] Abad J J and Koch H 2000 Renormalization and periodic orbits for Hamiltonian flows *Commun. Math. Phys.* **212** 371–94

[2] Cassels J W S 1957 *An Introduction to Diophantine Approximation* (Cambridge: Cambridge University Press)

[3] Gaidashev D 2005 Renormalization of isoenergetically degenerate Hamiltonian flows and associated bifurcations of invariant tori *Discrete Contin. Dyn. Syst.* **13** 63–102

[4] Herman M R 1989 Inégalités 'a priori' pour des tores lagrangiens invariants par des difféomorphismes symplectiques *Inst. Hautes Études Sci. Publ. Math.* **70** 47–101

[5] Hille E and Phillips R S 1974 *Functional Analysis and Semi-groups* (*AMS Colloquium Publications* vol 31) (revised edn of 1957)

[6] Khanin K, Lopes Dias J and Marklof J Multidimensional continued fractions, dynamical renormalization and KAM theory *Commun. Math. Phys.* (*Preprint* math-ph/0509007)

[7] Koch H 1999 A renormalization group for Hamiltonians, with applications to KAM tori *Ergod. Theory Dyn. Syst.* **19** 475–521

[8] Koch H 2004 A renormalization group fixed point associated with the breakup of golden invariant tori *Discrete Contin. Dyn. Syst.* **11** 881–909

[9] Koch H and Lopes Dias J 2005 Renormalization of diophantine skew flows, with applications to the reducibility problem *Preprint* mp-arc 05-285

[10] Kocić S 2005 Renormalization of Hamiltonians for diophantine frequency vectors and KAM tori *Nonlinearity* **18** 2513–44

[11] Lagarias J C 1994 Geodesic multidimensional continued fractions *Proc. London Math. Soc.* **69** 464–88

[12] Lopes Dias J 2002 Renormalization of flows on the multidimensional torus close to a KT frequency vector *Nonlinearity* **15** 647–64

[13] Lopes Dias J 2006 Brjuno condition and renormalization for Poincaré flows *Discrete Contin. Dyn. Syst.* **15** 641–56