

# MULTIDIMENSIONAL CONTINUED FRACTIONS, DYNAMICAL RENORMALIZATION AND KAM THEORY

KOSTYA KHANIN, JOÃO LOPES DIAS, AND JENS MARKLOF

ABSTRACT. The disadvantage of ‘traditional’ multidimensional continued fraction algorithms is that it is not known whether they provide simultaneous rational approximations for generic vectors. Following ideas of Dani, Lagarias and Kleinbock-Margulis we describe a simple algorithm based on the dynamics of flows on the homogeneous space  $SL(d, \mathbb{Z}) \backslash SL(d, \mathbb{R})$  (the space of lattices of covolume one) that indeed yields best possible approximations to any irrational vector. The algorithm is ideally suited for a number of dynamical applications that involve small divisor problems. As an example, we explicitly construct a renormalization scheme for the linearization of vector fields on tori of arbitrary dimension.

## 1. INTRODUCTION

The aims of this paper are two-fold. The first objective is to introduce a new multidimensional continued fraction algorithm that is ideally suited for different dynamical applications. The algorithm can be used in order to effectively deal with small divisors whenever quasi-periodicity with several frequencies is an essential feature of a problem. Our second goal is to demonstrate the strength of the algorithm by developing a renormalization approach to KAM theory. The method, being conceptually very simple, is also very general, and allows us to consider a wide class of frequency vectors. For reasons of clarity we restrict our attention to vectors satisfying an explicit Diophantine condition (valid for a set of vectors of full Lebesgue measure); the extension to more general frequency vectors follows straightforwardly from the approach presented here, and will be detailed in a separate publication.

**1.1. Continued fractions.** The classical continued fraction algorithm produces, for every irrational  $\alpha \in \mathbb{R}$ , a sequence of rational numbers  $p_n/q_n$  that approximate  $\alpha$  up to an error of order  $1/q_n^2$ . The first objective of this paper is to develop a multidimensional analogue that allows us to approximate any irrational  $\boldsymbol{\alpha} \in \mathbb{R}^{d-1}$  by rational vectors. The theory of multidimensional continued fractions has a long history going back to Jacobi and Perron (see [30] for extensive reference). The “traditional” m.c.f. algorithms (Jacobi-Perron algorithm etc) have many beautiful ergodic properties. Unfortunately, the quality of approximations provided by them is very difficult to control. In fact, it is not even known whether those algorithms give simultaneous rational approximations for Lebesgue almost all  $\boldsymbol{\alpha}$ . In the case  $d \geq 4$  the only result in this direction is a recent computer-assisted proof of the almost everywhere strong convergence for ordered Jacobi-Perron algorithm [11, 12]. However, even in this case an explicit description of the set of bad vectors seems difficult. For example, the existence of ‘noble’ vectors, that is vectors corresponding to a periodic continued-fraction expansion, for which approximations do not converge, is rather unsatisfactory. For these reasons, there has been only very limited success in the application of traditional multidimensional algorithms in renormalization schemes, in particular those involving typical frequency vectors.

The algorithm we employ here does not suffer from such pathologies. Following Lagarias' seminal ideas in [21], our approach is based on the dynamics of the geodesic<sup>1</sup> flow on the homogeneous space  $\Gamma \backslash G$  with  $G = \mathrm{SL}(d, \mathbb{R})$  and  $\Gamma = \mathrm{SL}(d, \mathbb{Z})$ . Notice that  $\Gamma \backslash G$  may be identified with the space of lattices in  $\mathbb{R}^d$  of covolume one or, equivalently, with the Teichmüller space of flat  $d$ -dimensional tori. Lagarias algorithm is, to the best of our knowledge, the first that provides a strongly convergent multidimensional continued fraction expansion for all vectors. The main advantage of our algorithm is that in addition it allows for effective hyperbolicity estimates which are crucial in dynamical renormalizations.

The problem of multidimensional continued fractions may be formulated in the following way. Given a vector  $\alpha \in \mathbb{R}^{d-1}$  find a sequence of matrices  $T^{(n)} \in \mathrm{GL}(d, \mathbb{Z})$ ,  $n \in \mathbb{N}$ , such that the 'cocycle' corresponding to the products  $P^{(n)} = T^{(n)}T^{(n-1)} \dots T^{(1)}$  exponentially contracts in the direction of the vector  $\omega = \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \in \mathbb{R}^d$  and exponentially expands in all other directions. Thus, the cocycle should have one negative Lyapunov exponent and  $d-1$  positive. In this spirit, our algorithm comprises the following steps:<sup>2</sup>

- (1) With every  $\alpha \in \mathbb{R}^{d-1}$  associate the orbit  $\{C(t) : t \geq 0\} \subset \mathrm{SL}(d, \mathbb{R})$ , where

$$C(t) = \begin{pmatrix} 1_{d-1} & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1_{d-1}e^{-t} & 0 \\ 0 & e^{(d-1)t} \end{pmatrix}, \quad (1.1)$$

$1_{d-1}$  denotes the  $(d-1) \times (d-1)$  unit matrix.

- (2) Fix a Siegel set  $\mathcal{S}$ .<sup>3</sup> Given a sequence of times  $t_1 < t_2 < \dots \rightarrow \infty$ , use classical reduction theory to find matrices  $P^{(n)} \in \mathrm{SL}(d, \mathbb{Z})$  that map the points  $C(t_n)$  to  $\mathcal{S}$ .

- (3) Define the  $n$ th continued fraction map by

$$\alpha^{(n-1)} \mapsto \alpha^{(n)} = \frac{T_{11}^{(n)} \alpha^{(n-1)} + \mathbf{t}_{12}^{(n)}}{\mathbf{t}_{21}^{(n)} \alpha^{(n-1)} + t_{22}^{(n)}}, \quad (1.2)$$

where  $\alpha^{(0)} = \alpha$  and

$$T^{(n)} = \begin{pmatrix} T_{11}^{(n)} & \mathbf{t}_{12}^{(n)} \\ \mathbf{t}_{21}^{(n)} & t_{22}^{(n)} \end{pmatrix} \in \mathrm{SL}(d, \mathbb{Z}) \quad (1.3)$$

is the  $n$ th transfer matrix defined by  $P^{(n)} = T^{(n)}P^{(n-1)}$ .

Note that the choice of matrices  $P^{(n)}$  is not unique, and thus different choices will lead to different algorithms. Uniqueness is guaranteed if one imposes the additional requirement that  $P^{(n)}C(t_n)$  belongs to a fundamental domain  $\mathcal{F}$ , as in Lagarias' algorithm. However, in this case, due to a very complicated geometry of fundamental domains, it is difficult to control hyperbolic properties of the matrices  $T^{(n)}$ . On the other hand, the use of Siegel sets allows us to relate these hyperbolic properties with the Diophantine properties of vector  $\alpha$ . Another important difference with the Lagarias algorithm is connected with a choice of times  $t_n$ . Lagarias requires a specific choice of  $t_n$  that correspond to the times the geodesic exits the fundamental domains. On the contrary, our algorithm, selects the  $t_n$  according to the diophantine properties of  $\alpha$ : the better approximable by rationals  $\alpha$  is, the faster the  $t_n$  have to increase with  $n$  in order to ensure the hyperbolicity of

<sup>1</sup>The term 'geodesic' is slightly inaccurate when  $d > 2$ . The orbits of the flow on  $\Gamma \backslash G$  we are discussing here correspond in fact to geodesics on the unit cotangent bundle of the space  $\Gamma \backslash G / \mathrm{SO}(d)$  for a certain family of initial conditions. Only for  $d = 2$  the cotangent bundle can be identified with  $\Gamma \backslash G$ .

<sup>2</sup>The fact that we restrict our attention here to  $\mathrm{SL}(d, \mathbb{Z})$  is not essential; the algorithm can readily be extended to allow for approximation by elements from  $\mathrm{GL}(d, \mathbb{Z})$ .

<sup>3</sup>A Siegel set has the property that (a) it contains a fundamental domain  $\mathcal{F}$  of  $\mathrm{SL}(d, \mathbb{Z})$  in  $\mathrm{SL}(d, \mathbb{R})$ , and (b) it is contained in a finite number of translates  $P\mathcal{F}$ ,  $P \in \mathrm{SL}(d, \mathbb{Z})$ .

the transfer matrices. Only in the special case of badly approximable  $\alpha$  the times  $t_n$  can be chosen to have bounded gaps  $t_n - t_{n-1}$ .

Notice that in dimension  $d = 2$  Lagarias' algorithm, and thus also our algorithm, do not reduce to the classical continued fraction dynamics, see [10] for details. However, a subsequence of exit times, called Hermite critical times, gives an accelerated version of the standard continued fractions [21].

Since the action of  $\mathrm{SL}(d, \mathbb{R})$  on  $\mathbb{R}^{d-1}$  by fractional linear transformation defines a group action, we have

$$\alpha^{(n)} = \frac{P_{11}^{(n)}\alpha + p_{12}^{(n)}}{p_{21}^{(n)}\alpha + p_{22}^{(n)}}, \quad (1.4)$$

where

$$P^{(n)} = T^{(n)}T^{(n-1)} \dots T^{(1)} = \begin{pmatrix} P_{11}^{(n)} & p_{12}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} \end{pmatrix} \in \mathrm{SL}(d, \mathbb{Z}). \quad (1.5)$$

**1.2. Renormalization.** Dani [5] and Kleinbock-Margulis [17] observed that Diophantine properties of  $\alpha$  translate to divergence properties of the corresponding orbit  $\{\Gamma C(t) : t \geq 0\} \subset \Gamma \backslash G$  in the cusps of  $\Gamma \backslash G$ . We exploit these results to show that, under mild Diophantine conditions on  $\alpha$  (satisfied by a set of  $\alpha$  of full Lebesgue measure, cf. Sec. 2.6), there is a sequence of times  $t_n$  such that the transfer matrices  $T^{(n)}$  are uniformly hyperbolic in a sense made precise in Sec. 2.8. This fact allows us to develop renormalization schemes for vector fields and Hamiltonian flows that had previously been constructed only in dimension one [23] or for very special choices of  $\alpha$  [1, 18, 19, 22]. In this paper we focus on the case of vector fields, where, as in the traditional approaches (cf. [28]), the proofs are technically slightly simpler than in the case of Hamiltonian flows. Full details of the latter are given in [15].

We say  $\omega \in \mathbb{R}^d$  is *Diophantine* if there are constants  $\epsilon > 0$ ,  $C > 0$  such that

$$\|\mathbf{k}\|^{(d-1)(1+\epsilon)} |\mathbf{k} \cdot \omega| > C, \quad (1.6)$$

for all  $\mathbf{k} \in \mathbb{Z}^d - \{\mathbf{0}\}$ . Note that we may assume without loss of generality that  $\omega$  is of the form  $\omega = \begin{pmatrix} \alpha \\ 1 \end{pmatrix}$  with  $\alpha \in \mathbb{R}^{d-1}$ . Condition (1.6) then translates to a standard Diophantine condition on  $\alpha$ , see Sec. (2.6) for details.

**Theorem 1.1.** *For any real analytic vector field  $\mathbf{v}$  on  $\mathbb{T}^d$ ,  $d \geq 2$ , sufficiently close to a constant vector field with Diophantine vector  $\omega \in \mathbb{R}^d$ , there is  $b > 0$ , an analytic curve  $\mathbf{p} : (-b, b) \rightarrow \mathbb{R}^d$ ,  $s \mapsto \mathbf{p}^s$ , and an analytic conjugacy  $h$  isotopic to the identity between the flow generated by  $\mathbf{v} + \mathbf{p}^s$  and the linear flow  $\phi_t(\mathbf{x}) = \mathbf{x} + t(1 + s)\omega$  on  $\mathbb{T}^d$ ,  $t \geq 0$ , for each  $|s| < b$ . Moreover, the maps  $\mathbf{v} \mapsto h$  and  $\mathbf{v} \mapsto \mathbf{p}$  are analytic.*

Let us emphasize that the result holds for all analytic vector fields close to a constant one without any additional conditions, such as preservation of volume etc.

The main strategy of the proof of the above theorem is as follows. Consider a vector field  $X(\mathbf{x}, \mathbf{y}) = \omega + f(\mathbf{x}) + \mathbf{y}$  where  $\mathbf{x} \in \mathbb{T}^d$ , and  $\mathbf{y} \in \mathbb{R}^d$  is an auxiliary parameter. The vector field  $f(\mathbf{x})$  is a sufficiently small analytic perturbation of a constant vector field. We may furthermore assume that  $\omega = \begin{pmatrix} \alpha \\ 1 \end{pmatrix}$  for some Diophantine  $\alpha \in \mathbb{R}^{d-1}$ ; this achieved by a rescaling of time. The aim is to find a value of parameter  $\mathbf{y} = \mathbf{y}_\omega$  such that the vector field  $X(\mathbf{x}, \mathbf{y}_\omega)$  is linearizable to a constant vector field identically equal to  $\omega$  by means of an analytic transformation of the coordinates on  $\mathbb{T}^d$ .

Renormalization is an iterative process, and we thus assume that after the  $(n - 1)$ th renormalization step the vector field is of the slightly more general form

$$X_{n-1}(\mathbf{x}, \mathbf{y}) = \omega^{(n-1)} + f_{n-1}(\mathbf{x}, \mathbf{y}) \quad (1.7)$$

where  $\boldsymbol{\omega}^{(n-1)} = (\boldsymbol{\alpha}_1^{(n-1)})$  and  $\boldsymbol{\alpha}^{(n-1)}$  is given by the continued fraction algorithm, cf. (1.4). The Fourier modes of  $f_{n-1}$  are smaller than in the previous step, and decay exponentially as  $\|\mathbf{k}\| \rightarrow +\infty$ . We define a cone of resonant modes by a relation  $I_{n-1}^+ = \{\mathbf{k} \in \mathbb{Z}^d : |\mathbf{k} \cdot \boldsymbol{\omega}^{(n-1)}| \leq \sigma_{n-1} \|\mathbf{k}\|\}$ . The  $n$ th step requires the following operations.

- (1) Eliminate all Fourier modes outside of the resonant cone  $I_{n-1}^+$ .
- (2) Apply a linear operator corresponding to a coordinate transformation given by the inverse transfer matrix  $T^{(n)-1}$ .
- (3) Rescale time to ensure that the frequency vector is of the form  $\boldsymbol{\omega}^{(n)} = (\boldsymbol{\alpha}_1^{(n)})$ .

The conjugate action on the Fourier modes is given by  $\mathbf{k} \mapsto {}^\top T^{(n)-1} \mathbf{k}$ . It follows from the hyperbolicity of  $T^{(n)}$  that this transformation contracts for  $\mathbf{k} \in I_{n-1}^+$  if  $\sigma_{n-1}$  is small enough. This gives a significant improvement of the analyticity domain which results in the decrease of the estimates for the corresponding Fourier modes. As a result, all Fourier modes apart from the zero modes get smaller. To decrease the size of the latter, we choose a parameter  $\mathbf{y} = \mathbf{y}_n$  in such a way that the corresponding zero modes vanish, and then consider a neighbourhood of  $\mathbf{y}$ -values centred at  $\mathbf{y}_n$ . That is, the auxiliary parameter  $\mathbf{y}$  is used to eliminate an instability in the direction of constant vector fields. To get enough control on the parameter dependence we perform an affine rescaling of this parameter on every renormalization step. One can then show that the corresponding sequence of parameter domains is nested and converges to a single point  $\mathbf{y} = \mathbf{y}_\omega$  for which the initial vector field is indeed linearizable.

In order for the scheme to be effective, the sequence of stopping times  $t_n$  and the sizes of the resonant cones defined by the sequence of  $\sigma_n$  must be chosen properly. Large intervals  $\delta t_n = t_n - t_{n-1}$  improve hyperbolicity but, on the other hand, worsen estimates for the norms  $\|T^{(n)}\|, \|T^{(n)-1}\|$ . Similarly, if  $\sigma_n$  is too small the elimination of non-resonant modes will give large contributions; on the other hand, for large values of  $\sigma_n$  the multiplication by  ${}^\top T^{(n+1)-1}$  will not yield a contraction for  $\mathbf{k} \in I_n^+$ . As we shall show below a right choice of sequences  $\{(t_n, \sigma_n)\}$  can be made, depending on the Diophantine properties of the vector  $\boldsymbol{\omega}$ .

The mechanism of convergence of renormalizations is well understood by now. The main framework was developed in [18] for special frequency vectors. Our proof of convergence of the renormalizations follows the same basic steps. However, there are several essential differences. The most important one is related to our choice of a sequence  $t_n$  which we discussed above. In fact, the use of long intervals  $\delta t_n = t_n - t_{n-1}$  simplifies the estimates and makes the proof more transparent. Another important point is connected with the construction of the analytic conjugacy  $h$ . That is built as  $h = \lim h_n$  where  $h_n$  is the composition of  $P^{(n)}$  with the coordinate transformations used in all renormalization steps up to the  $n$ th one. The domain of analyticity for each  $h_n$  is a complex strip around  $\mathbb{R}^d$  with width  $\rho_n \|P^{(n)}\|^{-1}$ , where  $\rho_n$  is the width of analyticity strip for the vector fields at the corresponding step. In  $d = 2$ , using standard continued fractions, the ratio  $\rho_n \|P^{(n)}\|^{-1}$  is bounded from below ([24]). However, in the multidimensional case, since we do not have such a bound, we have to consider shrinking domains. As a result, the conjugacy  $h$  is only  $C^1$  on  $\mathbb{R}^d$ . To recover analyticity, we use the analytic dependence of the conjugacy on the initial vector field (Theorem 3.19).

The above arguments are adapted in [15] to the case of Hamiltonian flows, which is technically slightly more challenging. Let  $B \subset \mathbb{R}^d$ ,  $d \geq 2$ , be an open set containing the origin, and let  $H^0$  be a real-analytic Hamiltonian function

$$H^0(\mathbf{x}, \mathbf{y}) = \boldsymbol{\omega} \cdot \mathbf{y} + \frac{1}{2} {}^\top \mathbf{y} Q \mathbf{y}, \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{T}^d \times B, \quad (1.8)$$

with  $\omega \in \mathbb{R}^d$  and a real symmetric  $d \times d$  matrix  $Q$ . It is said to be non-degenerate if  $\det Q \neq 0$ .

**Theorem 1.2.** *Suppose  $H^0$  is non-degenerate and  $\omega$  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  $d$ -dimensional torus where it is analytically conjugated to the linear flow  $\phi_t(\mathbf{x}) = \mathbf{x} + t\omega$  on  $\mathbb{T}^d$ ,  $t \geq 0$ . The conjugacy depends analytically on  $H$ .*

**1.3. A brief review.** The idea of renormalization was introduced to the theory of dynamical systems by Feigenbaum [6] in the late 1970's. In the case of Hamiltonian systems with two degrees of freedom MacKay proposed in the early 1980's a renormalization scheme for the construction of KAM invariant tori [25]. The scheme was realized for the construction of invariant curves for two-dimensional conservative maps of the cylinder. An important feature of MacKay's approach is the analysis of both smooth KAM invariant curves and so-called critical curves corresponding to critical values of a parameter above which invariant curves no longer exist. From the point of view of renormalization theory the KAM curves correspond to a trivial linear fixed point for the renormalization transformations, while critical curves give rise to very complicated fixed points with nontrivial critical behavior. MacKay's renormalization scheme was carried out only for a small class of Diophantine rotation numbers with periodic continued fraction expansion (such as the golden mean). Khanin and Sinai studied a different renormalization scheme for general Diophantine rotation numbers [16]. Both of the above early approaches were based on renormalization for maps or their generating functions. Essentially, the renormalization transformations are defined in the space of pairs of mappings which, being iterates of the same map, commute with each other. These commutativity conditions cause difficult technical problems, and led MacKay [26] to propose the development of alternative renormalization schemes acting directly on vector fields. The same idea was realized by Koch [18] who proves a KAM type result for analytic perturbations of linear Hamiltonians  $H^0(\mathbf{x}, \mathbf{y}) = \omega \cdot \mathbf{y}$ , for frequencies  $\omega$  which are eigenvectors of hyperbolic matrices in  $SL(d, \mathbb{Z})$  with only one unstable direction. Notice that the set of such frequencies has zero Lebesgue measure and in the case  $d = 2$  corresponds to vectors with a quadratic irrational slope. Further improvements and applications of Koch's techniques appeared in [1, 19, 22, 23, 7], emphasizing the connection between KAM and renormalization theories. The results of this paper illustrate that such a programme can indeed be carried out in considerable generality. Another direction was followed in [20], presenting a computer-assisted proof of the existence of MacKay's golden mean critical renormalization fixed point in the context of Hamiltonian vector fields with two degrees of freedom.

Other renormalization ideas have appeared in the context of the stability of invariant tori for nearly integrable Hamiltonian systems inspired by quantum field theory and an analogy with KAM theory (see e.g. [3], and [8, 9] where it is used a graph representation of the invariant tori in terms of Feynman diagrams).

Since the main focus of the paper is related to typical frequency vectors, we do not present here concrete examples of the work of the m.c.f. algorithm. Such examples are easy to construct only in the case of badly approximable  $\alpha$ . For the sake of transparency we have restricted our attention to Diophantine vectors  $\omega$ , which form a set of full Lebesgue measure and are thus 'typical.' A more detailed analysis under weaker Diophantine conditions is, in principle, possible within the present framework. It is however a fundamental open problem to state a sharp (i.e. the weakest possible) Diophantine condition under which the above conjugacy can be established. The answer to this question is known only in the case  $d = 2$ , namely in the discrete-time situation

for the Siegel problem [31] and for circle diffeomorphisms [32], where the Diophantine condition is of Brjuno type.

It would also be interesting to see whether the multidimensional continued fraction algorithm presented here will allow generalizations of other one-dimensional renormalization constructions. A concrete challenge is for instance the extension of the recent results on the reducibility of cocycles over irrational rotations by Ávila and Krikorian [2].

In the next section we introduce the multidimensional continued fraction algorithm, and include a discussion of its hyperbolicity properties required in the renormalization schemes. Section 3 provides a detailed account of one exemplary case, the renormalization of vector fields.

## 2. MULTIDIMENSIONAL CONTINUED FRACTIONS AND FLOWS ON HOMOGENEOUS SPACES

**2.1. Flows on homogeneous spaces.** Let us set  $G = \mathrm{SL}(d, \mathbb{R})$  and  $\Gamma = \mathrm{SL}(d, \mathbb{Z})$ , and define the diagonal subgroup  $\{E^t : t \in \mathbb{R}\}$  in  $G$ , where

$$E^t = \mathrm{diag}(e^{r_1 t}, \dots, e^{r_d t}) \quad (2.1)$$

with constants satisfying the conditions

$$r_1, \dots, r_{d-1} < 0 < r_d, \quad \sum_{j=1}^d r_j = 0. \quad (2.2)$$

The right action of  $E^t$  on the homogeneous space  $\Gamma \backslash G \backslash G$  generates the flow

$$\Phi^t : \Gamma \backslash G \rightarrow \Gamma \backslash G, \quad \Gamma M \mapsto \Gamma M E^t. \quad (2.3)$$

Since  $G$  is a simple Lie group with finite center,  $\Phi^t$  is ergodic and mixing [27].

Let  $\mathcal{F} \subset G$  be a fundamental domain of the left action of  $\Gamma$  on  $G$ . Recall that, by definition of the fundamental domain of a free group action,

$$\bigcup_{P \in \Gamma} P\mathcal{F} = G, \quad \mathcal{F} \cap P\mathcal{F} = \emptyset \text{ for all } P \in \Gamma - \{1\}, \quad (2.4)$$

and hence, for any given  $M \in G$ , there is a unique family of  $P(t) \in \Gamma$  such that

$$M(t) := P(t)M E^t \in \mathcal{F} \quad (2.5)$$

holds for all  $t \in \mathbb{R}$ .

**2.2. A convenient parametrization.** Let us consider those  $M \in G$  which can be written as

$$M = \begin{pmatrix} 1 & \boldsymbol{\alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & \mathbf{0} \\ \top \boldsymbol{\beta} & \gamma \end{pmatrix} \quad (2.6)$$

where  $A \in \mathrm{Mat}_{d-1}(\mathbb{R})$  (the space of real  $(d-1) \times (d-1)$  matrices),  $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^{d-1}$  are column vectors,  $\gamma \in \mathbb{R}$  with  $\gamma > 0$ . This yields a local parametrization of  $G$  for the set

$$G_+ := \left\{ \begin{pmatrix} 1 & \boldsymbol{\alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & \mathbf{0} \\ \top \boldsymbol{\beta} & \gamma \end{pmatrix} \in G : A \in \mathrm{Mat}_{d-1}(\mathbb{R}), \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^{d-1}, \gamma \in \mathbb{R}_{>0} \right\}, \quad (2.7)$$

which is particularly convenient for our purposes. All other matrices are either of the above form with  $\gamma < 0$  instead, or may be written as

$$M = S \begin{pmatrix} 1 & \boldsymbol{\alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & \mathbf{0} \\ \top \boldsymbol{\beta} & \gamma \end{pmatrix} \quad (2.8)$$

where  $S \in \Gamma$  is a suitably chosen ‘‘signed permutation matrix’’, i.e., every row and every column contains one and only one non-zero coefficient, which is either 1 or  $-1$ . In the

following we will stay clear of the parameter singularity at  $\gamma = 0$ , and thus may assume without loss of generality  $S = 1$ .

To work out the action of a general element  $T \in G$  in the above parametrization, consider

$$T : M \mapsto \tilde{M} := TM \quad (2.9)$$

where

$$T = \begin{pmatrix} T_{11} & \mathbf{t}_{12} \\ \top \mathbf{t}_{21} & t_{22} \end{pmatrix}, \quad (2.10)$$

$M$  is as above and

$$\tilde{M} = \begin{pmatrix} 1 & \tilde{\boldsymbol{\alpha}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{A} & \mathbf{0} \\ \top \tilde{\boldsymbol{\beta}} & \tilde{\gamma} \end{pmatrix}. \quad (2.11)$$

A short calculation yields the fractional linear action

$$\boldsymbol{\alpha} \mapsto \tilde{\boldsymbol{\alpha}} = \frac{T_{11}\boldsymbol{\alpha} + \mathbf{t}_{12}}{\top \mathbf{t}_{21}\boldsymbol{\alpha} + t_{22}}, \quad (2.12)$$

and

$$\gamma \mapsto \tilde{\gamma} = (\top \mathbf{t}_{21}\boldsymbol{\alpha} + t_{22})\gamma, \quad (2.13)$$

and more complicated expressions for  $\tilde{A}, \tilde{\boldsymbol{\beta}}$  which will not be needed in the following.

**2.3. Multidimensional continued fractions.** Let  $t_0 = 0 < t_1 < t_2 < \dots \rightarrow \infty$  be sequence of times, with gaps

$$\delta t_n := t_n - t_{n-1} \quad (2.14)$$

chosen large enough so that  $P(t_n) \neq P(t_{n-1})$ , where  $P(t)$  is defined by (2.5). The sequence  $P^{(n)} := P(t_n)$  of matrices in  $\Gamma$  may be viewed as the continued fraction approximants of the vector  $\boldsymbol{\alpha}$ , which are the best possible for suitable choices of a fundamental domain  $\mathcal{F}$  and times  $t_n$ , see [21]. Let us furthermore put  $M^{(n)} := M(t_n)$  with  $M(t)$  as in (2.5), and define  $\boldsymbol{\alpha}^{(n)}, \gamma^{(n)}$  by the decomposition (2.6), i.e., by

$$M^{(n)} = \begin{pmatrix} 1 & \boldsymbol{\alpha}^{(n)} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A^{(n)} & \mathbf{0} \\ \top \boldsymbol{\beta}^{(n)} & \gamma^{(n)} \end{pmatrix}. \quad (2.15)$$

From  $M^{(n)} = P^{(n)}ME^{t_n}$  and (2.12), (2.13) we deduce

$$\boldsymbol{\alpha}^{(n)} = \frac{P_{11}^{(n)}\boldsymbol{\alpha} + \mathbf{p}_{12}^{(n)}}{\top \mathbf{p}_{21}^{(n)}\boldsymbol{\alpha} + p_{22}^{(n)}}, \quad (2.16)$$

and

$$\gamma^{(n)} = (\top \mathbf{p}_{21}^{(n)}\boldsymbol{\alpha} + p_{22}^{(n)})e^{r_d t_n} \gamma \quad (2.17)$$

where

$$P^{(n)} = \begin{pmatrix} P_{11}^{(n)} & \mathbf{p}_{12}^{(n)} \\ \top \mathbf{p}_{21}^{(n)} & p_{22}^{(n)} \end{pmatrix}. \quad (2.18)$$

It is evident that if the components of  $(\top \boldsymbol{\alpha}, 1)$  are linearly independent over  $\mathbb{Q}$ , then  $\gamma \neq 0$  implies  $\gamma^{(n)} \neq 0$  for all  $n \geq 0$ .

We shall later employ the transfer matrices  $T^{(n)}$  defined by  $P^{(n)} = T^{(n)}P^{(n-1)}$ . Here,  $M^{(n)} = T^{(n)}M^{(n-1)}E^{\delta t_n}$  implies

$$\boldsymbol{\alpha}^{(n)} = \frac{T_{11}^{(n)}\boldsymbol{\alpha}^{(n-1)} + \mathbf{t}_{12}^{(n)}}{\top \mathbf{t}_{21}^{(n)}\boldsymbol{\alpha}^{(n-1)} + t_{22}^{(n)}}, \quad (2.19)$$

and

$$\gamma^{(n)} = (\top \mathbf{t}_{21}^{(n)}\boldsymbol{\alpha}^{(n-1)} + t_{22}^{(n)})e^{r_d \delta t_n} \gamma^{(n-1)}, \quad (2.20)$$

where

$$T^{(n)} = \begin{pmatrix} T_{11}^{(n)} & \mathbf{t}_{12}^{(n)} \\ \top \mathbf{t}_{21}^{(n)} & t_{22}^{(n)} \end{pmatrix}. \quad (2.21)$$

**2.4. Siegel sets.** In dimensions  $d > 2$  it is difficult to describe the geometry of a fundamental domain  $\mathcal{F}$ . To overcome this problem, C. Siegel introduced simply connected sets  $\mathcal{S}_d \subset G$  which have the property that they contain  $\mathcal{F}$  and are contained in a finite number of translates  $P\mathcal{F}$ ,  $P \in \Gamma$ . Consider the Iwasawa decomposition

$$M = nak \quad (2.22)$$

where

$$n = \begin{pmatrix} 1 & u_{12} & \dots & u_{1d} \\ & \ddots & \ddots & \vdots \\ & & \ddots & u_{d-1,d} \\ & & & 1 \end{pmatrix}, \quad a = \begin{pmatrix} a_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_d \end{pmatrix} \quad (2.23)$$

and  $k \in \mathrm{SO}(d)$ , with  $u_{ij}, a_j \in \mathbb{R}$ ,  $a_j > 0$ ,  $a_1 \cdots a_d = 1$ . Then

$$\mathcal{S}_d = \{nak : n \in \mathcal{F}_N, a_j \geq \frac{\sqrt{3}}{2} a_{j+1} > 0 \ (j = 1, \dots, d-1), k \in \mathrm{SO}(d)\} \quad (2.24)$$

is an example of a *Siegel set* [29]; here  $\mathcal{F}_N$  denotes a compact fundamental region of  $(\Gamma \cap N) \backslash N$ , where  $N$  is the upper triangular group of elements of the form  $n$  as above.

**2.5. Dani's correspondence.** We assume from now on that  $r_1, \dots, r_{d-1} = -1$ ,  $r_d = d-1$ , i.e.,

$$E^t = \mathrm{diag}(e^{-t}, \dots, e^{-t}, e^{(d-1)t}). \quad (2.25)$$

Let us denote by  $|\cdot|$  the maximum norm in  $\mathbb{R}^{d-1}$ . A vector  $\boldsymbol{\alpha} \in \mathbb{R}^{d-1}$  is called *badly approximable* or of *bounded type*, if one of the following equivalent conditions is satisfied.

(i) *There exists a constant  $C > 0$  such that*

$$|k\boldsymbol{\alpha} + \mathbf{m}|^{d-1}|k| > C \quad (2.26)$$

*for all  $\mathbf{m} \in \mathbb{Z}^{d-1}$ ,  $k \in \mathbb{Z} - \{0\}$ .*

(ii) *There exists a constant  $C > 0$  such that*

$$|\mathbf{m}|^{d-1}|\mathbf{m} \cdot \boldsymbol{\alpha} + k| > C \quad (2.27)$$

*for all  $\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}$ ,  $k \in \mathbb{Z}$ .*

The statements (i) and (ii) are equivalent in view of Khintchine's transference principle ([4] Chapter V).

We recall *Dani's correspondence* in the following proposition (cf. [5], Theorem 2.20).

**Proposition 2.1.** *The orbit  $\{\Gamma M E^t : t \geq 0\}$ , with  $M$  as in (2.6), is bounded in  $\Gamma \backslash G$  if and only if the vector  $\boldsymbol{\alpha}$  is of bounded type.*

The reason why the parameters  $A, \boldsymbol{\beta}, \gamma$  are irrelevant in the statement is that the family of matrices

$$W(t) = E^{-t} \begin{pmatrix} A & \mathbf{0} \\ \top \boldsymbol{\beta} & \gamma \end{pmatrix} E^t \quad (2.28)$$

is bounded in  $G$  for all  $t \geq 0$ .

The boundedness of the orbit  $\{\Gamma M E^t : t \geq 0\}$  implies of course that there is a compact set  $\mathcal{C} \in G$  such that  $M(t) \in \mathcal{C}$  for all  $t \geq 0$ , with  $M(t)$  as in (2.5).



**2.6. Diophantine conditions.** A vector  $\boldsymbol{\alpha} \in \mathbb{R}^{d-1}$  is called *Diophantine*, if there exist constants  $\epsilon > 0$ ,  $C > 0$  such that

$$|\mathbf{m}|^{(d-1)(1+\epsilon)} |\mathbf{m} \cdot \boldsymbol{\alpha} + k| > C \quad (2.29)$$

for all  $\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}$ ,  $k \in \mathbb{Z}$ . It is well known that Diophantine vectors form a set of full Lebesgue measure [4].

Let us show that (2.29) implies the inequality

$$\|\mathbf{k}\|^{(d-1)(1+\epsilon)} |\mathbf{k} \cdot \boldsymbol{\omega}| > C, \quad (2.30)$$

for all  $\mathbf{k} \in \mathbb{Z}^d - \{\mathbf{0}\}$ , where  $\boldsymbol{\omega} = \begin{pmatrix} \boldsymbol{\alpha} \\ 1 \end{pmatrix}$ , cf. (1.6). With  $\mathbf{k} = \begin{pmatrix} \mathbf{m} \\ k \end{pmatrix}$ , (2.29) yields

$$|\mathbf{k} \cdot \boldsymbol{\omega}| > C |\mathbf{m}|^{-(d-1)(1+\epsilon)} \geq C |\mathbf{k}|^{-(d-1)(1+\epsilon)} \geq C \|\mathbf{k}\|^{-(d-1)(1+\epsilon)} \quad (2.31)$$

for all  $\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}$ ,  $k \in \mathbb{Z}$ . In the case when  $\mathbf{m} = \mathbf{0}$ , we have  $k \neq 0$  (since  $\mathbf{k} \neq \mathbf{0}$ ) and thus (2.30) holds trivially.<sup>4</sup> Note also that (2.30) evidently implies (2.29), however with different choices for  $C$  in both inequalities.

Following [17] we define the following function on  $G$ ,

$$\delta(M) = \inf_{\mathbf{k} \in \mathbb{Z}^d - \{\mathbf{0}\}} |\mathbf{k}^\top M|. \quad (2.32)$$

It is easily checked that  $\delta(M)$  is invariant under left action of  $\Gamma$ , and may thus be viewed as a function on  $\Gamma \backslash G$ . In terms of the Iwasawa parametrization (2.22) and the Siegel set  $\mathcal{S}_d$  defined in (2.24) we have the following estimate.

**Lemma 2.2.** *For  $M = nak \in \mathcal{S}_d$  as in (2.22), (2.24), there are constants  $0 < C_1 \leq C_2$  such that for all  $0 < a_d \leq 1$*

$$C_1 a_d \leq \delta(M) \leq C_2 a_d. \quad (2.33)$$

*Proof.* Since  $\|\mathbf{x}\| \ll |\mathbf{x}| \ll \|\mathbf{x}\|$  for all  $\mathbf{x} \in \mathbb{R}^d$ , we may prove the statement of the lemma for the function

$$\tilde{\delta}_d(M) = \inf_{\mathbf{k} \in \mathbb{Z}^d - \{\mathbf{0}\}} \|\mathbf{k}^\top M\| \quad (2.34)$$

instead.<sup>5</sup> Due to the rotational invariance of the Euclidean distance we may assume that  $k \in \text{SO}(d)$  is the identity.

*Proof by induction.* The statement trivially holds for  $d = 1$ . Therefore let us assume the assertion is true for dimension  $d - 1$ . The  $j$ th coefficient of the vector  $\mathbf{k}^\top M$  is

$$(\mathbf{k}^\top M)_j = \left( k_j + \sum_{i=1}^{j-1} k_i u_{ij} \right) a_j. \quad (2.35)$$

Since  $a_1 \rightarrow \infty$  when  $a_d \rightarrow 0$ , this implies that when taking the infimum in (2.32) we must take  $k_1 = 0$  for all sufficiently small  $a_d$ . Thus we now need to estimate

$$\inf_{\tilde{\mathbf{k}} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}} \max_{2 \leq j \leq d} \left| \left( k_j + \sum_{i=1}^{j-1} k_i u_{ij} \right) a_j \right| = a_1^{-1/(d-1)} \inf_{\tilde{\mathbf{k}} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}} \max_{2 \leq j \leq d} \left| \left( k_j + \sum_{i=1}^{j-1} k_i u_{ij} \right) \tilde{a}_j \right| \quad (2.36)$$

where  $\mathbf{k}^\top = (k_2, \dots, k_d)$ ,  $\tilde{a}_j = a_1^{1/(d-1)} a_j$  so that  $\tilde{a}_2 \cdots \tilde{a}_d = 1$ . Now

$$\inf_{\tilde{\mathbf{k}} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}} \max_{2 \leq j \leq d} \left| \left( k_j + \sum_{i=1}^{j-1} k_i u_{ij} \right) \tilde{a}_j \right| =: \delta_{d-1}(\tilde{M}) \asymp \tilde{\delta}_{d-1}(\tilde{M}) \quad (2.37)$$

<sup>4</sup>Note that every admissible constant in (2.29) needs to satisfy  $C < 1/2$ ; to see this, choose  $\mathbf{m} = (1, 0, \dots, 0)$ , and  $k \in \mathbb{Z}$  such that  $|\alpha_1 + k| \leq 1/2$ .

<sup>5</sup>In the following,  $A \ll B$  means ‘there is a constant  $C > 0$  such that  $A \leq CB$ ’. If  $A \ll B \ll A$  we will also use the notation  $A \asymp B$ .

where  $\tilde{M} = \tilde{n}\tilde{a}$  with

$$\tilde{n} = \begin{pmatrix} 1 & u_{23} & \cdots & u_{2d} \\ & \ddots & \ddots & \vdots \\ & & \ddots & u_{d-1,d} \\ & & & 1 \end{pmatrix}, \quad \tilde{a} = \begin{pmatrix} \tilde{a}_2 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \tilde{a}_d \end{pmatrix}. \quad (2.38)$$

It is easily checked that  $\tilde{M} \in \mathcal{S}_{d-1}$ , so by the induction hypothesis, for suitable constants  $0 < C_{1,d-1} \leq C_{2,d-1}$ , we have

$$C_{1,d-1} \tilde{a}_d \leq \tilde{\delta}_{d-1}(\tilde{M}) \leq C_{2,d-1} \tilde{a}_d, \quad (2.39)$$

provided  $\tilde{a}_d = a_1^{1/(d-1)} a_d \leq 1$ . So for  $a_d$  sufficiently small and  $a_1^{1/(d-1)} a_d \leq 1$ , we have

$$C_{1,d-1} a_d \leq \tilde{\delta}_d(M) \leq C_{2,d-1} a_d. \quad (2.40)$$

In the remaining case  $\tilde{a}_d > 1$ , all  $\tilde{a}_j$  are bounded from above and below by positive constants, and hence  $\tilde{\delta}_{d-1}(\tilde{M})$  is bounded from above and below by positive constants. Furthermore  $\tilde{a}_d > 1$  implies  $a_1^{-1/(d-1)} < a_d$ , and, in view of our choice of the Siegel set,  $a_1^{-1} = a_2 \cdots a_d \gg a_d^{d-1}$ . So

$$a_d \ll a_1^{-1/(d-1)} < a_d \quad (2.41)$$

and the required bound follows from (2.36) also for the case  $\tilde{a}_d > 1$ .  $\square$

**Lemma 2.3.** *Choose  $M$  as in (2.6) and suppose  $\alpha$  satisfies condition (2.29). Then there exists a constant  $C' > 0$  such that for all  $t \geq 0$*

$$\delta(ME^t) > C' e^{-\theta t} \quad (2.42)$$

where

$$\theta = \frac{(d-1)\epsilon}{d + (d-1)\epsilon}. \quad (2.43)$$

*Proof.* Let us put  ${}^\top \mathbf{k} = ({}^\top \mathbf{m}, k)$  with  $\mathbf{m} \in \mathbb{Z}^{d-1}$  and  $k \in \mathbb{Z}$ . Then

$$\begin{aligned} \delta(ME^t) &= \inf_{(\mathbf{m}, k) \in \mathbb{Z}^d - \{\mathbf{0}\}} |({}^\top \mathbf{m} e^{-t}, ({}^\top \mathbf{m} \alpha + k) e^{(d-1)t}) W(t)| \\ &\gg \inf_{(\mathbf{m}, k) \in \mathbb{Z}^d - \{\mathbf{0}\}} |({}^\top \mathbf{m} e^{-t}, ({}^\top \mathbf{m} \alpha + k) e^{(d-1)t})| \end{aligned} \quad (2.44)$$

since  $W(t)$ , as defined in (2.28), is bounded in  $G$  for all  $t \geq 0$ . Furthermore for  $t$  sufficiently large

$$\inf_{(\mathbf{m}, k) \in \mathbb{Z}^d - \{\mathbf{0}\}} |({}^\top \mathbf{m} e^{-t}, ({}^\top \mathbf{m} \alpha + k) e^{(d-1)t})| = \inf_{\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}, k \in \mathbb{Z}} |({}^\top \mathbf{m} e^{-t}, ({}^\top \mathbf{m} \alpha + k) e^{(d-1)t})| \quad (2.45)$$

which, in view of the Diophantine condition (2.29), is bounded from below by

$$\geq \inf_{\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}} |({}^\top \mathbf{m} e^{-t}, C |\mathbf{m}|^{-(d-1)(1+\epsilon)} e^{(d-1)t})| = e^{-\theta t} \inf_{\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}} |({}^\top \mathbf{x}, C |\mathbf{x}|^{-(d-1)(1+\epsilon)})| \quad (2.46)$$

where  $\mathbf{x} = e^{(\theta-1)t} \mathbf{m}$ . We conclude the proof by noting that

$$\inf_{\mathbf{m} \in \mathbb{Z}^{d-1} - \{\mathbf{0}\}} |({}^\top \mathbf{x}, C |\mathbf{x}|^{-(d-1)(1+\epsilon)})| \geq \inf_{\mathbf{y} \in \mathbb{R}^{d-1} - \{\mathbf{0}\}} |({}^\top \mathbf{y}, C |\mathbf{y}|^{-(d-1)(1+\epsilon)})| > 0. \quad (2.47)$$

$\square$

The fact that  $\epsilon = 0$  implies  $\theta = 0$  is consistent with Dani's correspondence. On the other hand,  $\theta < 1$  for any  $\epsilon < \infty$ .

2.7. **Norm estimates.** Let  $\|\cdot\|$  denote the usual matrix norm

$$\|M\| := \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|M\mathbf{x}\|}{\|\mathbf{x}\|}. \quad (2.48)$$

**Proposition 2.4.** *Choose  $M = M^{(0)}$  as in (2.6), and suppose  $\alpha$  satisfies condition (2.29). Then there are constants  $c_1, c_2, c_3, c_4, c_5, c_6 > 0$  such that for all  $n \in \mathbb{N} \cup \{0\}$*

$$\|M^{(n)}\| \leq c_1 \exp[(d-1)\theta t_n], \quad (2.49)$$

$$\|M^{(n)-1}\| \leq c_2 \exp(\theta t_n), \quad (2.50)$$

$$\|P^{(n)}\| \leq c_3 \exp[(d\theta + 1 - \theta)t_n], \quad (2.51)$$

$$\|P^{(n)-1}\| \leq c_4 \exp[(d-1 + \theta)t_n], \quad (2.52)$$

$$\|T^{(n)}\| \leq c_5 \exp[(1 - \theta)\delta t_n + d\theta t_n], \quad (2.53)$$

$$\|T^{(n)-1}\| \leq c_6 \exp[(d-1)(1 - \theta)\delta t_n + d\theta t_n]. \quad (2.54)$$

*Proof.* For any  $M \in \mathcal{S}_d$  as in (2.22) we have, for all  $0 < a_d \leq 1$ ,

$$\|M\| \ll a_1 = (a_2 \cdots a_d)^{-1} \ll a_d^{-(d-1)}, \quad (2.55)$$

and

$$\|M^{-1}\| \ll a_d^{-1}. \quad (2.56)$$

Combine this with Lemmas 2.2 and 2.3 to obtain the bounds

$$\|M^{(n)-1}\| \ll C_2 \delta(M^{(n)})^{-1} = C_2 \delta(M^{(0)} E^{t_n})^{-1} < C_2 C'^{-1} \exp(\theta t_n) \quad (2.57)$$

and

$$\begin{aligned} \|M^{(n)}\| &\ll C_2^{d-1} \delta(M^{(n)})^{-(d-1)} = C_2^{d-1} \delta(M^{(0)} E^{t_n})^{-(d-1)} \\ &< C_2^{d-1} C'^{-(d-1)} \exp[(d-1)\theta t_n]. \end{aligned} \quad (2.58)$$

The remaining estimates follow immediately from (2.49), (2.50) and the equations

$$P^{(n)} = M^{(n)} E^{-t_n} M^{(0)-1}, \quad T^{(n)} = M^{(n)} E^{-\delta t_n} M^{(n-1)-1}. \quad (2.59)$$

□

**Proposition 2.5.** *Choose  $M = M^{(0)}$  as in (2.6), and suppose  $\alpha$  satisfies condition (2.29). Then there is a constant  $c_7 > 0$  such that for all  $n \in \mathbb{N} \cup \{0\}$ ,*

$$c_7 \exp\left[-\theta\left(\frac{d^2}{1-\theta} - (d-1)\right)t_n\right] \leq |\gamma^{(n)}| \leq c_1 \exp[(d-1)\theta t_n] \quad (2.60)$$

with  $c_1$  as in (2.49).

*Proof.* The upper bound for  $|\gamma^{(n)}|$  follows from (2.49), since  $\gamma^{(n)} = (M^{(n)})_{dd}$  and hence  $|\gamma^{(n)}| \leq \|M^{(n)}\|$ .

From (2.17) and the Diophantine condition (2.29) we have

$$|\gamma^{(n)}| = \gamma \exp[(d-1)t_n] \|\mathbf{p}_{21}^{(n)} \alpha + p_{22}^{(n)}\| > C \exp[(d-1)t_n] |\mathbf{p}_{21}^{(n)}|^{-(d-1)(1+\epsilon)}. \quad (2.61)$$

Since

$$1 + \epsilon = \frac{d-1+\theta}{(d-1)(1-\theta)} \quad (2.62)$$

and

$$|\mathbf{p}_{21}^{(n)}| \leq \|P^{(n)}\| \quad (2.63)$$

the proposition follows from the estimate (2.51). □

**2.8. Hyperbolicity of the transfer matrices.** Let

$$\omega_{\perp}^{(n)} = \{\boldsymbol{\xi} \in \mathbb{R}^d : \boldsymbol{\xi} \cdot \boldsymbol{\omega}^{(n)} = 0\} \quad (2.64)$$

be the orthogonal complement of the vector

$$\boldsymbol{\omega}^{(n)} = \begin{pmatrix} \boldsymbol{\alpha}^{(n)} \\ 1 \end{pmatrix} \in \mathbb{R}^d. \quad (2.65)$$

**Lemma 2.6.** *For all  $\boldsymbol{\xi} \in \omega_{\perp}^{(n-1)}$ ,  $n \in \mathbb{N}$ ,*

$${}^{\top}T^{(n)-1}\boldsymbol{\xi} = \exp(-\delta t_n) {}^{\top}(M^{(n-1)}M^{(n)-1})\boldsymbol{\xi} \quad (2.66)$$

*Proof.* This follows directly from the relation

$$E^{\delta t_n} {}^{\top}M^{(n-1)}\boldsymbol{\xi} = \exp(-\delta t_n) \begin{pmatrix} {}^{\top}A^{(n-1)}\boldsymbol{\xi}' \\ 0 \end{pmatrix} = \exp(-\delta t_n) {}^{\top}M^{(n-1)}\boldsymbol{\xi} \quad (2.67)$$

where  $\boldsymbol{\xi}' \in \mathbb{R}^{d-1}$  comprises the first  $d-1$  components of  $\boldsymbol{\xi}$ .  $\square$

**Proposition 2.7.** *Choose  $M = M^{(0)}$  as in (2.6), and suppose  $\boldsymbol{\alpha}$  satisfies condition (2.29). Then there is a constant  $\Lambda > 0$  such that for all  $\boldsymbol{\xi} \in \omega_{\perp}^{(n-1)}$ ,  $n \in \mathbb{N}$ ,*

$$\|{}^{\top}T^{(n)-1}\boldsymbol{\xi}\| \leq \frac{1}{2}\Lambda \exp(-\varphi_n)\|\boldsymbol{\xi}\| \quad (2.68)$$

with

$$\varphi_n = (1 - \theta)\delta t_n - d\theta t_{n-1}. \quad (2.69)$$

*Proof.* From Lemma 2.6,

$$\|{}^{\top}T^{(n)-1}\boldsymbol{\xi}\| \leq \exp(-\delta t_n)\|M^{(n-1)}\| \|M^{(n)-1}\| \|\boldsymbol{\xi}\|, \quad (2.70)$$

and the proposition follows from the bounds (2.49), (2.50).  $\square$

Given any positive sequence  $\varphi_0, \varphi_1, \dots$ , the values  $t_n$  that solve eq. (2.69) with  $t_0 = 0$  are

$$t_n = \frac{1}{1 - \theta} \sum_{j=1}^n (1 + \beta)^{n-j} \varphi_j. \quad (2.71)$$

where

$$\beta = \frac{d\theta}{1 - \theta}. \quad (2.72)$$

E.g., for constant  $\varphi_n = \varphi > 0$ , we have

$$t_n = \begin{cases} n\varphi & (\theta = 0) \\ \frac{\varphi}{d\theta}[(1 + \beta)^n - 1] & (0 < \theta < 1). \end{cases} \quad (2.73)$$

**2.9. The resonance cone.** As we shall see, a crucial step in our renormalization scheme is to eliminate all far-from-resonance modes in the Fourier series, i.e., all modes labeled by integer vectors outside the cone

$$K^{(n)} = \{\boldsymbol{\xi} \in \mathbb{R}^d : |\boldsymbol{\xi} \cdot \boldsymbol{\omega}^{(n)}| \leq \sigma_n \|\boldsymbol{\xi}\|\} \quad (2.74)$$

for a given  $\sigma_n > 0$ .

**Lemma 2.8.** *Choose  $M = M^{(0)}$  as in (2.6), and suppose  $\boldsymbol{\alpha}$  satisfies condition (2.29). Then*

$$\sup_{\boldsymbol{\xi} \in K^{(n-1)} - \{0\}} \frac{\|{}^{\top}T^{(n)-1}\boldsymbol{\xi}\|}{\|\boldsymbol{\xi}\|} \leq \left[ \frac{\Lambda}{2} + c_6 \sigma_{n-1} e^{d\delta t_n} \right] \exp[-(1 - \theta)\delta t_n + d\theta t_{n-1}], \quad (2.75)$$

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

*Proof.* We write  $\boldsymbol{\xi} = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$ , where

$$\boldsymbol{\xi}_1 = \frac{\boldsymbol{\xi} \cdot \boldsymbol{\omega}^{(n-1)}}{\|\boldsymbol{\omega}^{(n-1)}\|^2} \boldsymbol{\omega}^{(n-1)}, \quad \boldsymbol{\xi}_2 \in \boldsymbol{\omega}_\perp^{(n-1)}. \quad (2.76)$$

Firstly,

$$\|\mathbb{T}T^{(n)-1}\boldsymbol{\xi}_1\| \leq \|\mathbb{T}T^{(n)-1}\| \|\boldsymbol{\xi}_1\| = \|T^{(n)-1}\| \frac{|\boldsymbol{\xi} \cdot \boldsymbol{\omega}^{(n-1)}|}{\|\boldsymbol{\omega}^{(n-1)}\|} \leq \sigma_{n-1} \|T^{(n)-1}\| \|\boldsymbol{\xi}\| \quad (2.77)$$

since  $\boldsymbol{\xi} \in K^{(n-1)}$  and  $\|\boldsymbol{\omega}^{(n-1)}\| = \|(\mathbb{T}\boldsymbol{\alpha}^{(n-1)}, 1)\| \geq 1$ . Hence in view of (2.54)

$$\|\mathbb{T}T^{(n)-1}\boldsymbol{\xi}_1\| \leq c_6 \sigma_{n-1} \exp[(d-1)(1-\theta)\delta t_n + d\theta t_n] \|\boldsymbol{\xi}\|. \quad (2.78)$$

Secondly, from Proposition 2.7 we infer

$$\|\mathbb{T}T^{(n)-1}\boldsymbol{\xi}_2\| \leq \frac{1}{2} \Lambda \exp[-(1-\theta)\delta t_n + d\theta t_{n-1}] \|\boldsymbol{\xi}\|. \quad (2.79)$$

This proves (2.75).  $\square$

**Remark 2.1.** Note that if the  $t_n$  are chosen as in (2.73), and

$$\sigma_{n-1} \leq \frac{1}{2} c_6^{-1} \Lambda \exp(-d\delta t_n), \quad (2.80)$$

then

$$\|\mathbb{T}T^{(n)-1}\boldsymbol{\xi}\| \leq \Lambda \exp(-\varphi) \|\boldsymbol{\xi}\| \quad (2.81)$$

for all  $\boldsymbol{\xi} \in K^{(n-1)}$ ,  $n \in \mathbb{N}$  and  $\varphi > 0$ .

### 3. RENORMALIZATION OF VECTOR FIELDS

**3.1. Definitions.** The transformation of a vector field  $X$  on a manifold  $M$  by a diffeomorphism  $\psi: M \rightarrow M$  is given by the *pull-back* of  $X$  under  $\psi$ :

$$\psi^*X = (D\psi)^{-1}X \circ \psi.$$

As the tangent bundle of the  $d$ -torus is trivial,  $T\mathbb{T}^d \simeq \mathbb{T}^d \times \mathbb{R}^d$ , we identify the set of vector fields on  $\mathbb{T}^d$  with the set of functions from  $\mathbb{T}^d$  to  $\mathbb{R}^d$ , that can be regarded as maps of  $\mathbb{R}^d$  by lifting to the universal cover. We will make use of the analyticity to extend to the complex domain, so we will deal with complex analytic functions. We will also be considering an extra variable related to a parameter.

**Remark 3.1.** We will be using maps between Banach spaces over  $\mathbb{C}$  with a notion of analyticity stated as follows (cf. e.g. [14]): a map  $F$  defined on a domain is analytic if it is locally bounded and Gâteaux 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  $\lim_{k \rightarrow +\infty} F_k = F$  on  $D$ , then  $F$  is analytic on  $D$ .

Let  $\rho, a, b > 0$ ,  $r = (a, b)$  and consider the domain  $D_\rho \times B_r$ , where  $D_\rho = \{\boldsymbol{x} \in \mathbb{C}^d: \|\operatorname{Im} \boldsymbol{x}\| < \rho/2\pi\}$  for the norm  $\|\boldsymbol{u}\| = \sum_i |u_i|$  on  $\mathbb{C}^d$ , and

$$B_r = \left\{ \boldsymbol{y} = (y_1, \dots, y_d) \in \mathbb{C}^d: \sum_{i=1}^{d-1} |y_i| < a \text{ and } |y_d| < b \right\}. \quad (3.1)$$

Take complex analytic functions  $f: D_\rho \times B_r \rightarrow \mathbb{C}^d$  that are  $\mathbb{Z}^d$ -periodic on the first coordinate and on the form of the Fourier series

$$f(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} f_{\boldsymbol{k}}(\boldsymbol{y}) e^{2\pi i \boldsymbol{k} \cdot \boldsymbol{x}}. \quad (3.2)$$

Its coefficients are analytic functions  $f_{\mathbf{k}}: B_r \rightarrow \mathbb{C}^d$  with a continuous extension to the closure  $\overline{B_r}$ , endowed with the sup-norm:

$$\|f_{\mathbf{k}}\|_r = \sup_{\mathbf{y} \in B_r} \|f_{\mathbf{k}}(\mathbf{y})\|.$$

The Banach spaces  $\mathcal{A}_{\rho,r}$  and  $\mathcal{A}'_{\rho,r}$  are the subspaces of such functions such that the respective norms

$$\begin{aligned} \|f\|_{\rho,r} &= \sum_{\mathbf{k} \in \mathbb{Z}^d} \|f_{\mathbf{k}}\|_r e^{\rho\|\mathbf{k}\|}, \\ \|f\|'_{\rho,r} &= \sum_{\mathbf{k} \in \mathbb{Z}^d} (1 + 2\pi\|\mathbf{k}\|) \|f_{\mathbf{k}}\|_r e^{\rho\|\mathbf{k}\|} \end{aligned}$$

are finite. Also, write the constant Fourier mode of  $f \in \mathcal{A}_{\rho,r}$  through the projection

$$\mathbb{E}f(\mathbf{y}) = \int_{\mathbb{T}^d} f(\mathbf{x}, \mathbf{y}) d\mathbf{x} = f_0(\mathbf{y}) \quad (3.3)$$

into the projected space denoted by  $\mathbb{E}\mathcal{A}_r$ . The norm of its derivative  $Df_0$  is given by the operator norm  $\|Df_0\|_r = \sup_{\|g\|_r=1} \|Df_0 g\|_r$ .

Some of the properties of the above spaces are of easy verification. For instance, given any  $f, g \in \mathcal{A}'_{\rho,r}$  we have:

- $\|f(\mathbf{x}, \mathbf{y})\| \leq \|f\|_{\rho,r} \leq \|f\|'_{\rho,r}$  where  $(\mathbf{x}, \mathbf{y}) \in D_\rho \times B_r$ ,
- $\|f\|_{\rho-\delta,r} \leq \|f\|_{\rho,r}$  with  $\delta < \rho$ .

In order to setup notations write, according to section 2,  $\boldsymbol{\omega}^{(0)} = \boldsymbol{\omega} \in \mathbb{R}^d - \{0\}$ ,  $\lambda_0 = 1$  and, for  $n \in \mathbb{N}$ ,

$$\boldsymbol{\omega}^{(n)} = \gamma^{(n)-1} M^{(n)} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \lambda_n P^{(n)} \boldsymbol{\omega} = \eta_n T^{(n)} \boldsymbol{\omega}^{(n-1)}, \quad (3.4)$$

where

$$\lambda_n = \frac{\gamma}{\gamma^{(n)}} e^{(d-1)t_n} \quad \text{and} \quad \eta_n = \frac{\lambda_n}{\lambda_{n-1}}. \quad (3.5)$$

In the following, we will be interested in equilibria-free vector fields with a ‘‘twist’’ along the parameter direction. By rescaling this direction we will find the right parameter which guarantees the conjugacy to a linear flow. For a fixed  $n \in \mathbb{N} \cup \{0\}$ , we will be studying vector fields of the form

$$X(\mathbf{x}, \mathbf{y}) = X_n^0(\mathbf{y}) + f(\mathbf{x}, \mathbf{y}), \quad (\mathbf{x}, \mathbf{y}) \in D_\rho \times B_r, \quad (3.6)$$

where  $f \in \mathcal{A}_{\rho,r}$  and

$$X_n^0(\mathbf{y}) = \boldsymbol{\omega}^{(n)} + \gamma^{(n)-1} M^{(n)} \mathbf{y}. \quad (3.7)$$

(We drop the second coordinate of the vector field because it will always be equal to zero – there is no dynamics along the parameter direction.) The linear transformation on  $\mathbf{y}$  deforms the set  $B_r$  along the directions of the columns of  $M^{(n)}$  (see (2.15)). In particular, its  $d$ th column corresponds to  $\boldsymbol{\omega}^{(n)}$ .

For the space of the above vector fields we use the same notation  $\mathcal{A}_{\rho,r}$  and the same norm  $\|\cdot\|_{\rho,r}$  without ambiguity.

**3.2. Resonance modes.** Given  $\sigma_n > 0$  we define the *far from resonance* Fourier modes  $f_{\mathbf{k}}$  as in (3.2) with respect to  $\boldsymbol{\omega}^{(n)}$  to be the ones whose indices  $\mathbf{k}$  are in the cone

$$I_n^- = \{\mathbf{k} \in \mathbb{Z}^d: |\mathbf{k} \cdot \boldsymbol{\omega}^{(n)}| > \sigma_n \|\mathbf{k}\|\}. \quad (3.8)$$

Similarly, the *resonant* modes correspond to the cone

$$I_n^+ = \mathbb{Z}^d - I_n^-. \quad (3.9)$$

It is also useful to define the projections  $\mathbb{I}_n^+$  and  $\mathbb{I}_n^-$  on  $\mathcal{A}_{\rho,r}$  and  $\mathcal{A}'_{\rho,r}$  by restricting the Fourier modes to  $I_n^+$  and  $I_n^-$ , respectively. The identity operator is  $\mathbb{I} = \mathbb{I}_n^+ + \mathbb{I}_n^-$ .

Moreover, take

$$A_n = \sup_{\mathbf{k} \in I_n^+ - \{0\}} \frac{\|\mathbb{T}^{(n+1)-1} \mathbf{k}\|}{\|\mathbf{k}\|}. \quad (3.10)$$

A useful property of the above cones is included in the Lemma below.

**Lemma 3.2.** *If  $\mathbf{k} \in I_n^-$  and  $\mathbf{y} \in B_{r_n}$  with  $r_n = (a_n, b_n)$ ,*

$$a_n \leq \sigma_n \left( \frac{1}{2} - b_n \right) |\gamma^{(n)}| \|M^{(n)}\|^{-1} \quad \text{and} \quad b_n < \frac{1}{2}, \quad (3.11)$$

then

$$|\mathbf{k} \cdot X_n^0(\mathbf{y})| > \frac{\sigma_n}{2} \|\mathbf{k}\|. \quad (3.12)$$

*Proof.* For every  $\mathbf{y} \in B_{r_n}$  and  $\mathbf{k} \in I_n^-$ ,

$$\begin{aligned} |\mathbf{k} \cdot (\omega^{(n)} + \gamma^{(n)-1} M^{(n)} \mathbf{y})| &= |(1 + y_d) \mathbf{k} \cdot \omega^{(n)} + \gamma^{(n)-1} \mathbf{k} \cdot M^{(n)}(y_1, \dots, y_{d-1}, 0)| \\ &> (1 - b_n) \sigma_n \|\mathbf{k}\| - a_n |\gamma^{(n)}|^{-1} \|M^{(n)}\| \|\mathbf{k}\|. \end{aligned} \quad (3.13)$$

Our choice of  $a_n$  yields (3.12).  $\square$

**3.3. Basis change, time rescaling and reparametrization.** The fundamental step of the renormalization is a transformation of the domain of definition of our vector fields. This is done by a linear change of basis (coming essentially from the multidimensional continued fraction expansion of  $\omega$  – see section 2), a linear rescaling of time because the orbits take longer to cross the new torus, and a change of variables for the parameter  $\mathbf{y}$  in order to deal with the zero mode of the perturbation.

Let  $\rho_{n-1}, a_{n-1}, b_{n-1} > 0$ ,  $r_{n-1} = (a_{n-1}, b_{n-1})$  and consider a vector field

$$X(\mathbf{x}, \mathbf{y}) = X_{n-1}^0(\mathbf{y}) + f(\mathbf{x}, \mathbf{y}), \quad (\mathbf{x}, \mathbf{y}) \in D_{\rho_{n-1}} \times B_{r_{n-1}}, \quad (3.14)$$

with  $f \in \mathcal{A}_{\rho_{n-1}, r_{n-1}}$ . We are interested in the following coordinate and time linear changes:

$$\mathbf{x} \mapsto T^{(n)-1} \mathbf{x}, \quad t \mapsto \eta_n t. \quad (3.15)$$

Notice that negative time rescalings are possible, meaning that we are inverting the direction of time. In addition to (3.15) we will use a transformation on  $\mathbf{y}$ , a map  $\mathbf{y} \mapsto \Phi_n(X)(\mathbf{y})$  depending on  $X$  in a way to be defined later.

Therefore, consider the transformation

$$L_n(\mathbf{x}, \mathbf{y}) = (T^{(n)-1} \mathbf{x}, \Phi_n(X)(\mathbf{y})), \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{C}^{2d}, \quad (3.16)$$

that determines a vector field in the new coordinates as the image of the map

$$X \mapsto \mathcal{L}_n(X) = \eta_n L_n^* X.$$

That is, for  $(\mathbf{x}, \mathbf{y}) \in L_n^{-1} D_{\rho_{n-1}} \times B_{r_{n-1}}$ ,

$$\begin{aligned} \mathcal{L}_n(X)(\mathbf{x}, \mathbf{y}) &= \eta_n T^{(n)} [\omega^{(n-1)} + \gamma^{(n-1)-1} M^{(n-1)} \Phi_n(X)(\mathbf{y}) + f_0 \circ \Phi_n(X)(\mathbf{y})] \\ &\quad + \eta_n T^{(n)} (f - f_0) \circ L_n(\mathbf{x}, \mathbf{y}). \end{aligned} \quad (3.17)$$

In order to eliminate the  $\mathbf{k} = 0$  mode of the perturbation of  $X$  in the new coordinates and to normalise the linear term in  $\mathbf{y}$  to  $\gamma^{(n)-1} M^{(n)} \mathbf{y}$ , using the definitions of  $T^{(n)}$  and  $\eta_n$  we choose

$$\Phi_n(X): \mathbf{y} \mapsto \left( \text{Id} + \gamma^{(n-1)} M^{(n-1)-1} f_0 \right)^{-1} (e^{-d\delta t_n} y_1, \dots, e^{-d\delta t_n} y_{d-1}, y_d), \quad (3.18)$$

if possible. Hence,

$$\mathcal{L}_n(X)(\mathbf{x}, \mathbf{y}) = X_n^0(\mathbf{x}, \mathbf{y}) + \tilde{\mathcal{L}}_n(f - f_0)(\mathbf{x}, \mathbf{y}), \quad (3.19)$$

where

$$\tilde{\mathcal{L}}_n: f \mapsto \eta_n T^{(n)} f \circ L_n. \quad (3.20)$$

Denote by  $\Delta_\mu$  the set of  $X \in \mathcal{A}_{\rho_{n-1}, r_{n-1}}$  such that  $\|f_0\|_{r_{n-1}} < \mu$ .

**Lemma 3.3.** *Let  $r_n = (a_n, b_n)$  and  $\mu_{n-1} > 0$  such that*

$$\begin{aligned} a_n &\leq e^{d\delta t_n} \left[ a_{n-1} - \left( 1 + |\gamma^{(n-1)}| \|M^{(n-1)^{-1}}\| \right) \mu_{n-1} \right] \\ b_n &\leq b_{n-1} - \left( 1 + |\gamma^{(n-1)}| \|M^{(n-1)^{-1}}\| \right) \mu_{n-1}. \end{aligned} \quad (3.21)$$

*There exist an analytic map  $\Phi_n: \Delta_{\mu_{n-1}} \rightarrow \text{Diff}(B_{r_n}, \mathbb{C}^d)$  such that, for each  $X \in \Delta_{\mu_{n-1}}$ ,  $\Phi_n(X)$  is given by (3.18) and*

$$\Phi_n(X)(B_{r_n}) \subset B_{r_{n-1}}. \quad (3.22)$$

*In case  $f_0$  is real-analytic,  $\Phi_n(X)|_{\mathbb{R}^d}$  is also real-valued.*

*Proof.* For  $X \in \mathcal{A}_{\rho_{n-1}, r_{n-1}}$  with  $\|f_0\|_{r_{n-1}} < \mu_{n-1}$  and  $\hat{\delta}_{n-1} = (\delta_{n-1}, \delta_{n-1})$  with

$$\delta_{n-1} = \mu_{n-1} |\gamma^{(n-1)}| \|M^{(n-1)^{-1}}\|, \quad (3.23)$$

we have by the Cauchy estimate

$$\|Df_0\|_{r_{n-1} - \hat{\delta}_{n-1}} \leq \frac{\|f_0\|_{r_{n-1}}}{\delta_{n-1}} < \frac{1}{|\gamma^{(n-1)}| \|M^{(n-1)^{-1}}\|}.$$

So,  $F = \text{Id} + \gamma^{(n-1)} M^{(n-1)^{-1}} f_0$  is a diffeomorphism on  $B_{r_{n-1} - \hat{\delta}_{n-1}}$ . Now, if  $R_1 < a_{n-1} - \delta_{n-1} - \mu_{n-1}$ ,  $R_2 < b_{n-1} - \delta_{n-1} - \mu_{n-1}$  and  $R = (R_1, R_2)$ , we have  $B_R \subset F(B_{r_{n-1} - \hat{\delta}_{n-1}})$  and  $F^{-1}(B_R) \subset B_{r_{n-1} - \hat{\delta}_{n-1}}$ . Therefore,  $\Phi_n(X)$  as given by (3.18) is a diffeomorphism on  $B_{r_n}$  by choosing  $R = (e^{-d\delta t_n} a_n, b_n)$ , and thus we get (3.22). In addition,  $X \mapsto \Phi_n(X)$  is analytic from its dependence on  $f_0$ . When restricted to a real domain for a real-analytic  $f_0$ ,  $\Phi_n(X)$  is also real-analytic.  $\square$

Let the translation  $R_{\mathbf{z}}$  on  $\mathbb{C}^{2d}$  be defined for  $\mathbf{z} \in \mathbb{C}^d$  and given by

$$R_{\mathbf{z}}: (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x} + \mathbf{z}, \mathbf{y}). \quad (3.24)$$

Notice that we have the following ‘‘commutative’’ relation:

$$L_n^* R_{\mathbf{z}}^* = R_{T^{(n)} \mathbf{z}}^* L_n^*, \quad \mathbf{z} \in \mathbb{C}^d. \quad (3.25)$$

This also follows from the fact that  $\Phi_n$  is unchanged by the introduction of the translation  $R_{\mathbf{z}}$ .

### 3.4. Analyticity improvement.

**Lemma 3.4.** *If  $\delta > 0$  and*

$$\rho'_n \leq \frac{\rho_{n-1}}{A_{n-1}} - \delta, \quad (3.26)$$

*then  $\tilde{\mathcal{L}}_n$  as a map from  $(\mathbb{I}_{n-1}^+ - \mathbb{E}) \mathcal{A}_{\rho_{n-1}, r_{n-1}} \cap \Delta_{\mu_{n-1}}$  into  $(\mathbb{I} - \mathbb{E}) \mathcal{A}'_{\rho'_n, r_n}$  is continuous and compact with*

$$\|\tilde{\mathcal{L}}_n\| \leq |\eta_n| \|T^{(n)}\| \left( 1 + \frac{2\pi}{\delta} \right). \quad (3.27)$$

**Remark 3.5.** This result means that every vector field in  $\mathbb{I}_{n-1}^+ \mathcal{A}_{\rho_{n-1}, r_{n-1}} \cap \Delta_{\mu_{n-1}}$ , i.e. a function on  $D_{\rho_{n-1}} \times B_{r_{n-1}}$  into  $\mathbb{C}^d$ , has an analytic extension to  $T^{(n)^{-1}} D_{\rho'_n} \times B_{r_{n-1}}$ .



*Proof.* Let  $f \in (\mathbb{I}_{n-1}^+ - \mathbb{E})\mathcal{A}_{\rho_{n-1}, r_{n-1}} \cap \Delta_{\mu_{n-1}}$ . Then,

$$\|f \circ L_n\|'_{\rho'_n, r_n} \leq \sum_{\mathbf{k} \in I_{n-1}^+ - \{0\}} \left(1 + 2\pi \|{}^{\top}T^{(n)-1} \mathbf{k}\|\right) \|f_{\mathbf{k}} \circ M_n\|_{r_n} e^{(\rho'_n - \delta + \delta) \|{}^{\top}T^{(n)-1} \mathbf{k}\|}. \quad (3.28)$$

By using the relation  $\xi e^{-\delta \xi} \leq \delta^{-1}$  with  $\xi \geq 0$ , (3.10) and (3.22), we get

$$\begin{aligned} \|f \circ L_n\|'_{\rho'_n, r_n} &\leq (1 + 2\pi/\delta) \sum_{I_{n-1}^+ - \{0\}} \|f_{\mathbf{k}}\|_{r_{n-1}} e^{A_{n-1}(\rho'_n + \delta) \|\mathbf{k}\|} \\ &\leq (1 + 2\pi/\delta) \|f\|_{\rho_{n-1}, r_{n-1}}. \end{aligned} \quad (3.29)$$

Finally,  $\|\tilde{\mathcal{L}}_n f\|'_{\rho'_n, r_n} \leq |\eta_n| \|T^{(n)}\| \|f \circ L_n\|'_{\rho'_n, r_n}$ .

The above for  $D_{\rho'_n} \times B_{r_n}$  is also valid for  $D_{\zeta} \times B_{r_n}$ ,  $\zeta > \rho'_n$  but satisfying a similar inequality to (3.26). Therefore,  $\tilde{\mathcal{L}}_n = \mathcal{I} \circ \mathcal{J}$ , where  $\mathcal{J}: (\mathbb{I}_{n-1}^+ - \mathbb{E})\mathcal{A}_{\rho_{n-1}, r_{n-1}} \rightarrow \mathcal{A}'_{\zeta, r_n}$  is bounded as  $\tilde{\mathcal{L}}_n$ , and the inclusion map  $\mathcal{I}: \mathcal{A}'_{\zeta, r_n} \rightarrow \mathcal{A}'_{\rho'_n, r_n}$  is compact.  $\square$

For  $0 < \rho''_n \leq \rho'_n$ , consider the inclusion

$$\mathcal{I}_n: \mathcal{A}'_{\rho'_n, r_n} \rightarrow \mathcal{A}''_{\rho''_n, r_n} \quad (3.30)$$

by restricting  $X \in \mathcal{A}'_{\rho'_n, r_n}$  to the smaller domain  $D_{\rho''_n} \times B_{r_n}$ . When restricted to non-constant modes, its norm can be estimated as follows.

**Lemma 3.6.** *If  $\phi_n \geq 1$  and*

$$0 < \rho''_n \leq \rho'_n - \log(\phi_n), \quad (3.31)$$

*then*

$$\|\mathcal{I}_n(\mathbb{I} - \mathbb{E})\| \leq \phi_n^{-1}. \quad (3.32)$$

*Proof.* For  $f \in (\mathbb{I} - \mathbb{E})\mathcal{A}'_{\rho'_n, r_n}$ , we have

$$\|\mathcal{I}_n(f)\|'_{\rho''_n, r_n} \leq \sum_{\mathbf{k} \neq 0} (1 + 2\pi \|\mathbf{k}\|) \|f_{\mathbf{k}}\|_{r_n} e^{\rho'_n \|\mathbf{k}\|} \phi_n^{-\|\mathbf{k}\|} \leq \phi_n^{-1} \|f\|'_{\rho'_n, r_n}. \quad (3.33)$$

$\square$

**3.5. Elimination of far from resonance modes.** The theorem below (to be proven in Section A) states the existence of a nonlinear change of coordinates  $U$ , isotopic to the identity, that cancels the  $I_n^-$  modes of any  $X$  as in (3.6) with sufficiently small  $f$ . We are eliminating only the far from resonance modes, this way avoiding the complications usually related to small divisors. We remark that the “parameter” direction  $\mathbf{y}$  is not affected by this change of coordinates.

For given  $\rho_n, r_n, \varepsilon, \nu > 0$ , denote by  $\mathcal{V}_\varepsilon$  the open ball in  $\mathcal{A}'_{\rho_n + \nu, r_n}$  centred at  $X_n^0$  with radius  $\varepsilon$ .

**Theorem 3.7.** *Let  $r_n$  be as in (3.11),  $\sigma_n < \|\omega^{(n)}\|$  and*

$$\varepsilon_n = \frac{\sigma_n}{42} \min \left\{ \frac{\nu}{4\pi}, \frac{\sigma_n}{72 \|\omega^{(n)}\|} \right\}. \quad (3.34)$$

*For all  $X \in \mathcal{V}_{\varepsilon_n}$  there exists an isotopy*

$$\begin{aligned} U_t: D_{\rho_n} \times B_{r_n} &\rightarrow D_{\rho_n + \nu} \times B_{r_n}, \\ (\mathbf{x}, \mathbf{y}) &\mapsto (\mathbf{x} + u_t(\mathbf{x}, \mathbf{y}), \mathbf{y}), \end{aligned} \quad (3.35)$$

*of analytic diffeomorphisms with  $u_t$  in  $\mathcal{A}'_{\rho_n, r_n}$ ,  $t \in [0, 1]$ , satisfying*

$$\mathbb{I}_n^- U_t^* X = (1 - t) \mathbb{I}_n^- X, \quad U_0 = \text{Id}. \quad (3.36)$$

This defines the maps

$$\begin{aligned} \mathfrak{U}_t: \mathcal{V}_{\varepsilon_n} &\rightarrow \mathcal{A}'_{\rho_n, r_n} \\ X &\mapsto \text{Id} + u_t \end{aligned} \quad (3.37)$$

and

$$\begin{aligned} \mathcal{U}_t: \mathcal{V}_{\varepsilon_n} &\rightarrow \mathbb{I}^+ \mathcal{A}_{\rho_n, r_n} \oplus (1-t) \mathbb{I}_n^- \mathcal{A}'_{\rho_n + \nu, r_n} \\ X &\mapsto U_t^* X \end{aligned} \quad (3.38)$$

which are analytic, and satisfy the inequalities

$$\begin{aligned} \|\mathfrak{U}_t(X) - \text{Id}\|'_{\rho_n, r_n} &\leq \frac{42t}{\sigma_n} \|\mathbb{I}_n^- f\|_{\rho_n, r_n} \\ \|\mathcal{U}_t(X) - X_n^0\|_{\rho_n, r_n} &\leq (3-t) \|f\|'_{\rho_n + \nu, r_n}. \end{aligned} \quad (3.39)$$

If  $X$  is real-analytic, then  $\mathfrak{U}_t(X)(\mathbb{R}^{2d}) \subset \mathbb{R}^{2d}$ .

**Remark 3.8.** Further on we will be using the above result for  $t = 1$ . So that all far from resonance modes are eliminated.

Recall the definition of the translation  $R_{\mathbf{x}}$  in (3.24).

**Lemma 3.9.** *In the conditions of Theorem 3.7, if  $\mathbf{x} \in \mathbb{R}^d$  and  $X \in \mathcal{V}_{\varepsilon_n}$ , then*

$$\mathfrak{U}_t(X \circ R_{\mathbf{x}}) = R_{\mathbf{x}}^{-1} \circ \mathfrak{U}_t(X) \circ R_{\mathbf{x}} \quad (3.40)$$

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

*Proof.* Notice that  $R_{\mathbf{x}}(D_{\rho_n} \times B_{r_n}) = D_{\rho_n} \times B_{r_n}$ . If  $U_t = \mathfrak{U}_t(X)$  is a solution of the homotopy equation (3.36) on  $D_{\rho_n} \times B_{r_n}$ , then  $\tilde{U}_t = R_{\mathbf{x}}^{-1} \circ \mathfrak{U}_t(X) \circ R_{\mathbf{x}}$  solves the same equation for  $\tilde{X} = X \circ R_{\mathbf{x}}$ , i.e.  $\mathbb{I}_n^- \tilde{X} \circ \tilde{U}_t = (1-t) \mathbb{I}_n^- \tilde{X}$ , on  $D_{\rho_n} \times B_{r_n}$ .  $\square$

**3.6. Trivial limit of renormalization.** Let a sequence of “widths”  $0 < \sigma_n < 1$  of the resonance cones  $I_n^+$  be given. The  $n$ th step renormalization operator is thus

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

where  $\mathcal{U}_n$  is the full elimination of the modes in  $I_n^-$  as in Theorem 3.7 (for  $t = 1$ ). Notice that  $\mathcal{R}_n(X^0 + \mathbf{v}) = X_n^0$ , for every  $\mathbf{v} \in \mathbb{C}^d$ . From the previous sections the map  $\mathcal{R}_n$  on its domain is analytic. Also, in case a vector field  $X$  is real-analytic, the same is true for  $\mathcal{R}_n(X)$ .

Fix the constants  $\nu$  and  $\delta$  as in Theorem 3.7 and Lemma 3.4, respectively, and choose  $0 < \lambda < 1$ . Take

$$\Theta_n = \min \left\{ \varepsilon_n, \frac{\lambda^n \sigma_n^2}{\prod_{i=1}^n \|T^{(i)-1}\|^2}, \lambda^n \frac{\frac{\sigma_n |\gamma^{(n)}|}{\|M^{(n)}\|} - \frac{\sigma_{n+1} |\gamma^{(n+1)}|}{e^{\delta \delta t_{n+1}} \|M^{(n+1)}\|}}{1 + |\gamma^{(n)}| \|M^{(n)-1}\|} \right\} \quad (3.41)$$

by assuming that the sequence of times  $t_n$  guarantees that  $\Theta_n > 0$ . Now, write

$$B_n = \prod_{i=0}^n A_i. \quad (3.42)$$

with  $A_i$  given by (3.10). By recalling the inequalities (3.26) and (3.31) we choose, for a given  $\rho_0 > 0$ ,

$$\rho_n = \frac{1}{B_{n-1}} \left[ \rho_0 - \sum_{i=0}^{n-1} B_i \log(\phi_{i+1}) - (\delta + \nu) \sum_{i=0}^{n-1} B_i \right], \quad (3.43)$$

where

$$\phi_n = \max \left\{ 2|\eta_n| \|T^{(n)}\| (1 + 2\pi\delta^{-1}) \frac{\Theta_{n-1}}{\Theta_n}, 1 \right\} \geq 1 \quad (3.44)$$

is to be used in Lemma 3.6.

Define the following function for every  $\omega \in \mathbb{R}^d$  associated to the choice of  $\sigma_n$ :

$$\mathcal{B}(\omega) = \sum_{i=0}^{+\infty} B_i \log(\phi_{i+1}) + (\delta + \nu) \sum_{i=0}^{+\infty} B_i. \quad (3.45)$$

The convergence of the renormalization scheme now follows directly from our construction.

**Theorem 3.10.** *Suppose that*

$$\mathcal{B}(\omega) < +\infty \quad (3.46)$$

and  $\rho > \mathcal{B}(\omega) + \nu$ . There is  $K, b > 0$  and  $r_n = (a_n, b_n)$  with  $a_n > 0$  and  $b_n > b > 0$ , such that if  $X$  is in a sufficiently small open ball around  $X^0$  in  $\mathcal{A}_{\rho, r_0}$ , then

(i)  $X$  is in the domain of  $\mathcal{R}_n$  and

$$\|\mathcal{R}_n(X) - \mathcal{R}_n(X^0)\|_{\rho_n, r_n} \leq K\Theta_n \|X - X^0\|_{\rho, r_0}, \quad n \in \mathbb{N} \cup \{0\}, \quad (3.47)$$

(ii) for each  $|s| < b$  there exists in  $B_{r_{n-1}} \subset \mathbb{C}^d$  the limits

$$p_n^s(X) = \lim_{m \rightarrow +\infty} \Phi_n(\mathcal{R}_{n-1}(X)) \dots \Phi_m(\mathcal{R}_{m-1}(X))(0, \dots, 0, s) \quad (3.48)$$

and

$$\lim_{n \rightarrow +\infty} \|p_n^s(X) - (0, \dots, 0, s)\| = 0, \quad (3.49)$$

(iii) the map  $X \mapsto p_n(X)$  is analytic and takes any real-analytic  $X$  into an analytic curve  $s \mapsto p_n^s(X)$  in  $\mathbb{R}^d$ .

*Proof.* Let  $\xi > 0$  and  $\rho_0 = \rho - \nu - \xi > 0$  such that  $\rho_0 > \mathcal{B}(\omega)$ . Hence, by (3.43), we have  $R > 0$  satisfying  $\rho_n > RB_{n-1}^{-1}$  for all  $n \in \mathbb{N}$ .

Denote by  $c$  the radius of an open ball in  $\mathcal{A}_{\rho, r_0}$  centred at  $X^0$  and containing  $X$ . If  $c \leq \varepsilon_0$  we can use Theorem 3.7 to obtain  $\mathcal{R}_0(X) \in \mathbb{I}_0^+ \mathcal{A}_{\rho_0, r_0}$  with  $r_0 = (a_0, b_0)$  satisfying (3.11) and

$$\|\mathcal{R}_0(X) - \mathcal{R}_0(X^0)\|_{\rho_0, r_0} \leq 2\|X - X^0\|'_{\rho+\xi, r_0} \leq 2\xi^{-1}\|X - X^0\|_{\rho, r_0}.$$

Let  $K = 2(\xi\Theta_0)^{-1}$  and assume that  $c \leq K^{-1} \min\{b_0(1-\lambda), \frac{1}{2} - b_0\}$ . So, (3.47) holds for  $n = 0$ .

Now, with  $n \in \mathbb{N}$  we choose the following  $r_n$ :

$$a_n = \sigma_n \left( \frac{1}{2} - b_0 \right) \frac{|\gamma^{(n)}|}{\|M^{(n)}\|} \quad \text{and} \quad b_n = b_0 - cK \sum_{i=0}^{n-1} \lambda^i, \quad (3.50)$$

so that  $1/2 > b_n > b = b_0 - cK(1-\lambda)^{-1}$ . The inequalities in (3.11) follow immediately. Moreover, (3.21) is also satisfied with  $\mu_{n-1} = cK\Theta_{n-1}$  because

$$\begin{aligned} a_{n-1} - e^{-d\delta t_n} a_n &\geq \left( \frac{1}{2} - b_0 \right) (1 + |\gamma^{(n-1)}| \|M^{(n-1)}\|^{-1}) \Theta_{n-1} \\ &\geq (1 + |\gamma^{(n-1)}| \|M^{(n-1)}\|^{-1}) cK \Theta_{n-1}, \\ b_{n-1} - b_n &= cK \lambda^{n-1} \\ &\geq (1 + |\gamma^{(n-1)}| \|M^{(n-1)}\|^{-1}) cK \Theta_{n-1}. \end{aligned} \quad (3.51)$$

Suppose that  $X_{n-1} = \mathcal{R}_{n-1}(X) \in \mathbb{I}_{n-1}^+ \mathcal{A}_{\rho_{n-1}, r_{n-1}}$  and

$$\|X_{n-1} - X_{n-1}^0\|_{\rho_{n-1}, r_{n-1}} \leq K\Theta_{n-1} \|X - X^0\|_{\rho, r}.$$

Since (3.21) holds, Lemmas 3.3 and 3.4 are valid and, together with (3.19) and Lemma 3.6, can be used to estimate  $\mathcal{I}_n \circ \mathcal{L}_n(X_{n-1})$ :

$$\begin{aligned} \|\mathcal{I}_n \circ \mathcal{L}_n(X_{n-1}) - X_n^0\|'_{\rho_n, r_n} &\leq |\eta_n| \|T^{(n)}\| (1 + 2\pi\delta^{-1}) \phi_n^{-1} K \Theta_{n-1} \|X - X^0\|_{\rho, r_0} \\ &= \frac{1}{2} K \Theta_n \|X - X^0\|_{\rho, r_0}. \end{aligned} \quad (3.52)$$

This vector field is inside the domain of  $\mathcal{U}_n$  as (3.11) and  $\frac{1}{2}c K \Theta_n < \varepsilon_n$  are satisfied. Thus (3.47) follows from (3.39).

Denote by  $f_0^{(n)}$  the constant mode of the perturbation term of  $X_n$ . By Lemma 3.3,  $\Phi_n(X_{n-1}): B_{r_n} \rightarrow B_{r_{n-1}}$  is given by

$$\mathbf{y} \mapsto (\text{Id} + g_n) \text{diag}(e^{-d\delta t_n}, \dots, e^{-d\delta t_n}, 1) \mathbf{y},$$

where

$$g_n = \left( \text{Id} + \gamma^{(n-1)} M^{(n-1)-1} f_0^{(n-1)} \right)^{-1} - \text{Id} \quad (3.53)$$

is defined on  $B_{r'_n}$  with  $r'_n = (e^{-d\delta t_n} a_n, b_n)$ . So, for  $\mathbf{z} \in B_{r'_n}$  there is  $\boldsymbol{\xi} \in B_{r'_n}$  such that

$$\begin{aligned} g_n(\mathbf{z}) &= [I + \gamma^{(n-1)} M^{(n-1)-1} Df_0^{(n-1)}(\boldsymbol{\xi})]^{-1} [\mathbf{z} - \gamma^{(n-1)} M^{(n-1)-1} f_0^{(n-1)}(0)] - \mathbf{z} \\ &= -[I + \gamma^{(n-1)} M^{(n-1)-1} Df_0^{(n-1)}(\boldsymbol{\xi})]^{-1} \gamma^{(n-1)} M^{(n-1)-1} [Df_0^{(n-1)}(\boldsymbol{\xi}) \mathbf{z} + f_0^{(n-1)}(0)] \end{aligned} \quad (3.54)$$

and

$$\|g_n\|_{r'_n} \leq \frac{|\gamma^{(n-1)}| \|M^{(n-1)-1}\|}{1 - |\gamma^{(n-1)}| \|M^{(n-1)-1}\| \|Df_0^{(n-1)}\|_{r'_n}} \left( \|r'_n\| \|Df_0^{(n-1)}\|_{r'_n} + \|f_0^{(n-1)}\|_{r'_n} \right). \quad (3.55)$$

The choice of  $r_n$  means that

$$\min\{a_{n-1} - e^{-d\delta t_n} a_n, b_{n-1} - b_n\} \gg \min\left\{ \frac{\sigma_{n-1} |\gamma^{(n-1)}|}{\|M^{(n-1)}\|}, \lambda^{n-1} \right\}. \quad (3.56)$$

By using (3.41) and the Cauchy estimate,

$$\|Df_0^{(n-1)}\|_{r'_n} \leq \frac{\|f_0^{(n-1)}\|_{r_{n-1}}}{\min\{a_{n-1} - e^{-d\delta t_n} a_n, b_{n-1} - b_n\}} \ll \frac{\lambda^{n-1}}{|\gamma^{(n-1)}| \|M^{(n-1)-1}\|} \quad (3.57)$$

Thus,

$$\|g_n\|_{r'_n} \ll \lambda^{n-1}. \quad (3.58)$$

Writing  $\mathbf{y}_s = (0, \dots, 0, s)$ , by induction we have

$$\Phi_n(X_{n-1}) \dots \Phi_m(X_{m-1})(\mathbf{y}_s) = \mathbf{y}_s + \sum_{i=n}^m \text{diag}(e^{-d(t_{i-1}-t_{n-1})}, \dots, e^{-d(t_{i-1}-t_{n-1})}, 1) g_i(\boldsymbol{\xi}_i), \quad (3.59)$$

for some  $\boldsymbol{\xi}_k \in B_{r'_k}$ . Therefore, from (3.58), there exists  $p_n^s(X) \in \mathbb{C}^d$  unless  $X$  is real which clearly gives  $p_n^s(X) \in \mathbb{R}^d$ . In addition,

$$\|p_n^s(X) - \mathbf{y}_s\| \leq \sum_{i=n}^{+\infty} \|g_i\|_{r'_i} \ll \frac{\lambda^{n-1}}{1 - \lambda}. \quad (3.60)$$

The maps  $X \mapsto p_n^s(X)$  are analytic since the convergence is uniform. Lemma 3.3 gives us the nested sequence  $\Phi_n(X_{n-1})(B_{r_n}) \subset B_{r_{n-1}}$ . So, as  $\mathbf{y}_s \in \cap_{i \in \mathbb{N}} B_{r_i}$ , it follows that  $p_n^s(X) \in B_{r_{n-1}}$ .  $\square$

**Remark 3.11.** The above can be generalised for a small analyticity radius  $\rho$  by considering a sufficiently large  $N$  and applying the above theorem to  $\tilde{X} = \mathcal{U}_N \mathcal{L}_N \dots \mathcal{U}_1 \mathcal{L}_1 \mathcal{U}_0(X)$ , where  $X$  is close enough to  $X^0$ . We recover the large strip case since  $\rho_N$  is of the order of  $B_{N-1}^{-1}$ . It remains to check that  $\rho_N > \mathcal{B}(\boldsymbol{\omega}^{(N)}) + \nu$ . This follows from the fact that  $\mathcal{B}(\boldsymbol{\omega}^{(N)}) = B_{N-1}^{-1}[\mathcal{B}(\boldsymbol{\omega}) - \mathcal{B}_N(\boldsymbol{\omega})]$  where  $\mathcal{B}_N(\boldsymbol{\omega})$  is the sum of the first  $N$  terms of  $\mathcal{B}(\boldsymbol{\omega})$  so that  $\mathcal{B}_N(\boldsymbol{\omega}) \rightarrow \mathcal{B}(\boldsymbol{\omega})$  as  $N \rightarrow +\infty$ .

**Lemma 3.12.** *If  $\boldsymbol{\omega} = \begin{pmatrix} \alpha \\ 1 \end{pmatrix}$  in  $\mathbb{R}^d$  is diophantine, i.e.  $\alpha$  satisfies (2.29) with exponent  $\epsilon$  (related to  $\theta$  by (2.43) and to  $\beta$  by (2.72)), then (3.46) is verified.*

*Proof.* Let us set  $\delta t_n = \xi t_{n-1}$ ,  $\sigma_n = \exp(-c\delta t_n)$ ,  $n \geq 1$ , where positive constants  $\xi, c$  will be chosen later. Obviously,  $t_n = (1 + \xi)t_{n-1} = [(1 + \xi)/\xi]\delta t_n$  and  $\delta t_n = (1 + \xi)\delta t_{n-1}$ . We shall assume that

$$c < d(1 + \xi), \quad (3.61)$$

so that  $\sigma_{n-1} \exp(d\delta t_n) = \exp(-c\delta t_{n-1} + d\delta t_n) = \exp[(d - c/(1 + \xi))\delta t_n]$  is much larger than  $\Lambda$  given by Proposition 2.7. Hence, using (2.75) we have

$$A_{n-1} \ll \exp \left[ \left( -\frac{c}{1 + \xi} + d - (1 - \theta) + \frac{d\theta}{\xi} \right) \delta t_n \right]. \quad (3.62)$$

We next estimate  $\|\boldsymbol{\omega}^{(n)}\|$  and  $\varepsilon_n$ . It follows from (3.4) that  $\|\boldsymbol{\omega}^{(n)}\| \ll \|M^{(n)}\| |\gamma^{(n)}|^{-1}$ . Thus, using (2.49), (2.60) we have

$$\|\boldsymbol{\omega}^{(n)}\| \ll \exp \left( \frac{\theta}{1 - \theta} d^2 t_n \right) = \exp \left( \frac{\theta}{1 - \theta} d^2 \frac{1 + \xi}{\xi} \delta t_n \right). \quad (3.63)$$

Since  $\|\boldsymbol{\omega}^{(n)}\| \geq 1$  one gets from (3.34) that  $\varepsilon_n \sim \sigma_n^2 / \|\boldsymbol{\omega}^{(n)}\|$  which together with (3.63) implies

$$\exp \left[ \left( -2c - \frac{\theta}{1 - \theta} d^2 \frac{1 + \xi}{\xi} \right) \delta t_n \right] \ll \varepsilon_n \ll \exp(-2c\delta t_n). \quad (3.64)$$

Here  $X \sim Y$  means that there exist two positive constants  $C_1, C_2 > 0$  such that  $C_1 Y < X < C_2 Y$ . Using again (2.49), (2.60) we get

$$\frac{\sigma_n |\gamma^{(n)}|}{\|M^{(n)}\|} \gg \exp \left[ \left( -c - \frac{\theta}{1 - \theta} d^2 \frac{1 + \xi}{\xi} \right) \delta t_n \right]. \quad (3.65)$$

Also, since  $\|M^{(n+1)}\| \geq \gamma^{(n+1)}$ ,

$$\frac{\sigma_{n+1} |\gamma^{(n+1)}|}{\|M^{(n+1)}\| \exp(d\delta t_{n+1})} \ll \exp[-(c + d)(1 + \xi)\delta t_n]. \quad (3.66)$$

We shall assume that  $c$  and  $\xi$  are chosen in such a way that

$$-c - \frac{\theta}{1 - \theta} d^2 \frac{1 + \xi}{\xi} > -(c + d)(1 + \xi), \quad (3.67)$$

so that

$$\frac{\sigma_n |\gamma^{(n)}|}{\|M^{(n)}\|} - \frac{\sigma_{n+1} |\gamma^{(n+1)}|}{\|M^{(n+1)}\| \exp(dt_{n+1})} \gg \frac{\sigma_n |\gamma^{(n)}|}{\|M^{(n)}\|}. \quad (3.68)$$

Inequality (3.67) is equivalent to the following condition

$$c > \frac{\theta}{1 - \theta} \frac{1 + \xi}{\xi^2} d^2 - \frac{1 + \xi}{\xi} d. \quad (3.69)$$

Finally, we want  $A_n$  to be small and, hence, require the exponent in (3.62) to be negative

$$-\frac{c}{1 + \xi} + d - (1 - \theta) + \frac{d\theta}{\xi} < 0. \quad (3.70)$$

Suppose that conditions (3.61), (3.69), (3.70) are satisfied. It follows immediately from the estimates above and (3.5), (2.49), (2.50), (2.54), (2.60) that

$$|\log \Theta_n|, |\log \Theta_{n-1}|, \log \|T^{(n)}\|, |\log |\eta_n|| \ll \delta t_n. \quad (3.71)$$

At the same time

$$B_n = \prod_{i=0}^n A_i \ll C^n \exp(-\alpha t_{n+1}), \quad (3.72)$$

where

$$\alpha = \frac{c}{1+\xi} - d + (1-\theta) - \frac{d\theta}{\xi} > 0. \quad (3.73)$$

Since  $B_n$  decays exponentially with  $t_n$  and  $\log \phi_n$  grows at most linearly the series (3.46) converges. To finish the proof it is enough to show that conditions (3.61), (3.69), (3.70) can be satisfied. Indeed, since  $0 < \theta < 1$  we can choose  $\xi$  so large that  $1 - \theta - d\theta/\xi > 0$  and

$$\frac{\theta}{1-\theta} \frac{1+\xi}{\xi^2} d^2 - \frac{1+\xi}{\xi} d < 0. \quad (3.74)$$

It is easy to see that all three inequalities (3.61), (3.69), (3.70) are satisfied if  $(1+\xi)(d-\beta) < c < (1+\xi)d$ , where  $\beta = 1 - \theta - d\theta/\xi > 0$ .  $\square$

**3.7. Analytic conjugacy to linear flow.** As a consequence of Theorem 3.10, we obtain an analytic conjugacy between a vector field and the linear flow, thus proving Theorem 1.1. In the following we always assume to be in the conditions of Section 3.6.

Let  $r = r_0$  and

$$\Delta = \{X \in \mathcal{A}_{\rho,r} : \|X - X^0\|_{\rho,r} < c\} \quad (3.75)$$

inside the domain of  $\mathcal{R}_n$  for all  $n \in \mathbb{N} \cup \{0\}$ . By taking  $X \in \Delta$ , we denote  $X_n = \mathcal{R}_n(X) \in \mathbb{I}_n^+ \mathcal{A}_{\rho_n, r_n}$  so that

$$X_n = \lambda_n (U_0 \circ L_1 \circ U_1 \cdots L_n \circ U_n)^*(X), \quad (3.76)$$

where  $U_k = \mathfrak{U}_k(\mathcal{I}_k \mathcal{L}_k(X_{k-1}))$  is given by Theorem 3.7 for  $t = 1$  at the  $k$ th step and  $L_k$  is the linear rescaling as in (3.16) for  $X_{k-1}$ .

Denote by  $V_n$  the coordinate change

$$V_n : (\mathbf{x}, \mathbf{y}) \mapsto (P^{(n)})^{-1} \mathbf{x}, \Phi_1(X_0) \dots \Phi_n(X_{n-1})(\mathbf{y}) \quad (3.77)$$

and set  $V_0 = \text{Id}$ . Thus,  $L_n = V_{n-1}^{-1} \circ V_n$  and

$$X_n = \lambda_n (V_n \circ U_n)^*(V_{n-1} \circ U_{n-1} \circ V_{n-1}^{-1})^* \cdots (V_1 \circ U_1 \circ V_1^{-1})^* U_0^*(X). \quad (3.78)$$

In particular, the  $\mathbf{y}$ -coordinate is only transformed by the second component of  $V_n$ .

Notice that if  $X_n = X_n^0$  for some  $n \in \mathbb{N}$ ,

$$\mathbf{y} = \Phi_1(X_0) \dots \Phi_n(X_{n-1})(0, \dots, 0, s) \in \mathbb{C}^d,$$

with  $|s| < b$ , corresponds to the parameter for which  $X$  is conjugated to  $(1+s)\omega^{(n)}$ . The parameter value for the general case  $X_n - X_n^0 \rightarrow 0$  as  $n \rightarrow +\infty$  is  $p^s(X) = p_1^s(X)$ .

**Lemma 3.13.** *There is an open ball  $B$  about  $X^0$  in  $\Delta$  such that we can find a sequence  $R_n > 0$  satisfying  $R_{-1} = \rho$ ,*

$$R_n + 2\pi 42K \Theta_n^{1/2} \|X - X^0\|_{\rho,r} \leq R_{n-1} \leq \frac{\rho_{n-1}}{\|P^{(n-1)}\|}, \quad X \in B, \quad (3.79)$$

and

$$\lim_{n \rightarrow +\infty} R_n^{-1} \Theta_n^{1/2} = 0. \quad (3.80)$$

*Proof.* Let  $\rho_* = \min \rho_n$ . It is enough to check that  $\Theta_n^{1/2} \ll \lambda^n \rho_* \prod_{i=1}^n \|T^{(i)}\|^{-1}$  with  $0 < \lambda < 1$  and taking  $R_n = c\lambda^{-n}\Theta_n^{1/2}$  for some positive constant  $c$ . This immediately implies (3.80) and (3.79) by considering a small enough upper bound for  $\|X - X^0\|_{\rho,r}$ .  $\square$

Let  $\text{Diff}_{\text{per}}(D_\zeta, \mathbb{C}^d)$ ,  $\zeta > 0$ , be the Banach space of  $\mathbb{Z}^d$ -periodic diffeomorphisms  $g: D_\zeta \rightarrow \mathbb{C}^d$  with finite norm  $\|g\|_\zeta = \sum_{\mathbf{k}} \|g_{\mathbf{k}}\| e^{\zeta\|\mathbf{k}\|}$ , where  $g_{\mathbf{k}} \in \mathbb{C}^d$  are the coefficients of the Fourier representation. It is simple to check that  $\|g \circ P^{(n)}\|_{R_n} \leq \|g\|_{\rho_n}$ .

Denote by  $u_n$  the analytic function

$$\begin{aligned} u_n: \Delta &\rightarrow \text{Diff}_{\text{per}}(D_{\rho_n}, \mathbb{C}^d) \\ X &\mapsto \mathfrak{U}_n(\mathcal{I}_n \mathcal{L}_n(X_{n-1}))(\cdot, p_{n+1}^s(X)). \end{aligned} \quad (3.81)$$

As  $p_{n+1}^s(X) \in B_{r_n}$ ,  $D_{\rho_n} \times \{\mathbf{y} = p_{n+1}^s(X)\}$  is inside the domain  $\mathcal{D}_{\rho_n, r_n}$  of  $\mathfrak{U}_n(\mathcal{I}_n \mathcal{L}_n(X_{n-1}))$  given in Theorem 3.7. Now, for each  $X$ , define the isotopic to the identity diffeomorphism

$$W_n(X) = P^{(n)-1} \circ u_n(X) \circ P^{(n)}, \quad (3.82)$$

on  $P^{(n)-1}D_{\rho_n}$ . If  $X$  is real-analytic, then  $W_n(X)(\mathbb{R}^d) \subset \mathbb{R}^d$ , since this property holds for  $u_n(X)$ . We also have  $W_n(X^0) = \text{Id}$ .

**Lemma 3.14.** *For all  $n \in \mathbb{N} \cup \{0\}$ ,  $W_n: B \rightarrow \text{Diff}_{\text{per}}(D_{R_n}, \mathbb{C}^d)$  is analytic satisfying  $W_n(X): D_{R_n} \rightarrow D_{R_{n-1}}$  and*

$$\|W_n(X) - \text{Id}\|_{R_n} \leq 42K\Theta_n^{1/2}\|X - X^0\|_{\rho,r}, \quad X \in B. \quad (3.83)$$

*Proof.* For any  $X \in \Delta$ , in view of (3.39) we get

$$\begin{aligned} \|W_n(X) - \text{Id}\|_{R_n} &= \|P^{(n)-1} \circ [u_n(X) - \text{Id}] \circ P^{(n)}\|_{R_n} \\ &\leq \frac{42}{\sigma_n} \|P^{(n)-1}\| \|\mathcal{I}_n \mathcal{L}_n(X_{n-1}) - X_n^0\|_{\rho_n, r_n}. \end{aligned}$$

We can bound the above by (3.83).

Now, for  $\mathbf{x} \in D_{R_n}$  and  $X \in B \subset \Delta$ ,

$$\begin{aligned} \|\text{Im } W_n(X)(\mathbf{x})\| &\leq \|\text{Im}(W_n(X)(\mathbf{x}) - \mathbf{x})\| + \|\text{Im } \mathbf{x}\| \\ &< \|W_n(X) - \text{Id}\|_{R_n} + R_n/2\pi \leq R_{n-1}/2\pi. \end{aligned}$$

So we have  $W_n(X): D_{R_n} \rightarrow D_{R_{n-1}}$  and  $W_n(X) \in \text{Diff}_{\text{per}}(D_{R_n}, \mathbb{C}^d)$ . From the properties of  $\mathfrak{U}_n$ ,  $W_n$  is analytic as a map from  $B$  into  $\text{Diff}_{\text{per}}(D_{R_n}, \mathbb{C}^d)$ .  $\square$

Consider the analytic map  $H_n: B \rightarrow \text{Diff}_{\text{per}}(D_{R_n}, \mathbb{C}^d)$  defined by the coordinate transformation  $H_n(X): D_{R_n} \rightarrow D_{\rho_0}$  as

$$H_n(X) = W_0(X) \circ \cdots \circ W_n(X). \quad (3.84)$$

**Lemma 3.15.** *There exists  $c > 0$  such that for  $X \in B$  and  $n \in \mathbb{N}$ ,*

$$\|H_n(X) - H_{n-1}(X)\|_{R_n} \leq c\Theta_n^{1/2}\|X - X^0\|_{\rho,r}. \quad (3.85)$$

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

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

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

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

If the image of  $D_{R_n}$  under  $\text{Id} + G_{k+1}(z, X)$  is inside the domain of  $W_k(X)$ , or simply

$$\|G_{k+1}(z, X)\|_{R_n} \leq (R_k - R_n)/2\pi,$$

then  $G_k$  is well-defined as an analytic map into  $\text{Diff}_{per}(D_{R_n}, \mathbb{C}^d)$ , and

$$\|G_k(z, X)\|_{R_n} \leq \|W_k(X) - \text{Id}\|_{R_k} + \|G_{k+1}(z, X)\|_{R_n}.$$

An inductive scheme shows that

$$\begin{aligned} \|G_n(z, X)\|_{R_n} &\leq (R_{n-1} - R_n)/2\pi, \\ \|G_k(z, X)\|_{R_n} &\leq \sum_{i=k}^{n-1} \|W_i(X) - \text{Id}\|_{R_i} + |z| \|W_n(X) - \text{Id}\|_{R_n} \\ &\leq (R_{k-1} - R_n)/2\pi. \end{aligned}$$

By Cauchy's formula

$$\begin{aligned} \|H_n(X) - H_{n-1}(X)\|_{R_n} &= \|G_0(1, X) - G_0(0, X)\|_{R_n} \\ &= \left\| \frac{1}{2\pi i} \oint_{|z|=1+d_n/2} \frac{G_0(z, X)}{z(z-1)} dz \right\|_{R_n}, \end{aligned}$$

and

$$\begin{aligned} \|H_n(X) - H_{n-1}(X)\|_{R_n} &\leq \frac{2}{d_n} \sup_{|z|=1+d_n/2} \|G_0(z, X)\|_{R_n} \\ &\ll \Theta_n^{1/2} \|X - X^0\|_{\rho, r}. \end{aligned}$$

□

Consider  $C_{per}^1(\mathbb{R}^d, \mathbb{C}^d)$  to be the Banach space of the  $\mathbb{Z}^d$ -periodic  $C^1$  functions between  $\mathbb{R}^d$  and  $\mathbb{C}^d$  with norm

$$\|f\|_{C^1} = \max_{k \leq 1} \max_{\mathbf{x} \in \mathbb{R}^d} \|D^k f(\mathbf{x})\|. \quad (3.86)$$

**Lemma 3.16.** *There exists  $C > 0$ , an open ball  $B' \subset B$  about  $X^0$  and an analytic map  $H: B' \rightarrow \text{Diff}_{per}(\mathbb{R}^d, \mathbb{C}^d)$  such that for  $X \in B'$ ,  $H(X) = \lim_{n \rightarrow +\infty} H_n(X)$  and*

$$\|H(X) - \text{Id}\|_{C^1} \leq C \|X - X^0\|_{\rho, r}. \quad (3.87)$$

If  $X \in B'$  is real-analytic, then  $H(X) \in \text{Diff}_{per}(\mathbb{R}^d, \mathbb{R}^d)$ .

*Proof.* As the domains  $D_{R_n}$  are shrinking, we consider the restrictions of  $W_n(X)$  and  $H_n(X)$  to  $\mathbb{R}^d$ , and estimate their  $C^1$  norms from the respective norms in  $\text{Diff}_{per}(D_{R_n}, \mathbb{C}^d)$ . More precisely, for any  $X \in B$ , making use of Lemma 3.15,

$$\begin{aligned} \|H_n(X) - H_{n-1}(X)\|_{C^1} &\leq \max_{k \leq 1} \sup_{\mathbf{x} \in D_{R_n/2}} \|D^k [H_n(X)(\mathbf{x}) - H_{n-1}(X)(\mathbf{x})]\| \\ &\leq \frac{4}{R_n} \|H_n(X) - H_{n-1}(X)\|_{R_n}, \end{aligned} \quad (3.88)$$

which goes to zero by (3.80). Notice that here we have used Cauchy's estimate  $\|D^1 g\|_{\zeta} \leq (2\pi/e\delta) \|g\|_{\zeta+\delta}$  with  $\zeta, \delta > 0$ .

Therefore, it is shown the existence of the limit  $H_n(X) \rightarrow H(X)$  as  $n \rightarrow +\infty$ , in the Banach space  $C_{per}^1(\mathbb{R}^d, \mathbb{C}^d)$ . Moreover,  $\|H(X) - \text{Id}\|_{C^1} \ll \|X - X^0\|_{\rho, r}$ . The convergence of  $H_n$  is uniform in  $B$  so  $H$  is analytic. As the space of close to identity diffeomorphisms is closed for the  $C^1$  norm,  $H(X)$  is a diffeomorphism for any  $X$  sufficiently close to  $X^0$ , i.e.  $X \in B'$ . The fact that, for real-analytic  $X$ ,  $H(X)$  takes real values for real arguments, follows from the same property of each  $W_n(X)$ . □

To simplify notation, write  $\pi_{\mathbf{y}} X = X(\cdot, \mathbf{y})$ .

**Lemma 3.17.** *For every real-analytic  $X \in B'$  and  $|s| < b$ ,  $[H(X)]^*(\pi_{p^s(X)} X) = (1+s)\omega$  on  $\mathbb{R}^d$ .*



*Proof.* For each  $n \in \mathbb{N}$  the definition of  $H_n(X)$  and (3.78) imply that

$$H_n(X)^*(\pi_{p^s(X)}X) = \lambda_n^{-1}\pi_{p^s(X)}V_n^{-1*}(X_n). \quad (3.89)$$

The r.h.s. can be written as

$$\begin{aligned} & \lambda_n^{-1}P^{(n)-1}[\boldsymbol{\omega} + \gamma^{(n)-1}M^{(n)}\Phi_n(X_{n-1})^{-1} \cdots \Phi_1(X_0)^{-1}p^s(X)] + \lambda_n^{-1}\pi_{p^s(X)}V_n^{-1*}(X_n - X_n^0) = \\ & = (1+s)\boldsymbol{\omega} + \lambda_n^{-1}P^{(n)-1}\gamma^{(n)-1}M^{(n)}p_{n+1}^s(X) - s\boldsymbol{\omega} + \lambda_n^{-1}\pi_{p^s(X)}V_n^{-1*}(X_n - X_n^0). \end{aligned} \quad (3.90)$$

Its terms can be estimated, for  $\mathbf{x} \in \mathbb{R}^d$ , by

$$\|\lambda_n^{-1}\pi_{p^s(X)}V_n^{-1*}(X_n - X_n^0)(\mathbf{x})\| \leq |\lambda_n^{-1}| \|P^{(n)-1}\| \|X_n - X_n^0\|_{\rho_n, r_n} \ll \Theta_n^{1/2}, \quad (3.91)$$

and

$$\begin{aligned} \frac{1}{|\lambda_n\gamma^{(n)}|} \|P^{(n)-1}M^{(n)}[p_{n+1}^s(X) - (0, \dots, 0, s)]\| &= \frac{\|M^{(0)}\|}{|\gamma|e^{(d-1)t_n}} \|E^{t_n}[p_{n+1}^s(X) - (0, \dots, 0, s)]\| \\ &\ll \|p_{n+1}^s(X) - (0, \dots, 0, s)\| \end{aligned} \quad (3.92)$$

which is controlled by (3.49). Consequently, the limit of (3.89) as  $n \rightarrow +\infty$  is  $(1+s)\boldsymbol{\omega}$ . Using the convergence of  $H_n$  we complete the proof.  $\square$

**Lemma 3.18.** *If  $X \in B'$  and  $\mathbf{x} \in \mathbb{R}^d$ , then*

$$H(X \circ R_{\mathbf{x}}) = \widehat{R}_{\mathbf{x}}^{-1} \circ H(X) \circ \widehat{R}_{\mathbf{x}}, \quad (3.93)$$

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

*Proof.* For each  $n \in \mathbb{N}$ , (3.40) and (3.25) yield that  $\mathcal{U}_n(X \circ R_{\mathbf{x}}) = \mathcal{U}_n(X) \circ R_{\mathbf{x}}$  and  $\mathcal{L}_n(X \circ R_{\mathbf{x}}) = \mathcal{L}_n(X) \circ R_{T^{(n)}\mathbf{x}}$ . This implies immediately that

$$\mathcal{R}_n(X \circ R_{\mathbf{x}}) = \mathcal{R}_n(X) \circ R_{P^{(n)}\mathbf{x}}. \quad (3.94)$$

Next, from a simple adaptation of (3.40) and the formula  $\widehat{R}_{P^{(n)}\mathbf{z}} = P^{(n)}\widehat{R}_{\mathbf{z}}P^{(n)-1}$  for  $\mathbf{z} \in \mathbb{C}^d$ , we get

$$\begin{aligned} W_n(X \circ R_{\mathbf{x}}) &= P^{(n)-1} \circ \mathfrak{U}_{a_n}(\mathcal{L}_n \mathcal{R}_{n-1}(X \circ R_{\mathbf{x}})) \circ P^{(n)} \\ &= \widehat{R}_{\mathbf{x}}^{-1} \circ W_n(X) \circ \widehat{R}_{\mathbf{x}}. \end{aligned} \quad (3.95)$$

Thus,  $H_n(X \circ R_{\mathbf{x}}) = \widehat{R}_{\mathbf{x}}^{-1} \circ H_n(X) \circ \widehat{R}_{\mathbf{x}}$ . The convergence of  $H_n$  implies (3.93).  $\square$

**Theorem 3.19.** *If  $\mathbf{v} \in \text{Vect}^\omega(\mathbb{T}^d)$  is sufficiently close to  $\boldsymbol{\omega}$ , then there exists an analytic curve  $\mathbf{p}: (-b, b) \rightarrow \mathbb{R}^d$ ,  $s \mapsto \mathbf{p}^s$ , and  $h \in \text{Diff}^\omega(\mathbb{T}^d)$  homotopic to the identity such that*

$$h^*(\mathbf{v} + \mathbf{p}^s) = (1+s)\boldsymbol{\omega}. \quad (3.96)$$

*The maps  $\mathbf{v} \mapsto \mathbf{p}$  and  $\mathbf{v} \mapsto h$  are analytic.*

*Proof.* The lift  $\tilde{\mathbf{v}}$  to  $\mathbb{R}^d$  of  $\mathbf{v}$  is assumed to have an analytic extension in  $D_\rho$ . Consider the real-analytic vector field  $Y(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{v}}(\mathbf{x}) + \mathbf{y}$  in  $\mathcal{A}_{\rho, r}$ . Suppose that  $\mathbf{v}$  is close enough to  $\boldsymbol{\omega}$  such that  $Y \in B'$  and  $Y \circ R_{\mathbf{z}} \in B'$  for some  $\eta > 0$  and  $\mathbf{z} \in D_\eta$ . Then, the parameter  $\mathbf{p}^s = p^s(Y) \in \mathbb{R}^d$  and the  $C^1$ -diffeomorphism  $h = H(Y) \bmod 1$  verify (3.96).

We now want to extend  $h$  analytically to a complex neighbourhood of its domain. Take  $\tilde{h}(\mathbf{z}) = \mathbf{z} + H(Y \circ R_{\mathbf{z}})(0)$ ,  $\mathbf{z} \in D_\eta$ . The maps  $\mathbf{z} \mapsto Y \circ R_{\mathbf{z}}$  and  $X \mapsto H(X)$  are analytic and  $C_{per}^1(\mathbb{R}^d, \mathbb{C}^d) \ni g \mapsto g(0)$  is bounded. As  $\tilde{h}$  involves their composition, it is analytic on the domain  $D_\eta$  and  $\mathbb{Z}^d$ -periodic. From (3.93), for any  $\mathbf{x} \in \mathbb{R}^d$ , we have

$$\begin{aligned} \tilde{h}(\mathbf{x}) \bmod 1 &= (\mathbf{x} + \widehat{R}_{\mathbf{x}}^{-1} \circ H(Y) \circ \widehat{R}_{\mathbf{x}}(0)) \bmod 1 \\ &= (\mathbf{x} + H(Y)(\mathbf{x}) - \mathbf{x}) \bmod 1 \\ &= h(\mathbf{x}). \end{aligned} \quad (3.97)$$

The extension of  $h$  is a complex analytic diffeomorphism, thus  $h$  is a real-analytic diffeomorphism.  $\square$

## APPENDIX A. HOMOTOPY METHOD FOR VECTOR FIELDS

In this section we prove Theorem 3.7 using the homotopy method (cf. [22]). As  $n$  is fixed, we will drop it from our notations. In addition we write  $\rho' = \rho_n$  and  $\rho = \rho_n + \nu$ . We will be using the symbol  $D_{\mathbf{x}}$  for the derivative with respect to  $\mathbf{x}$ .

Firstly, we include a technical lemma that will be used in the following.

**Lemma A.1.** *Let  $f \in \mathcal{A}'_{\rho,r}$ . If  $U = \text{Id} + (u, 0)$  where  $u: D_{\rho'} \times B_r \rightarrow D_{(\rho-\rho')/2}$  is in  $\mathcal{A}'_{\rho',r}$  and  $\|u\|_{\rho',r} < (\rho - \rho')/4\pi$ , then*

- $\|f \circ U\|_{\rho',r} \leq \|f\|_{(\rho+\rho')/2,r}$ ,
- $\|D_{\mathbf{x}}f \circ U\| \leq \|f\|'_{(\rho+\rho')/2,r}$ ,
- $\|f \circ U - f\|_{\rho',r} \leq \|f\|'_{(\rho+\rho')/2,r} \|u\|_{\rho',r}$ ,
- $\|D_{\mathbf{x}}f \circ U - D_{\mathbf{x}}f\| \leq \frac{4\pi}{\rho-\rho'} \|f\|'_{\rho,r} \|u\|_{\rho',r}$ .

The proof of these inequalities is straightforward and thus will be omitted. Now, assume that  $\delta = 42\varepsilon/\sigma < 1/2$ . For vector fields in the form  $X = \boldsymbol{\omega} + \pi_2 + f$ , where  $\pi_2: (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{y}$  is seen as a function in  $\mathcal{A}'_{\rho,r}$ , consider  $f$  to be in the open ball in  $\mathcal{A}'_{\rho,r}$  centred at the origin with radius  $\varepsilon$ . The coordinate transformation  $U$  is written as  $U = \text{Id} + (u, 0)$ , with  $u$  in

$$\mathcal{B} = \{u \in \mathbb{I}^- \mathcal{A}'_{\rho',r} : u: D_{\rho'} \times B_r \rightarrow D_{\rho}, \|u\|'_{\rho',r} < \delta\}.$$

Notice that we have

$$\begin{aligned} \mathbb{I}^- U^*(X) &= \mathbb{I}^- (DU)^{-1}(\boldsymbol{\omega} + \pi_2 + f \circ U, 0) \\ &= (\mathbb{I}^- (I + D_{\mathbf{x}}u)^{-1}(\boldsymbol{\omega} + \pi_2 + f \circ U), 0). \end{aligned}$$

From now on the parameter  $r$  is omitted whenever there is no ambiguity. Define the operator  $F: \mathcal{B} \rightarrow \mathbb{I}^- \mathcal{A}_{\rho'}$ ,

$$F(u) = \mathbb{I}^- (I + D_{\mathbf{x}}u)^{-1}(\boldsymbol{\omega} + \pi_2 + f \circ U). \quad (\text{A.1})$$

$F(u)$  takes real values for real arguments whenever  $u$  has that property. It is easy to see that the derivative of  $F$  at  $u$  is the linear map from  $\mathbb{I}^- \mathcal{A}'_{\rho'}$  to  $\mathbb{I}^- \mathcal{A}_{\rho'}$ :

$$\begin{aligned} DF(u)h &= \mathbb{I}^- (I + D_{\mathbf{x}}u)^{-1} [D_{\mathbf{x}}f \circ U h \\ &\quad - D_{\mathbf{x}}h (I + D_{\mathbf{x}}u)^{-1}(\boldsymbol{\omega} + \pi_2 + f \circ U)]. \end{aligned} \quad (\text{A.2})$$

We want to find a solution of

$$F(u_t) = (1 - t)F(u_0), \quad (\text{A.3})$$

with  $0 \leq t \leq 1$  and ‘‘initial’’ condition  $u_0 = 0$ . Differentiating the above equation with respect to  $t$ , we get

$$DF(u_t) \frac{du_t}{dt} = -F(0). \quad (\text{A.4})$$

**Proposition A.2.** *If  $u \in \mathcal{B}$ , then  $DF(u)^{-1}$  is a bounded linear operator from  $\mathbb{I}^- \mathcal{A}_{\rho'}$  to  $\mathbb{I}^- \mathcal{A}'_{\rho'}$  and  $\|DF(u)^{-1}\| < \delta/\varepsilon$ .*

From the above proposition (to be proved in Section A.0.1) we integrate (A.4) with respect to  $t$ , obtaining the integral equation:

$$u_t = - \int_0^t DF(u_s)^{-1} F(0) ds. \quad (\text{A.5})$$

In order to check that  $u_t \in \mathcal{B}$  for any  $0 \leq t \leq 1$ , we estimate its norm:

$$\begin{aligned} \|u_t\|'_{\rho'} &\leq t \sup_{v \in \mathcal{B}} \|DF(v)^{-1}F(0)\|'_{\rho'} \\ &\leq t \sup_{v \in \mathcal{B}} \|DF(v)^{-1}\| \|\mathbb{I}^- f\|_{\rho'} < t\delta \|f\|_{\rho'}/\varepsilon, \end{aligned}$$

so,  $\|u_t\|'_{\rho'} < \delta$ . Therefore, the solution of (A.3) exists in  $\mathcal{B}$  and is given by (A.5). Moreover, if  $X$  is real-analytic, then  $u_t$  takes real values for real arguments.

It is now easy to see that

$$U_t^* X - X^0 = \mathbb{I}^+ \sum_{n \geq 2} (-D(U_t - \text{Id}))^n X^0 + \mathbb{I}^+ U_t^* f + (1-t)\mathbb{I}^- f.$$

So, using Lemma A.1,

$$\begin{aligned} \|U_t^* X - X^0\|_{\rho'} &\leq \frac{1}{1 - \|u_t\|'_{\rho'}} (\|\boldsymbol{\omega}\| \|u_t\|'_{\rho'}{}^2 + \|f\|_{\rho}) + (1-t)\|f\|_{\rho'} \\ &< \frac{1}{1 - \delta} (\delta^2 \|\boldsymbol{\omega}\| \|f\|_{\rho'}/\varepsilon^2 + 1) \|f\|_{\rho} + (1-t)\|f\|_{\rho'} \\ &< \left[ \frac{1}{1 - \delta} \left( \frac{\delta^2 \|\boldsymbol{\omega}\|}{\varepsilon} + 1 \right) + 1 - t \right] \|f\|'_{\rho}. \end{aligned}$$

Moreover,  $\|U_t^* X - X^0 - \mathbb{I}^+ f - (1-t)\mathbb{I}^- f\|_{\rho'} = \mathcal{O}(\|f\|_{\rho}^2)$ , hence the derivative of  $X \mapsto U_t^* X$  at  $X^0$  is  $\mathbb{I} - t\mathbb{I}^-$ .

A.0.1. *Proof of Proposition A.2.*

**Lemma A.3.** *If  $\|f\|'_{\rho} < \varepsilon < \sigma/6$ , then  $DF(0)^{-1}: \mathbb{I}^- \mathcal{A}_{\rho'} \rightarrow \mathbb{I}^- \mathcal{A}'_{\rho'}$  is continuous and*

$$\|DF(0)^{-1}\| < \frac{3}{\sigma - 6\|f\|'_{\rho}}.$$

*Proof.* From (A.2) one has

$$\begin{aligned} DF(0) h &= \mathbb{I}^- (\widehat{f} - D_{\boldsymbol{\omega}}) h \\ &= - \left( \mathbb{I} - \mathbb{I}^- \widehat{f} D_{\boldsymbol{\omega}}^{-1} \right) D_{\boldsymbol{\omega}} h, \end{aligned}$$

where  $\widehat{f} h = Df h - Dh f$  and  $D_{\boldsymbol{\omega}} h = D_{\mathbf{x}} h(\boldsymbol{\omega} + \pi_2)$ . Thus, the inverse of this operator, if it exists, is given by

$$DF(0)^{-1} = -D_{\boldsymbol{\omega}}^{-1} \left( \mathbb{I} - \mathbb{I}^- \widehat{f} D_{\boldsymbol{\omega}}^{-1} \right)^{-1}.$$

The inverse of  $D_{\boldsymbol{\omega}}$  is the linear map from  $\mathbb{I}^- \mathcal{A}_{\rho'}$  to  $\mathbb{I}^- \mathcal{A}'_{\rho'}$ :

$$D_{\boldsymbol{\omega}}^{-1} g(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{k} \in I^-} \frac{g_{\mathbf{k}}(\mathbf{y})}{2\pi i \mathbf{k} \cdot X^0(\mathbf{y})} e^{2\pi i \mathbf{k} \cdot \mathbf{x}},$$

and is well-defined since Lemma 3.2 implies that  $|\mathbf{k} \cdot X^0(\mathbf{y})| > \sigma \|\mathbf{k}\|/2$ , with  $\mathbf{k} \in I^-$  and  $\mathbf{y} \in B_r$ . So,

$$\begin{aligned} \|D_{\boldsymbol{\omega}}^{-1} g\|'_{\rho'} &= \sum_{\mathbf{k} \in I^-} \frac{1 + 2\pi \|\mathbf{k}\|}{2\pi} \sup_{\mathbf{y} \in B_r} \left\| \frac{g_{\mathbf{k}}(\mathbf{y})}{\mathbf{k} \cdot X^0(\mathbf{y})} \right\| e^{\rho' \|\mathbf{k}\|} \\ &< \sum_{\mathbf{k} \in I^-} \frac{1 + 2\pi \|\mathbf{k}\|}{\pi \sigma \|\mathbf{k}\|} \|g_{\mathbf{k}}\|_r e^{\rho' \|\mathbf{k}\|} \leq \frac{3}{\sigma} \|g\|_{\rho'}. \end{aligned}$$

Hence,  $\|D_{\omega}^{-1}\| < 3/\sigma$ . It is possible to bound from above the norm of  $\widehat{f}: \mathbb{I}^- \mathcal{A}'_{\rho'} \rightarrow \mathcal{A}'_{\rho'}$  by  $\|\widehat{f}\| \leq 2\|f\|'_{\rho'}$ . Therefore,  $\|\mathbb{I}^- \widehat{f} D_{\omega}^{-1}\| < \frac{6}{\sigma} \|f\|'_{\rho'} < 1$ , and

$$\left\| \left( \mathbb{I} - \mathbb{I}^- \widehat{f} D_{\omega}^{-1} \right)^{-1} \right\| < \frac{\sigma}{\sigma - 6\|f\|'_{\rho'}}.$$

The statement of the lemma is now immediate.  $\square$

**Lemma A.4.** *Given  $u \in \mathcal{B}$ , the linear operator  $DF(u) - DF(0)$  mapping  $\mathbb{I}^- \mathcal{A}'_{\rho'}$  into  $\mathbb{I}^- \mathcal{A}'_{\rho'}$ , is bounded and*

$$\|DF(u) - DF(0)\| < \frac{\|u\|'_{\rho'}}{1 - \|u\|'_{\rho'}} \left[ \left( \frac{4\pi}{\rho - \rho'} + \frac{4 - 2\|u\|'_{\rho'}}{1 - \|u\|'_{\rho'}} \right) \|f\|'_{\rho'} + \frac{2 - \|u\|'_{\rho'}}{1 - \|u\|'_{\rho'}} \|\omega + \pi_2\| \right].$$

*Proof.* The formula (A.2) gives

$$\begin{aligned} [DF(u) - DF(0)] h &= \mathbb{I}^- (I + D_{\mathbf{x}}u)^{-1} [D_{\mathbf{x}}f \circ U h - (I + D_{\mathbf{x}}u) D_{\mathbf{x}}f h \\ &\quad - D_{\mathbf{x}}h (I + D_{\mathbf{x}}u)^{-1} (\omega + \pi_2 + f) \circ U \\ &\quad + (I + D_{\mathbf{x}}u) D_{\mathbf{x}}h (\omega + \pi_2 + f)] \\ &= \mathbb{I}^- (I + D_{\mathbf{x}}u)^{-1} \{A + B + C\}, \end{aligned}$$

where

$$\begin{aligned} A &= [D_{\mathbf{x}}f \circ U - D_{\mathbf{x}}f - D_{\mathbf{x}}u D_{\mathbf{x}}f] h \\ B &= D_{\mathbf{x}}u D_{\mathbf{x}}h (\omega + \pi_2 + f) \\ C &= -D_{\mathbf{x}}h (I + D_{\mathbf{x}}u)^{-1} [f \circ U - f - D_{\mathbf{x}}u (\omega + \pi_2 + f)]. \end{aligned}$$

Using Lemma A.1,

$$\begin{aligned} \|A\|_{\rho'} &\leq \left( \frac{4\pi}{\rho - \rho'} \|f\|'_{\rho'} \|u\|_{\rho'} + \|f\|'_{\rho'} \|u\|'_{\rho'} \right) \|h\|_{\rho'}, \\ \|B\|_{\rho'} &\leq (\|\omega + \pi_2\| + \|f\|_{\rho'}) \|u\|'_{\rho'} \|h\|'_{\rho'}, \\ \|C\|_{\rho'} &\leq \frac{1}{1 - \|u\|'_{\rho'}} [\|f\|'_{(\rho+\rho')/2} \|u\|_{\rho'} + \|u\|'_{\rho'} (\|\omega + \pi_2\|_r + \|f\|_{\rho'})] \|h\|'_{\rho'}. \end{aligned}$$

$\square$

To conclude the proof of Proposition A.2, notice that

$$\begin{aligned} \|DF(u)^{-1}\| &\leq (\|DF(0)^{-1}\|^{-1} - \|DF(u) - DF(0)\|)^{-1} \\ &< \left\{ \frac{\sigma}{3} - 2\varepsilon - \frac{\delta}{1 - \delta} \left[ \left( \frac{4\pi}{\rho - \rho'} + \frac{4 - 2\delta}{1 - \delta} \right) \varepsilon + \frac{2 - \delta}{1 - \delta} \|\omega + \pi_2\|_r \right] \right\}^{-1} \\ &< \frac{\delta}{\varepsilon}. \end{aligned}$$

The last inequality is true if

$$\varepsilon < \delta \left[ \frac{\sigma}{3} - \frac{2\delta}{(1 - \delta)^2} \|\omega + \pi_2\|_r \right] \left[ 1 + 2\delta + \frac{\delta^2}{1 - \delta} \left( \frac{4\pi}{\rho - \rho'} + \frac{4 - 2\delta}{1 - \delta} \right) \right]^{-1}$$

with a positive numerator  $N$  and denominator  $D$  in the r.h.s. This is so for our choices of  $\varepsilon$  and  $\delta < \frac{1}{2}$ , by observing that

$$\|\pi_2\|_r = \sup_{\mathbf{y} \in B_r} \|\mathbf{y}\| \leq a |\gamma^{(n)}|^{-1} \|M^{(n)}\| + b \|\omega\| < \frac{1}{2} \|\omega\|, \quad (\text{A.6})$$

thus  $\|\omega + \pi_2\|_r < \frac{3}{2}\|\omega\|$  and

$$\frac{2\delta}{(1-\delta)^2}\|\omega + \pi_2\|_r < 12\delta\|\omega\| < \frac{\sigma}{6}.$$

So,  $N > \delta\sigma/6$ ,  $D < 7$ , and finally  $\varepsilon \leq \frac{\sigma^2}{42\|\omega\|} < \frac{\sigma}{42} < N/D$ .

#### ACKNOWLEDGEMENTS

We would like to express our gratitude to D. Kleinbock, H. Koch, R. S. MacKay, G. Margulis, Ya. Sinai and J.-C. Yoccoz for useful discussions and comments.

JLD was supported by Fundação para a Ciência e a Tecnologia, and JM by an EPSRC Advanced Research Fellowship. We would also like to thank the Isaac Newton Institute, CAMGSD/IST and Cemapre/ISEG (through FCT's Program POCTI/FEDER) for travel support.

#### REFERENCES

- [1] J. J. Abad and H. Koch. Renormalization and periodic orbits for Hamiltonian flows. *Commun. Math. Phys.*, 212:371–394, 2000.
- [2] A. Ávila and R. Krikorian. Reducibility or non-uniform hyperbolicity for quasiperiodic Schrödinger cocycles. *Ann. of Math.*, to be published, 2005.
- [3] J. Brémont, K. Gawędzki and A. Kupiainen. KAM theorem and quantum field theory. *Comm. Math. Phys.*, 201:699–727, 1999.
- [4] J. W. S. Cassels. *An Introduction to Diophantine Approximation*. Cambridge University Press, 1957.
- [5] S. G. Dani. Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *J. Reine Angew. Math.*, 359:55–89, 1985.
- [6] M. J. Feigenbaum. Quantitative universality for a class of non-linear transformations. *J. Stat. Phys.*, 19:25–52, 1978.
- [7] D. G. Gaidashev. Renormalization of isoenergetically degenerate Hamiltonian flows and associated bifurcations of invariant tori. *Discrete Contin. Dyn. Syst.* 13:63–102, 2005.
- [8] G. Gallavotti. Twistless KAM tori. *Comm. Math. Phys.* 164:145–156, 1994.
- [9] G. Gentile and V. Mastropietro. Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications. *Rev. Math. Phys.* 8:393–444, 1996.
- [10] D. J. Grabiner and J. C. Lagarias. Cutting sequences for geodesic flow on the modular surface and continued fractions. *Monatsh. Math.* 133:295–339, 2001.
- [11] D. M. Hardcastle. The three-dimensional Gauss algorithm is strongly convergent almost everywhere. *Experiment. Math.* 11:131–141, 2002.
- [12] D. M. Hardcastle and K. Khanin, K. The  $d$ -dimensional Gauss transformation: strong convergence and Lyapunov exponents. *Experiment. Math.* 11:119–129, 2002.
- [13] M. R. Herman. Inégalités “a priori” pour des tores lagrangiens invariants par des difféomorphismes symplectiques. *Inst. Hautes Études Sci. Publ. Math.*, (70):47–101 (1990), 1989.
- [14] E. Hille and R. S. Phillips. *Functional analysis and semi-groups*, volume 31. AMS Colloquium Publications, rev. ed. of 1957, 1974.
- [15] K. Khanin, J. Lopes Dias and J. Marklof. Renormalization of multidimensional Hamiltonian flows. *preprint*, 2005.
- [16] K. Khanin and Ya. Sinai. The renormalization group method and Kolmogorov-Arnold-Moser theory. In R. Z. Sagdeev, editor, *Nonlinear phenomena in plasma physics and hydrodynamics*, pages 93–118. Mir Moscow, 1986.
- [17] D. Y. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math (2)*, 148:339–360, 1998.
- [18] H. Koch. A renormalization group for Hamiltonians, with applications to KAM tori. *Erg. Theor. Dyn. Syst.*, 19:475–521, 1999.
- [19] H. Koch. On the renormalization of Hamiltonian flows, and critical invariant tori. *Discrete Contin. Dyn. Syst.*, 8(3):633–646, 2002.
- [20] H. Koch. A renormalization group fixed point associated with the breakup of golden invariant tori. *Discrete Contin. Dyn. Syst.* 11:881–909, 2004.

- [21] J. C. Lagarias. Geodesic multidimensional continued fractions. *Proc. London Math. Soc.*, 69:464–488, 1994.
- [22] J. Lopes Dias. Renormalization of flows on the multidimensional torus close to a  $KT$  frequency vector. *Nonlinearity*, 15:647–664, 2002.
- [23] J. Lopes Dias. Renormalization scheme for vector fields on  $\mathbb{T}^2$  with a diophantine frequency. *Nonlinearity*, 15:665–679, 2002.
- [24] J. Lopes Dias. Brjuno condition and renormalization for Poincaré flows. *Discrete Contin. Dyn. Syst.*, to appear.
- [25] R. S. MacKay. *Renormalisation in area-preserving maps*. World Scientific Publishing Co. Inc., River Edge, NJ, 1993.
- [26] R. S. MacKay. Three topics in Hamiltonian dynamics. In Y. Aizawa, S. Saito, and K. Shiraiwa, editors, *Dynamical Systems and Chaos*, volume 2. World Scientific, 1995.
- [27] C. C. Moore. Ergodicity of flows on homogeneous spaces. *Am. J. Math.*, 88:154–178, 1966.
- [28] J. Moser. *Stable and random motions in dynamical systems*. Annals Math. Studies, Princeton Univ. Press, 1973.
- [29] M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, 1972.
- [30] F. Schweiger. *Multidimensional continued fractions*. Oxford Science Publications. Oxford University Press, Oxford, 2000.
- [31] J.-C. Yoccoz. Petits diviseurs en dimension 1 (Small divisors in dimension one). *Astérisque*, 231, 1995.
- [32] J.-C. Yoccoz. Analytic linearization of circle diffeomorphisms. *in Dynamical systems and small divisors*, Ed. Marmi and Yoccoz, Lecture Notes in Mathematics, 1784, Springer-Verlag, 2002.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, 100 ST GEORGE STREET, TORONTO, ONTARIO M5S 3G3, CANADA

*E-mail address:* `khanin@math.toronto.edu`

DEPARTAMENTO DE MATEMÁTICA, ISEG, UNIVERSIDADE TÉCNICA DE LISBOA, RUA DO QUELHAS 6, 1200-781 LISBOA, PORTUGAL

*E-mail address:* `jldias@iseg.utl.pt`

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL BS8 1TW, U.K.

*E-mail address:* `j.marklof@bristol.ac.uk`