QUANTITATIVE ALMOST REDUCIBILITY AND MÖBIUS DISJOINTNESS FOR ANALYTIC QUASIPERIODIC SCHRODINGER COCYCLES

WEN HUANG, JING WANG, ZHIREN WANG AND QI ZHOU

ABSTRACT. Sarnak's Möbius disjointness conjecture states that Möbius function is disjoint to any zero entropy dynamics. We prove that Möbius disjointness conjecture holds for one-frequency analytic quasi-periodic cocycles which are almost reducible, which extend [35, 43] to the noncommutative case. The proof relies on quantitative version of almost reducibility.

1. INTRODUCTION

A quasi-periodic $SL(2,\mathbb{R})$-cocycle is a linear skew product of the form

$$(x,\varpi)\mapsto(x+\alpha,A(x)\varpi)$$

where $x\in\mathbb{T}^1:=\mathbb{R}^1/\mathbb{Z}^1$, $\varpi\in\mathbb{R}^2$, $\alpha\in\mathbb{R}\backslash\mathbb{Q}$ and $A\in C^0(\mathbb{T}^1,SL(2,\mathbb{R}))$, denoted by (α,A). We consider the projective action of the $SL(2,\mathbb{R})$-cocycles (α,A) on $\mathbb{RP}^1$. Those are quasiperiodically forced (qpf) circle homeomorphisms of the form

$$T_{(\alpha,A)}\ :\ \mathbb{T}^1\times\mathbb{RP}^1\to\mathbb{T}^1\times\mathbb{RP}^1,\quad (x,\varphi)\mapsto(x+\alpha,\frac{A(x)\cdot\varphi}{\|A(x)\cdot\varphi)\|}),$$

where $\cdot$ denotes the Möbius transformation.

The topological dynamics of the projective action of a quasi-periodic $SL(2,\mathbb{R})$ cocycle is a very interesting subject in itself [13, 21, 22, 24, 28, 41, 42, 46], which has zero topological entropy [24]. In this paper, we focus on the question that whether this action satisfies Sarnak's Möbius disjointness conjecture.

The Möbius function $\mu:\mathbb{N}\to\{-1,0,1\}$ is defined by $\mu(1)=1$ and

$$\mu(n)=\begin{cases}(-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes;}\\ 0 & \text{otherwise.}\end{cases}$$

The orthogonality of μ with other sequences is a very important issue in number theory. For instance, the disjointness of Möbius function with constant sequence, that is $\sum_{n\le N}\mu(n)=o(N)$, is equivalent to the prime number theorem. The Möbius randomness law [27] suggests that $\mu(n)$ has significant cancellations with reasonable sequences $\xi(n)$, which means

$$\sum_{n\le N}\mu(n)\xi(n)=o(N). \tag{1.1}$$

Sarnak's Möbius disjointness conjecture [39] expects the disjointness holds for deterministic sequences, namely:

Date: June 10, 2021.

Conjecture 1.1 (Möbius disjointness conjecture)**.** *For a topological dynamical system (t.d.s.) (X,T) of zero topological entropy, all $f \in C(X)$ and all $x \in X$, the sequence $\xi(n) = f(T^n x)$ satisfies* (1.1).

There have been many partial results on Möbius disjointness conjecture. For the progress in this conjecture, we will simply refer to the recent comprehensive survey [19] for brevity, and we only discuss the historical developments that are more relevant to this paper. We remark that in the special case when $A(x) \in SO(2)$:

$$A(x) = R_{g(x)} := \begin{pmatrix} \cos(2\pi g(x)) & -\sin(2\pi g(x)) \\ \sin(2\pi g(x)) & \cos(2\pi g(x)) \end{pmatrix}$$

for some $g : \mathbb{T}^1 \to \mathbb{R}$, i..e., $T_{(\alpha,A)}$ has the form $(x,\varphi) \mapsto (x+\alpha, \varphi + g(x))$, then the action $T_{(\alpha,A)}$ is known to satisfy the Möbius disjointness conjecture as long as g is $C^{1+\epsilon}$ by the recent work of de Faveri [16] (see also the earlier works [25, 30, 32, 35, 43], which assumed higher regularity and/or other conditions).

Therefore the natural question is whether Möbius function is disjoint for any cocycle (α, A), or what happens if the non-commutative part appears? It is easy to see, if the cocycles has positive Lyapunov exponent, then Möbius disjointness conjecture holds (§3.3). Thus the real difficulty lies in if the cocycle has zero Lyapunov exponent. From this respect, we mention the following important concept raised by Eliasson [17, 18]. Recall that (α, A) is C^ω-conjugated to $(\alpha, \tilde{A})$, if there exists $B \in C^\omega(\mathbb{T}, PSL(2,\mathbb{R}))$ such that

$$B(\cdot+\alpha)^{-1} A(\cdot) B(\cdot) = \tilde{A}(\cdot).$$

Then (α, A) is said to be C^ω-almost reducible if its closure of analytic conjugate class contains a constant cocycle.

Theorem 1.2. *Let $(\alpha, A) \in \mathbb{R}\backslash\mathbb{Q} \times C^\omega(\mathbb{T}^1, SL(2,\mathbb{R}))$. If (α, A) is almost reducible, then the Möbius disjointness conjecture holds for $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$.*

One may ask, compared to the whole space of analytic cocycle, how large is the space of almost reducible cocycle is. Indeed, Avila's global theory of analytic $SL(2,\mathbb{R})$ cocycles [2, 3, 4] stated that typical cocycle is almost reducible or has positive Lyapunov exponent. To precise this, let's take quasi-periodic Schrödinger cocycles

$$A(x) = S_{v,E}(x) = \begin{pmatrix} E - v(x) & -1 \\ 1 & 0 \end{pmatrix},$$

as a typical example. Quasi-periodic Schrödinger cocycles are the main source of examples for quasi-periodic $SL(2,\mathbb{R})$ cocycles, because of their relation to one-dimensional discrete quasi-periodic Schrödinger operators

$$(H_{v,\alpha,x}u)_n = u_{n+1} + u_{n-1} + v(x+n\alpha)u_n, \forall n \in \mathbb{Z}, \tag{1.2}$$

where $x \in \mathbb{T}^1$ is the phase, $v \in C^0(\mathbb{T}^1, \mathbb{R})$ is the potential, and $\alpha \in \mathbb{T}^1$ is the frequency. Then as a direct corollary of Avila's global theory and Theorem 1.2, we have the following:

Corollary 1.3. *Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. Then for a (measure-theoretically) typical $v \in C^\omega(\mathbb{T}^1, \mathbb{R})$, the Möbius disjointness conjecture holds for the dynamical system $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha, S_{v,E})})$ induced by the projective action of $(\alpha, S_{v,E})$ for any $E \in \mathbb{R}$.*

Let us explain the meaning of "typical" in the theorem as in [4]. Since in the infinite-dimensional settings one lacks a translation-invariant measure (Haar measure), it is common to replace the notion of *almost every* by *prevalence*: one fixes some probability measure μ of compact support (a set of admissible perturbations w), and declares a property to be typical if it is satisfied for almost every perturbation $v+w$ of every starting condition w. In finite-dimensional vector spaces, prevalence implies full Lebesgue measure. For instance, one can set $\Delta = \mathbb{D}^{\mathbb{N}}$ endowed with the probability measure μ given by the product of normalized Lebesgue measure. Given an arbitrary function $\epsilon : \mathbb{N} \to \mathbb{R}_+$ which decays exponentially fast, we associate a probability measure μ_ϵ with compact support on $C^\omega(\mathbb{T}^1, \mathbb{R})$ by push forward of μ under the map

$$\{t_m\}_{m\in\mathbb{Z}} \to \sum_{m\geq 1} \epsilon(m) 2\Re[t_m e^{2\pi i m x}].$$

Thus Corollary 1.3 says that, for any $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, and for every $v \in C^\omega(\mathbb{T}^1, \mathbb{R})$, the Möbius disjointness conjecture holds for the dynamics $T_{(\alpha, S_{v+w,E})}$ for μ_ϵ-almost every w and every $E \in \mathbb{R}$.

Remark 1.4. *In essence, Corollary 1.3 relies on Avila's Almost Reducibility Conjecture(ARC), which asserts that any subcritical cocycle*[1] *is almost reducible. A proof of ARC is announced in* [4]*, to appear in* [2, 3]*. ARC has many important dynamical and spectral consequences* [2, 4, 8, 9, 11, 20, 34, 36]*, indeed, it was already stated as Almost Reducibility Theorem(ART) in* [8]*.*

If we take $v(x) = 2\lambda\cos(2\pi x)$, which corresponds to almost Mathieu operators, then we have the following:

Corollary 1.5. *Let* $\alpha \in \mathbb{R}\backslash\mathbb{Q}$*, the Möbius disjointness conjecture holds for the dynamics* $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha, S_{2\lambda\cos(2\pi\cdot),E})})$ *for any* $\lambda \neq \pm 1$*, and for any* $E \in \mathbb{R}$*.*

We take Schrödinger cocycles (and almost Mathieu operators) as typical examples, since it is interesting from both physical and mathematical points of view. Physically, quasi-periodic Schrödinger operator is a mathematical model for quasicrystals, and it plays a central role in explaining the quantum Hall effect (which was Thouless's Nobel Prize work) [33, 40]. From the mathematical side, Schrödinger cocycles (even almost Mathieu cocycles) could display all the dynamical behavior of general $SL(2,\mathbb{R})$ cocycles [4]. For example, Avila-Fayad-Krikorian [5] proved that for any $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, $v \in C^\omega(\mathbb{T}^1, \mathbb{R})$ and for almost every energy $E \in \mathbb{R}$ the analytic cocycle $(\alpha, S_{v,E})$ is either C^ω-rotations reducible or has positive Lyapunov exponent. Recall (α, A) is called C^ω*-rotations reducible*, if there exists $B \in C^\omega(\mathbb{T}, PSL(2,\mathbb{R}))$ such that

$$B(\cdot+\alpha)^{-1}A(\cdot)B(\cdot) \in SO(2,\mathbb{R}).$$

Thus Theorem 1.3 can be viewed as a direct non-commutative extension of the results from [35, 43]. Secondly, to better understand zero topological entropy, complexity is a very useful concept, while complexity is closely related to the growth of cocycles, and Schrödinger cocycles (even almost Mathieu cocycles) are believed to display high order complexity [10].

Finally, we outline the novelty and difficulty of the proof. The concept of almost reducibility generalizes the scope of applicability of the local theories. While quantitative

[1]Consult section 2.4 for precise definition.

estimates were not involved in the definition, we stress that the main novelty of the paper is that quantitative version of almost reducibility implies the disjointness. Theorem 1.2 covers all frequencies, and the difficulty in its proof lies in two possible issues: the frequency being Liouvillean and the cocycle being almost reducible to parabolic cocycles.

If the frequency is Diophantine, then in the commutative case [35, 43], the Möbius disjointness is a straightforward corollary of Davenport's theorem [15]; while in our case, quantitative almost reducibility was obtained based on almost localization of the dual model, as a consequence, we prove that the cocycle has sub-polynomial complexity, which leads to Möbius disjointness by [25]. When the frequency is Liouvillean, almost reducibility was previous obtained by different methods: periodic approximation as in [2], and non-standard KAM as in [23]. However, these estimates are far from our needs. Indeed, uniform estimates in the parameters are usually not useful, since counterexamples [10] and even phase transitions [11] could happen. On the other hand, to get precise estimates for any frequency, one has to interpolate non-standard KAM schemes [23] with classical KAM schemes [17].

Another difficulty is that (α, A) may be almost reducible to parabolic cocycles, i.e. there exists conjugacies B_n which conjugate (α, A) to perturbations of (α, A_n), where $A_n = \begin{pmatrix} 1 & c_n \\ 0 & 1 \end{pmatrix}$ is parabolic. This is a new difficulty in the non-commutative case. In this case, one needs delicate control on the sizes of c_n and the conjugacy B_n (Proposition 3.3). We point out the same difficulty also occurs in Avila-You-Zhou's work [12] on the dry Martini problem. However, Aubry duality, a key tool in [12], cannot be used here. To deal with this obstruction, we distinguish two cases: if c_n is relatively small, then one can consider it as a perturbation of identity, and still apply the measurable complexity criterion of Huang-Wang-Ye [25]; if c_n is relatively large, we will approximate the trajectories of the projective cocycles with either rotational or periodic trajectories on sufficiently long intervals, and use the Matomäki-Radziwiłł-Tao estimate; however, if c_n is too large (even of constant size), then one will lose control of the conjugacy B_n, which destroy the whole KAM scheme. Our quantitative argument will provide bounds on c_n that prevents this from happening.

The remainder of the paper is organized as follows. In section 2, we give some notations and collect some necessary background which will be used in what follows. In section 3, we give the proof of the main theorems, mainly Theorem 1.2. If the cocycle is almost reducible and not uniformly hyperbolic, we get the quantitative almost reducibility result for $\beta(\alpha) = 0$ and $\beta(\alpha) > 0$ respectively, and in the end deduce Möbius disjointness according to the type of the constant matrices in the conjugated systems (elliptic or parabolic). In section 4, we give the quantitative almost reducibility for $\beta(\alpha) = 0$, using Aubry duality and quantitative almost localization. In section 5, we give the quantitative almost reducibility for $\beta(\alpha) > 0$ by KAM scheme. In section 6, we get the measure complexity of the cocycle for the case that the constant matrices in the conjugated cocycles are elliptic. In section 7, we prove the linear disjointness of the dynamics of the system with Möbius function for the case that the constant matrices in the conjugated cocycles are parabolic.

2. Notations and preliminaries

2.1. Notations on functions.

Denote by $C^\omega_h(\mathbb{T}^d,*)$ the set of all $*$-valued functions ($*$ will usually denote $\mathbb{R}$, $sl(2,\mathbb{R})$ $SL(2,\mathbb{R})$ and etc.) which are analytic and bounded in $\{x\in\mathbb{C}^d/\mathbb{Z}^d \ : \ |\Im x|<h\}$, and for any $F\in C^\omega_h(\mathbb{T}^d,*)$, we define the norm

$$\|F\|_h := \sup_{|\Im x|<h}\|F(x)\|,$$

(where $\|\cdot\|$ denotes absolute value or the usual matrix norm).

Moreover, an integrable real-valued function f on $\mathbb{T}^d$ has the Fourier expansion $f(\phi)=\sum_{k\in\mathbb{Z}^d}\hat{f}(k)e^{2\pi i\langle k,\phi\rangle}$ with $\hat{f}(k)=\int_{\mathbb{T}^d}f(\phi)e^{-2\pi i\langle k,\phi\rangle}d\phi$. For any $K>0$, $\mathcal{T}_K$ and $\mathcal{R}_K$ are used to denote the truncation operators:

$$\mathcal{T}_N(f)=\sum_{|k|<K}\hat{f}(k)e^{2\pi i\langle k,\phi\rangle},\ \mathcal{R}_N(f)=\sum_{|k|\geq K}\hat{f}(k)e^{2\pi i\langle k,\phi\rangle}.$$

2.2. Continued fraction expansion and CD bridge.

Let $\alpha\in\mathbb{R}^1$ be irrational. Define $a_0=[\alpha],\alpha_0=\alpha-a_0$, and inductively for $k\geq 1$,

$$a_k=[\alpha_{k-1}^{-1}],\qquad \alpha_k=\{\alpha_{k-1}^{-1}\}=\alpha_{k-1}^{-1}-a_k.$$

We define $p_0=0,p_1=1,q_0=1,q_1=a_1$ and inductively,

$$p_k=a_kp_{k-1}+p_{k-2},\quad q_k=a_kq_{k-1}+q_{k-2}.$$

Then the sequence $\{q_n\}_{n\in\mathbb{N}}$ satisfies

$$\forall 1\leq k<q_n,\quad \|k\alpha\|_{\mathbb{T}}\geq\|q_{n-1}\alpha\|_{\mathbb{T}},$$
$$\frac{1}{q_n+q_{n+1}}\leq\inf_{p\in\mathbb{Z}}|q_n\alpha-p|\leq\frac{1}{q_{n+1}}.$$

Let $\beta(\alpha):=\limsup_{n\to\infty}\frac{\ln q_{n+1}}{q_n}$. Equivalently, we have

$$\beta(\alpha)=\limsup_{k\to\infty}\frac{1}{|k|}\ln\frac{1}{\|k\alpha\|_{\mathbb{T}}},$$

where $\|k\alpha\|_{\mathbb{T}}:=\inf_{j\in\mathbb{Z}}|k\alpha-j|$.

Let us now introduce the CD bridge.

Definition 2.1 ([5]). *Let $0<\mathcal{A}\leq\mathcal{B}\leq\mathcal{C}$. We say that the pair of denominators (q_l,q_n) forms a* $\mathrm{CD}(\mathcal{A},\mathcal{B},\mathcal{C})$ *bridge if*

- $q_{i+1}\leq q_i^{\mathcal{A}},\quad \forall i=l,\dots,n-1$
- $q_l^{\mathcal{C}}\geq q_n\geq q_l^{\mathcal{B}}$.

In the following, for simplicity, we will fix a subsequence $\{q_{n_k}\}_k$ of $\{q_n\}_n$, denoted by $\{Q_k\}_k$, and the subsequence $\{q_{n_k+1}\}_k$, denoted by $\{\overline{Q}_k\}_k$.

The following lemma was proved by Avila-Fayad-Krikorian in [5], and we include the proof for completeness.

Lemma 2.2 ([5]). *For any $\mathcal{A}>0$, there exists a subsequence $\{Q_k\}_{k\in\mathbb{N}_0}$ such that $Q_0=1$ and for each $k\geq 1$, $Q_k\leq\overline{Q}_{k-1}^{\mathcal{A}^4}$. Furthermore, either $\overline{Q}_k\geq Q_k^{\mathcal{A}}$, or the pairs $(\overline{Q}_{k-1},Q_k)$ and (Q_k,Q_{k+1}) are both* $\mathrm{CD}(\mathcal{A},\mathcal{A},\mathcal{A}^3)$ *bridges.*

Proof. Suppose for $l \leq k$, we have already selected Q_l satisfying the conditions. If $q_{n+1} \leq q_n^{\mathcal{A}}$ for any $q_n > Q_k$, then we can find $q_{n_0} = \overline{Q}_k, q_{n_1}, q_{n_2}, \dots$ that $(q_{n_j}, q_{n_{j+1}})$ is CD$(\mathcal{A}, \mathcal{A}, \mathcal{A}^3)$ bridge for $j = 0, 1, \dots$. In this case, we let $Q_{k+j} = q_{n_j}$ for $j \geq 1$. Otherwise, we let $q_n > Q_k$ be the smallest denominator that satisfies $q_{n+1} > q_n^{\mathcal{A}}$. In this situation, if $q_n \leq \overline{Q}_k^{\mathcal{A}^4}$, then let $Q_{k+1} = q_n$. Otherwise, we can find $q_{n_0} = \overline{Q}_k < q_{n_1} < \dots$ that $(q_{n_j}, q_{n_{j+1}})$ is CD$(\mathcal{A}, \mathcal{A}, \mathcal{A}^2)$ bridge. Let $q_{n_l} < q_n$ be the smallest denominator that $q_{n_l}^{\mathcal{A}^2} \geq q_n$, by which we know that $l \geq 1$. If $q_n \geq q_{n_l}^{\mathcal{A}}$, let $Q_{k+j} = q_{n_j}$ for $j = 1, \dots, l$ and $Q_{k+l+1} = q_n$. Then $(\overline{Q}_k, Q_{k+1})$, (Q_{k+j}, Q_{k+j+1}) for $j = 1, \dots, l$ are CD$(\mathcal{A}, \mathcal{A}, \mathcal{A}^2)$ bridges and thus CD$(\mathcal{A}, \mathcal{A}, \mathcal{A}^3)$ bridges. If $q_n < q_{n_l}^{\mathcal{A}}$ (in this case, $l \geq 2$), we let $Q_{k+j} = q_{n_j}$ for $j = 1, \dots, l-1$ and $Q_{k+l} = q_n$. Then $(\overline{Q}_k, Q_{k+1})$, (Q_{k+j}, Q_{k+j+1}) for $j = 1, \dots, l-2$ are CD$(\mathcal{A}, \mathcal{A}, \mathcal{A}^2)$ bridges and (Q_{k+l-1}, Q_{k+l}) is CD$(\mathcal{A}, \mathcal{A}, \mathcal{A}^3)$ bridge. This completes the proof. □

Remark 2.3. *By the selection of* $\{Q_k\}_{k \in \mathbb{N}_0}$*, those* q_n *satisfying* $q_{n+1} > q_n^{\mathcal{A}}$ *belong to* $\{Q_k\}_{k \in \mathbb{N}_0}$.

Corollary 2.4 ([31])**.** *For the subsequence* $\{Q_k\}_{k \in \mathbb{N}_0}$ *selected in Lemma 2.2 we have* $Q_k \geq Q_{k-1}^{\mathcal{A}}$ *for every* $k \geq 1$.

2.3. **Lyapunov exponent and uniformly hyperbolic.** Given $A \in C^\omega(\mathbb{T}, SL(2, \mathbb{R}))$ and $\alpha \in \mathbb{R}$ irrational, the iterates of the quasi-periodic cocycle (α, A) are of the form $(\alpha, A)^n = (n\alpha, A_n)$, where

$$A_n(x) := \begin{cases} A(x + (n-1)\alpha)A(x + (n-2)\alpha) \cdots A(x), & n \geq 0 \\ A(x + n\alpha)^{-1}A(x + (n+1)\alpha)^{-1} \cdots A(x - \alpha)^{-1}, & n < 0 \end{cases}.$$

The *Lyapunov exponent* is defined as $L(\alpha, A) := \lim_{n \to \infty} \frac{1}{n} \int_{\mathbb{T}^1} \ln \|A_n(x)\| dx \geq 0$.

We say the cocycle (α, A) is *uniformly hyperbolic* if there exists a continuous splitting $\mathbb{R}^2 = E^s(x) \oplus E^u(x)$, and $C > 0, 0 < \lambda < 1$ such that for every $n \geq 1$ we have

$$\|A_n(x) \cdot w\| \leq C\lambda^n \|w\|, \quad w \in E^s(x),$$
$$\|A_{-n}(x) \cdot w\| \leq C\lambda^n \|w\|, \quad w \in E^u(x).$$

Such a splitting is automatically unique and thus invariant, i.e., $A(x)E^s(x) = E^s(x + \alpha)$ and $A(x)E^u(x) = E^u(x + \alpha)$.

2.4. **Global theory of one-frequency $SL(2, \mathbb{R})$-cocycles.** Avila's global theory classified the analytic $SL(2, \mathbb{R})$ cocycles that are not uniformly hyperbolic to three cases: *supercritical*, *subcritical* and *critical*. A cocycle (α, A) which is not uniformly hyperbolic is said to be *supercritical* if $L(\alpha, A) > 0$; it is *subcritical*, if there exists $\delta > 0$ such that $L(\alpha, A(z)) = 0$ for $|\Im z| \leq \delta$; and it is *critical* otherwise. In particular, we say the Schrödinger operator $H_{v,\alpha}$ is *acritical* if for any E in the spectrum of the cocycle $(\alpha, S_{v,E})$ is not critical. To give an example of subcritical cocycles, we consider a cocycle (α, R_g) which is homotopic to constant. It is subcritical and the projective action is of the form

$$T : \mathbb{T}^1 \times \mathbb{RP}^1 \circlearrowleft \text{ via } (x, \varphi) \mapsto (x + \alpha, \varphi + g(x)).$$

Moreover, if (α, A) is supercritical, then it is *non-uniformly hyperbolic*, that is, there is a splitting as in the uniformly hyperbolic case, except that this splitting is not continuous. It is equivalent to the existence of a *strange non-chaotic attractor* ϕ^- (An SNA in a qpf system T is a T-invariant graph which has negative Lyapunov exponent and is not continuous.) and a *strange non-chaotic repeller* ϕ^+ (An SNR is a non-continuous T-invariant graph with positive Lyapunov exponent.) for the qpf circle homeomorphism $T_{(\alpha,A)}$ induced by the projective action of (α, A) (For a detailed discussion of this relation, see [28, Section 1.3.2]).

The following theorem of Avila shows that the critical case is very rare.

Theorem 2.5 ([4]). *Let α be irrational. Then for a (measure-theoretically) typical $v \in C^\omega(\mathbb{T}^1, \mathbb{R})$, the operator $H_{v,\alpha}$ is acritical.*

2.5. **Measure complexity of a t.d.s.** Let (X, T) be a topological dynamical system(t.d.s.) with a metric d and let $\mathcal{M}(X, T)$ be the set of all T-invariant Borel probability measures on X. For $\rho \in \mathcal{M}(X, T)$ and any $n \in \mathbb{N}$, we consider the metric

$$\bar{d}_n(x, y) = \frac{1}{n} \sum_{i=0}^{n-1} d(T^i x, T^i y)$$

for any $x, y \in X$. For $\epsilon > 0$, let

$$S_n(d, \rho, \epsilon) = \min\{m \in \mathbb{N} \ : \ \exists x_1, x_2, \cdots, x_m \ s.t. \ \rho(\bigcup_{i=1}^{m} B_{\bar{d}_n}(x_i, \epsilon)) > 1 - \epsilon\},$$

where $B_{\bar{d}_n}(x, \epsilon) := \{y \in X \ : \ \bar{d}_n(x, y) < \epsilon\}$ for any $x \in X$.

Definition 2.6 ([25]). *Let $U : \mathbb{N} \to [1, +\infty)$ be an increasing sequence with $\lim_{n\to+\infty} U(n) = +\infty$. We say the measure complexity of (X, d, T, ρ) is weaker than $U(n)$ if*

$$\liminf_{n\to+\infty} \frac{S_n(d, \rho, \epsilon)}{U(n)} = 0, \quad \forall \epsilon > 0.$$

Remark 2.7. *By* [25, Prop. 2.2]*, the measure complexity of (X, d, T, ρ) is weaker than $U(n)$ if and only if the measure complexity of (X, d', T, ρ) is weaker than $U(n)$ for any compatible metric d' on X. Thus we can simply say the measure complexity of (X, T, ρ) is weaker than $U(n)$.*

Definition 2.8 ([25]). *Let (X, T) be a t.d.s. and $\rho \in \mathcal{M}(X, T)$. The measure complexity of (X, T, ρ) is sub-polynomial if it is weaker than $U_\tau(n) = n^\tau$ for any $\tau > 0$.*

Using the measure complexity, Huang-Wang-Ye[25] provided a criterion for a t.d.s. satisfying the required disjointness, which says:

Theorem 2.9 ([25]). *Let (X, T) be a t.d.s. such that the measure complexity of (X, T, ρ) is sub-polynomial for any $\rho \in \mathcal{M}(X, T)$. Then the Möbius disjointness conjecture holds.*

Using Theorem 2.9, [25] got Möbius disjointness for systems with discrete spectrum. Recall that a t.d.s. (X, T, ρ) has *discrete spectrum* if $L^2(X, \mathcal{B}_X, \rho)$ is spanned by the set of eigenfunctions for T, where $\mathcal{B}_X$ stands for the Borel σ-algebra of X and $\rho \in \mathcal{M}(X, T)$. Let (X, T) and (Y, S) be two t.d.s., and let $\rho \in \mathcal{M}(X, T)$ and $\nu \in \mathcal{M}(Y, S)$. If $(X, \mathcal{B}_X, T, \rho)$ is measurably isomorphic to $(Y, \mathcal{B}_Y, S, \mu)$, then (X, T, ρ) has discrete spectrum iff (Y, S, ν) has discrete spectrum. Moreover, the following theorem implies that a t.d.s which has discrete spectrum for any invariant measure is linear disjoint with the Möbius function.

Theorem 2.10 ([25]). *Let (X, T) be a t.d.s. and $\rho \in \mathcal{M}(X, T)$. If ρ has discrete spectrum, then the measure complexity of (X, T, ρ) is sub-polynomial.*

The following fact is evident and we omit the proof.

Proposition 2.11. *Let $(X,T),(Y,S)$ be two t.d.s.. If they are conjugate, i.e. there exists a homeomorphism $h: X \to Y$ such that $h \circ T = S \circ h$, then the Möbius disjointness conjecture holds for (X,T) if and only if it is true for (Y,S).*

3. Proof of the main theorems

In this section, we give the proof of Theorem 1.2 and Corollary 1.3, Corollary 1.5. In order to prove Theorem 1.2, we first consider the case that (α, A) is not uniformly hyperbolic, since the case that (α, A) is uniformly hyperbolic is relatively easy (Lemma 3.7).

Theorem 3.1. *Let $(\alpha, A) \in \mathbb{R}\backslash\mathbb{Q} \times C^\omega(\mathbb{T}^1, SL(2,\mathbb{R}))$. If (α, A) is almost reducible and not uniformly hyperbolic, then Sarnak's Möbius conjecture holds for $T_{(\alpha,A)}$.*

Strategy of the proof. As we mentioned before, quantitative estimates are not involved in the definition of almost reducibility, and the key point in the proof is to get the quantitative almost reducibility of the cocycle with delicate control on the conjugations and perturbations. To get the appropriate quantitative almost reducibility, if $\beta(\alpha) = 0$, the proof will be based on a quantitative version of Aubry duality and almost localization, and if $\beta(\alpha) > 0$, we will perform a KAM scheme. In the end, we deduce the desired result from the quantitative almost reducibility.

3.1. Quantitative almost reducibility.

First we state the quantitative almost reducibility results that we need, and leave the full proof to section 4 and section 5. Both the quantitative almost reducibility result and the proof will depend on the arithmetic property of the frequency α. If $\beta(\alpha) = 0$, then we have the following:

Proposition 3.2. *Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ with $\beta(\alpha) = 0$. If (α, A) is almost reducible and not uniformly hyperbolic, then either it is reducible, or there exist $r_* > 0, n_* \in \mathbb{N}, \varrho \in \mathbb{R}$, and an infinite sequence $\{n_j\}_{j\in\mathbb{N}} \subseteq \mathbb{N}$ such that for $n_j \geq n_*$, there exists $W_j : \mathbb{T}^1 \to SL(2,\mathbb{R})$ analytic with $\|W_j\|_{r_*} \leq Ce^{o(n_j)}$ that*

$$W_j(\cdot + \alpha)^{-1} A(\cdot) W_j(\cdot) = R_{\pm\varrho}(I + G_j(\cdot)),$$

with $\|G_j\|_{r_*} \leq Ce^{-cn_j}$, where c, C are constants not depending on j.

If $\beta(\alpha) > 0$, let $\mathcal{A} > 2$ and $\{Q_k\}_{k\in\mathbb{N}_0}$ be the subsequence selected in Lemma 2.2. Then we will have the following:

Proposition 3.3. *Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ with $\beta(\alpha) > 0$. If (α, A) is almost reducible and not uniformly hyperbolic, then*

(1) *either there exist sequences $\{W_j\}_{j\in\mathbb{N}} \subseteq C^\omega(\mathbb{T}^1, PSL(2,\mathbb{R}))$, $\{\varrho_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}$ and $\{n_j\}_{j\in\mathbb{N}} \subseteq \mathbb{N}$ such that*

$$W_j(\cdot + \alpha)^{-1} A(\cdot) W_j(\cdot) = R_{\varrho_j}(I + G_j(\cdot))$$

with $\|W_j\|_{C^1} \leq e^{n_j \eta_j}$, $\|G_j\|_{C^0} \leq e^{-n_j \tau_j}$, where $\eta_j = o(\tau_j) < 1$, and $\tau_j n_j^{1/2} > 1$;

(2) *or there is $k_* \in \mathbb{N}$ that for $k \geq k_*$, there exists $W_k \in C^\omega(\mathbb{T}^1, PSL(2,\mathbb{R}))$ that*

$$W_k(\cdot + \alpha)^{-1} A(\cdot) W_k(\cdot) = \begin{pmatrix} 1 & c_k \\ 0 & 1 \end{pmatrix} (I + G_k(\cdot))$$

with estimates $\|G_k\|_{C^0} \le Ce^{-Q_k\tau_k}$, $\|W_k\|_{C^1} \le Ce^{CQ_k\eta_k}$, *where* $e^{-\frac{Q_k\tau_k}{10}} \le |c_k| \le e^{-Q_k\eta_k}$, $\eta_k = o(\tau_k), \tau_k = o(1)$, $\tau_k Q_k^{\frac{1}{2A}} > 1$ *and* C *is a global constant not depending on* k*.*

3.2. **Proof of Theorem 3.1.** Let (α, A) be almost reducible and not uniformly hyperbolic. By Proposition 3.2 and Proposition 3.3, the cocycle (α, A) is either reducible, or almost reducible with corresponding quantitative estimations.

For reducible cocycles, we have the following measure complexity estimate:

Lemma 3.4. *If* (α, A) *is reducible and not uniformly hyperbolic, then the Möbius disjointness holds for the t.d.s.* $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$ *induced by the projective action of* (α, A)*.*

Proof. Assume that (α, A) is conjugate to (α, D) with $D \in SL(2,\mathbb{R})$. By Proposition 2.11, it suffices to show the Möbius disjointness conjecture holds for (α, D). As (α, A) is not uniformly hyperbolic, then D is elliptic or parabolic. Without loss of generality, we assume D is in the real normal form R_ϱ or $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$, where $\varrho \in \mathbb{T}^1$, and $0 \neq c \in \mathbb{R}$.

Case (1): $D = R_\varrho$. In this case,

$$T^n_{(\alpha,D)}(x, \varphi) = (x + n\alpha, \varphi + n\varrho).$$

As $T_{(\alpha,D)}$ is a translation on $\mathbb{T}^1 \times \mathbb{RP}^1$, it is a straightforward corollary of Davenport's theorem [15] that it satisfies the Möbius disjointness conjecture (see e.g [39]).

Case (2): $D = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$. In this case, we have

$$T^n_{(\alpha,D)}(x, \varphi) = (x + n\alpha, \Pi \circ \left(\frac{1}{2\pi} \arctan \frac{1}{\cot \hat{\varphi} + nc}\right)),$$

where $\Pi : (-\frac{1}{4}, \frac{1}{4}] \to \mathbb{RP}^1$ is the canonical projection, and $\hat{\varphi} := 2\pi\gamma(\varphi)$ with $\gamma : \mathbb{RP}^1 \to (-\frac{1}{4}, \frac{1}{4}]$ being the lift of the identity map on $\mathbb{RP}^1$. Then $\Pi \circ \left(\frac{1}{2\pi} \arctan \frac{1}{\cot \hat{\varphi}+nc}\right) \to 0$ as $n \to \infty$. Given $f \in C(\mathbb{T}^1 \times \mathbb{RP}^1)$ and $(x, \varphi) \in \mathbb{T}^1 \times \mathbb{RP}^1$, as f is uniformly continuous, $f(T^n_{(\alpha,D)}(x, \varphi)) - f(x + n\alpha, 0) \to 0$. It follows that $\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^N \mu(n) f(T^n_{(\alpha,D)}(x, \varphi))$ coincides with $\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^N \mu(n) f(x + n\alpha, 0)$. Again, one can easily deduce from Davenports theorem as in Case (1) that this later limit vanishes. Hence the Möbius disjointness holds in this case. □

Lemma 3.4 deals with the case that (α, A) is reducible. If (α, A) is almost reducible with suitable estimates, that is

$$W_j(\cdot + \alpha)^{-1} A(\cdot) W_j(\cdot) = A_j(I + G_j(\cdot)),$$

then the proof of Möbius disjointness conjecture not only depends on the quantitative estimates of the conjugation W_j and the perturbation G_j, but most importantly, also depends on the structure of A_j. If A_j is elliptic, then one should estimate the measure complexity of the dynamics as follows:

Proposition 3.5. *Let* $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ *and* $A \in C^\omega(\mathbb{T}^1, SL(2,\mathbb{R}))$*. If there exists* $j_* \in \mathbb{N}$ *and a sequence* $\{\varrho_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}$ *such that for* $j \ge j_*$*, there is* $W_j \in C^\omega(\mathbb{T}^1, PSL(2,\mathbb{R}))$ *that*

$$W_j(\cdot + \alpha)^{-1} A(\cdot) W_j(\cdot) = R_{\varrho_j}(I + G_j(\cdot)),$$

with $\|G_j\|_{C^0} \to 0$ *and* $\|W_j\|_{C^1}\|G_j\|_{C^0}^{\eta} \to 0$ *as* $j \to \infty$ *for any* $\eta > 0$*, then the measure complexity of* $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)}, \rho)$ *is sub-polynomial for any* $\rho \in \mathcal{M}(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$.

If A_j is parabolic and $\beta(\alpha) > 0$, then one should use periodic approximation, and decompose the periodic sequence into short average of Dirichlet characters, reducing the problem to control the average of multiplicative function on a typical interval to finish the proof.

For $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ with $\beta := \beta(\alpha) > 0$, we assume $\{\frac{p_k}{q_k}\}_{k\in\mathbb{N}_0}$ is the convergents of α, and denote

$$\mathfrak{Q} := \{q_k \mid q_{k+1} \geq e^{\frac{\beta}{2}q_k}\}.$$

Proposition 3.6. *For the analytic cocycle* (α, A) *with* $\beta := \beta(\alpha) > 0$*, if there exists a sequence* $\{q_{k_j}\}_j \subseteq \mathfrak{Q}$ *and* $\{W_j\}_j \subseteq C^{\omega}(\mathbb{T}^1, PSL(2,\mathbb{R}))$ *that*

$$W_j(\cdot+\alpha)^{-1}A(\cdot)W_j(\cdot) = A_j(I + G_j(\cdot))$$

with $A_j = \begin{pmatrix} 1 & c_j \\ 0 & 1 \end{pmatrix}$, $\|G_j\|_{C^0} \leq \tilde{C}e^{-q_{k_j}\tau_j}$, $\|W_j\|_{C^1} \leq \tilde{C}e^{\tilde{C}q_{k_j}\eta_j}$, $\tilde{c}e^{-(\frac{1}{7}-\xi)q_{k_j}\tau_j} \leq |c_j| \leq 1$ *for some* $0 < \xi < \frac{1}{7}$, $\eta_j = o(\tau_j)$, $\tau_j = o(1)$, *and* $\tau_j q_{k_j} > q_{k_j}^{\frac{1}{2}+\epsilon}$ *for some* $\epsilon > 0$*, where* $\tilde{c}, \tilde{C} > 0$ *are absolute constants, then the Möbius disjointness holds for* $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$.

We leave the proof of Proposition 3.5 and Proposition 3.6 in section 6 and section 7 respectively.

Proof of Theorem 3.1. For $\beta(\alpha) = 0$, if (α, A) is reducible, then by Lemma 3.4, the Möbius disjointness conjecture holds for $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$; otherwise, by Proposition 3.2, there exist $r_* > 0, n_* \in \mathbb{N}, \varrho \in \mathbb{R}$, and a sequence $\{n_j\}_{j\in\mathbb{N}}$ such that for $n_j \geq n_*$, there exists $W_j \in C^{\omega}(\mathbb{T}^1, SL(2,\mathbb{R}))$ such that $W_j(\cdot+\alpha)^{-1}A(\cdot)W_j(\cdot) = R_{\pm\varrho}(I + G_j(\cdot))$ with $\|W_j\|_{C^1} \leq \frac{C}{r_*}e^{o(n_j)}$ and $\|G_j\|_{C^0} \leq Ce^{-cn_j}$, which means that $\|W_j\|_{C^1}\|G_j\|_{C^0}^{\eta} \leq \frac{C^2}{r_*}e^{-c\eta n_j}e^{o(n_j)} \to 0$ as $j \to \infty$ for any $\eta > 0$. Then by Proposition 3.5, the measure complexity of $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)}, \rho)$ is sub-polynomial for any $\rho \in \mathcal{M}(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$ and hence the Möbius disjointness holds by Theorem 2.9.

For $\beta(\alpha) > 0$, if there exist sequences $\{W_j\}_{j\in\mathbb{N}} \subseteq C^{\omega}(\mathbb{T}^1, PSL(2,\mathbb{R}))$, $\{\varrho_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}$, and $\{n_j\}_{j\in\mathbb{N}} \subseteq \mathbb{N}$ such that $W_j(\cdot+\alpha)^{-1}A(\cdot)W_j(\cdot) = R_{\varrho_j}(I+G_j(\cdot))$ with $\|W_j\|_{C^1} \leq e^{n_j\eta_j}$, $\|G_j\|_{C^0} \leq e^{-n_j\tau_j}$, $\eta_j = o(\tau_j)$ and $\tau_j n_j > n_j^{1/2}$, then by Proposition 3.5 and Theorem 2.9, the Möbius disjointness conjecture holds. Otherwise, by Proposition 3.3, there is $k_* \in \mathbb{N}$ that for $k \geq k_*$, there exists $W_k \in C^{\omega}(\mathbb{T}^1, PSL(2,\mathbb{R}))$ that $W_k(\cdot+\alpha)^{-1}A(\cdot)W_k(\cdot) = A_k(I + G_k(\cdot))$ with $\|G_k\|_{C^0} \leq Ce^{-Q_k\tau_k}$, $\|W_k\|_{C^1} \leq Ce^{CQ_k\eta_k}$, $A_k = \begin{pmatrix} 1 & c_k \\ 0 & 1 \end{pmatrix}$, where $e^{-\frac{Q_k\tau_k}{10}} \leq |c_k| \leq e^{-Q_k\eta_k}$, $\eta_k = o(\tau_k) < 1$, $\tau_k = o(1)$, $\tau_k Q_k^{\frac{1}{2A}} > 1$ and C is a global constant not depending on k. Moreover, by Remark 2.3, we have that for large enough $q_{k_j} \in \mathfrak{Q}$, it also belongs to $\{Q_k\}_{k\in\mathbb{N}_0}$. Hence, by Proposition 3.6, the Möbius disjointness holds for the corresponding t.d.s. $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$.

□

3.3. **Proof of Theorem 1.2.** If (α, A) is almost reducible and not uniformly hyperbolic, then by Theorem 3.1, the Möbius disjointness conjecture holds. Now we consider the cases (α, A) is uniformly hyperbolic. Actually, for any cocycle (α, A) that has positive Lyapunov exponent, the Möbius disjointness holds for its projective action, which is the content of the following lemma. Thus, we finish the proof.

Lemma 3.7. *Let (α, A) be a continuous $SL(2,\mathbb{R})$-cocycle. If $L(\alpha, A) > 0$, then the Möbius disjointness holds for $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)})$.*

Proof. It follows from the Oseledets Theorem that there exist two $T_{(\alpha,A)}$-invariant graphs $\phi^{\pm}$[28, Section 1.3.2]. Then we can associate a measure $\mu_{\phi^{\pm}}$ respectively given by

$$\mu_{\phi^{\pm}}(\Omega) = \mathrm{Leb}_{\mathbb{T}^1}(\pi_1(\Omega \cap \Phi^{\pm}))$$

for every Lebesgue-measurable set $\Omega \subseteq \mathbb{T}^1 \times \mathbb{RP}^1$, where π_j is the projection to the j-th variable, and $\Phi^{\pm} := \{(x, \phi^{\pm}(x)) : x \in \mathbb{T}^1\} \subseteq \mathbb{T}^1 \times \mathbb{RP}^1$. It is easy to see that $\mu_{\phi^{\pm}}$ are $T_{(\alpha,A)}$-invariant and ergodic. Moreover, $\mu_{\phi^{\pm}}$ are the two only $T_{(\alpha,A)}$-invariant ergodic measures. Let

$$H^{\pm} : \Phi^{\pm} \to \mathbb{T}^1, \ \textit{via} \ \ (x, \phi^{\pm}(x)) \mapsto x.$$

Then $(\mathbb{T}^1 \times \mathbb{RP}^1, \mathcal{B}_{\mathbb{T}^1\times\mathbb{RP}^1}, T_{(\alpha,A)}, \mu_{\phi^{\pm}})$ are measurably-isomorphic to $(\mathbb{T}^1, \mathcal{B}_{\mathbb{T}^1}, R_\alpha, \mathrm{Leb}_{\mathbb{T}^1})$ by $H^{\pm}$, where $R_\alpha : \mathbb{T}^1 \to \mathbb{T}^1$, via $x \mapsto x + \alpha$. In particular, as $(\mathbb{T}^1, R_\alpha, \mathrm{Leb}_{\mathbb{T}^1})$ have discrete spectrum, so do $(\mathbb{T}^1 \times \mathbb{RP}^1, \mathcal{B}_{\mathbb{T}^1\times\mathbb{RP}^1}, T_{(\alpha,A)}, \mu_{\phi^{\pm}})$. As an application of Matomäki-Radziwiłł-Tao's proof of averaged Chowla conjecture [37], it was known [26, Thm. 1.3] that Möbius disjointness conjecture holds for a t.d.s., if it has only countably many ergodic invariant probability measures, all of which have discrete spectrum. This proves the lemma. □

3.4. **Proof of Corollary 1.3.** By Theorem 2.5, for (measure-theoretically) typical $v \in C^\omega(\mathbb{T}^1, \mathbb{R})$, for any $E \in \Sigma_{v,\alpha}$, where $\Sigma_{v,\alpha}$ is the spectrum of $H_{v,\alpha}$, the cocycle $(\alpha, S_{v,E})$ is either supercritical or subcritical. Moreover, it is well known (see [29]) that

$$\Sigma_{v,\alpha} = \{E \in \mathbb{R} \ : \ (\alpha, S_{v,E}) \text{ is not uniformly hyperbolic}\}. \tag{3.1}$$

Therefore, if $E \notin \Sigma_{v,\alpha}$, then $(\alpha, S_{v,E})$ is uniformly hyperbolic. Thus, for (measure-theoretically) typical $v \in C^\omega(\mathbb{T}^1, \mathbb{R})$ and for any $E \in \mathbb{R}$, the cocycle $(\alpha, S_{v,E})$ is either uniformly hyperbolic, supercritical or subcritical. For the case $(\alpha, S_{v,E})$ is supercritical or uniformly hyperbolic, the Möbius disjointness conjecture holds for $T_{(\alpha,S_{v,E})}$ by Lemma 3.7, since $L(\alpha, S_{v,E}) > 0$. If $(\alpha, S_{v,E})$ is subcritical, then by Avila's ART [2, 3], it is almost reducible. Then Theorem 1.2 yields Corollary 1.3.

3.5. **Proof of Corollary 1.5.** Denote $v(x) = \cos(2\pi x)$. We only need to consider the case $\lambda \neq 0$. Otherwise if $\lambda = 0$, then the cocycle is a constant one, and of course it is almost reducible.

If $E \notin \Sigma_{\lambda v,\alpha}$, then by (3.1), the cocycle $(\alpha, S_{\lambda v,E})$ is uniformly hyperbolic. And if $|\lambda| > 1$, then by Corollary 2 in [14],

$$L(\alpha, S_{\lambda v,E}) = \max\{0, \ln|\lambda|\} > 0.$$

Then these two cases follow from Lemma 3.7.

If $0 < |\lambda| < 1$, then by Theorem 19 in [4], the cocycle $(\alpha, S_{\lambda v,E})$ is subcritical. We divide the proof into two cases: $\beta(\alpha) = 0$ and $\beta(\alpha) > 0$. For $\beta(\alpha) > 0$, the cocycle $(\alpha, S_{\lambda v,E})$ is almost reducible by Theorem 1.1 in [2]. For $\beta(\alpha) = 0$, the cocycle $(\alpha, S_{\lambda v,E})$ is almost reducible by Theorem 3.8 in [1]. Then the result follows from Theorem 1.2.

4. Quantitative almost reducibility for $\beta(\alpha)=0$: proof of Proposition 3.2

In the case $\beta(\alpha)=0$, we will use the quantitative almost localization of the dual model to get the desired result of the original system.

We now give some concepts and known results related to our proof.

4.1. Aubry duality and almost localization.

Let $v:\mathbb{T}^1\to\mathbb{R}$ be analytic and $H=H_{v,\alpha,x}:\ell^2(\mathbb{Z})\to\ell^2(\mathbb{Z})$ be the quasiperiodic Schrödinger operator with $v(x)=\sum_{k\in\mathbb{Z}}\hat{v}_k e^{2\pi ikx}$. Then the dual operator $\hat{H}_{v,\alpha,\theta}$ given by

$$(\hat{H}_{v,\alpha,\theta}\hat{u})_n=\sum_{k\in\mathbb{Z}}\hat{v}_k\hat{u}_{n-k}+2\cos(2\pi(\theta+n\alpha))\hat{u}_n,$$

is defined on $\ell^2(\mathbb{Z})$. It has the property that if $u:\mathbb{T}^1\to\mathbb{C}$ is an L^2 function such that its Fourier coefficients satisfy $\hat{H}_{v,\alpha,\theta}\hat{u}=E\hat{u}$, then

$$S_{v,E}(x)\cdot U(x)=e^{2\pi i\theta}U(x+\alpha),$$

where $U(x)=\begin{pmatrix} e^{2\pi i\theta}u(x)\\ u(x-\alpha)\end{pmatrix}$.

Now, we give some necessary and useful concepts.

Definition 4.1 (Resonances)**.** Fix $\epsilon_0>0$ and $\theta\in\mathbb{R}$. An integer $k\in\mathbb{Z}$ is called an ϵ_0-resonance of θ if $\|2\theta-k\alpha\|_{\mathbb{T}}\le e^{-\epsilon_0|k|}$ and $\|2\theta-k\alpha\|_{\mathbb{T}}=\min_{|j|\le|k|}\|2\theta-j\alpha\|_{\mathbb{T}}$.

We denote by $\{n_j\}_{j\in\mathbb{Z}}$ the set of ϵ_0-resonances of θ, ordered in such a way that $|n_1|\le|n_2|\le\cdots$. We say that θ is ϵ_0-resonant if the set $\{n_j\}_j$ is infinite. In particular, by direct computation, one can see that if $\beta(\alpha)=0$, then $\|2\theta-k\alpha\|_{\mathbb{T}}\le e^{-\epsilon_0|k|}$ implies $\|2\theta-k\alpha\|_{\mathbb{T}}=\min_{|j|\le|k|}\|2\theta-j\alpha\|_{\mathbb{T}}$ for k large. Moreover, the following inequality holds.

Lemma 4.1 ([1])**.** *If $\beta(\alpha)=0$, then we have*

$$e^{\epsilon_0|n_j|}\le L_j:=\|2\theta-n_j\alpha\|_{\mathbb{T}}^{-1}\le e^{o(|n_{j+1}|)}.$$

Definition 4.2 (Almost localization)**.** The family $\{\hat{H}_{v,\alpha,\theta}\}_{\theta\in\mathbb{R}}$ is almost localized if there exist constants $C_0,C_1,\epsilon_0,\epsilon_1>0$ such that for every solution $\hat{u}$ of $\hat{H}\hat{u}=E\hat{u}$ satisfying $\hat{u}_0=1$ and $|\hat{u}_k|\le 1+|k|$, we have

$$|\hat{u}_k|\le C_1e^{-\epsilon_1|k|},\quad \forall\ C_0(1+|n_j|)\le|k|\le C_0^{-1}|n_{j+1}|, \tag{4.1}$$

where $\{n_j\}_j$ is the set of ϵ_0-resonances of θ.

In order to get the almost reducibility of the Schrödinger cocycles, we will actually need the almost localization of the dual operators, which is the following result of Avila and Jitomirskaya [6]:

Theorem 4.2 ([6])**.** *Let $\alpha\in\mathbb{R}\backslash\mathbb{Q}$ satisfy $\beta(\alpha)=0$. There exists an absolute constant $c_0>0$ such that for any given $0<r_0<1, C_0>1$, there exist $\epsilon_0=\epsilon_0(r_0)>0,\epsilon_1=\epsilon_1(r_0,C_0)\in(0,r_0)$ and $C_1=C_1(\alpha,r_0,C_0)>0$ such that the following holds: given any $v\in C^\omega(\mathbb{T}^1,\mathbb{R})$ satisfying $\|v\|_{r_0}\le c_0r_0^3$, the family $\{\hat{H}_{v,\alpha,\theta}\}_{\theta\in\mathbb{R}}$ is almost localized with parameters $C_0,C_1,\epsilon_0,\epsilon_1$ as in (4.1). For $v(x)=2\lambda cos(2\pi x)$, the conclusion holds for $0<|\lambda|<1$.*

While almost reducibility allows one to conjugate the dynamics of cocycles close to constant ones, it is rather convenient to have the conjugated cocycles in Schrödinger form, since many results (particularly the ones depending on Aubry duality, c.f.[6]) are obtained only in this setting. The following lemma obtained by Avila and Jitomirskaya [7] takes care of this:

Lemma 4.3 ([7])**.** *Let $(\alpha, A) \in \mathbb{R}\backslash\mathbb{Q} \times C^\omega(\mathbb{T}^1, SL(2,\mathbb{R}))$ be almost reducible. Then there exists $h_* > 0$ such that for every $\kappa > 0$, there is $v \in C^\omega_{h_*}(\mathbb{T}^1, \mathbb{R})$ with $\|v\|_{h_*} < \kappa$, $E \in \mathbb{R}$ and $B \in C^\omega_{h_*}(\mathbb{T}^1, PSL(2,\mathbb{R}))$ such that*

$$B(\cdot+\alpha)A(\cdot)B(\cdot)^{-1} = S_{v,E}(\cdot).$$

4.2. **Proof of Proposition 3.2.** By Lemma 4.3, there exists $h_* > 0$ such that for any $\kappa > 0$, there is $v \in C^\omega_{h_*}(\mathbb{T}^1, \mathbb{R})$ with $\|v\|_{h_*} < \kappa$, $E \in \mathbb{R}$ and $B \in C^\omega_{h_*}(\mathbb{T}^1, PSL(2,\mathbb{R}))$ such that

$$B(x+\alpha)A(x)B(x)^{-1} = S_{v,E}(x). \tag{4.2}$$

Now, we let $r_0 = \min\{h_*, \frac{1}{2}\}$, $\kappa = c_0 r_0^3$ and $c_0, \epsilon_0, \epsilon_1, C_0, C_1$ be the parameter defined as in Theorem 4.2. Then by Theorem 4.2, the family $\{\hat{H}_{v,\alpha,\theta}\}_{\theta\in\mathbb{R}}$ is almost localized with parameters $\epsilon_0, \epsilon_1, C_0, C_1$. Since (α, A) is not uniformly hyperbolic, it is the same for $(\alpha, S_{v,E})$, and then $E \in \Sigma_{v,\alpha}$. By Theorem 3.3 in [6], there exist some $\theta = \theta(E) \in \mathbb{R}$ and $\hat{u} = (\hat{u}_k)_{k\in\mathbb{Z}}$ such that $\hat{H}_{v,\alpha,\theta}\hat{u} = E\hat{u}$ with $\hat{u}_0 = 1$ and $|\hat{u}_k| \leq 1$, $(\forall k \in \mathbb{Z})$. Moreover, Theorem 4.2 implies that $|\hat{u}_k| \leq C_1 e^{-\epsilon_1|k|}$, $\forall\ C_0(1+|n_j|) \leq |k| \leq C_0^{-1}|n_{j+1}|$, where $\{n_j\}_j$ is the set of ϵ_0-resonances of θ. We fix $\theta = \theta(E)$ in the sequel. We divide the proof into two cases: non-ϵ_0-resonant case and ϵ_0-resonant case.

Case (i): θ is not ϵ_0-resonant.

If θ is not ϵ_0-resonant, then by almost localization estimate, we obtain that $\hat{u}$ is localized, i.e. $|\hat{u}_k| \leq C_1 e^{-\epsilon_1|k|}$ for large enough k, that is, it is the Fourier coefficients of an analytic function. Classical Aubry duality yields a connection between localization and reducibility: Indeed by Theorem 2.5 in [6], one has the following:

If $2\theta \notin \alpha\mathbb{Z}+\mathbb{Z}$, then there exists $B : \mathbb{T}^1 \to SL(2,\mathbb{R})$ analytic such that

$$B(x+\alpha)A(x)B(x)^{-1} = R_{\pm\theta}.$$

If $2\theta \in \alpha\mathbb{Z}+\mathbb{Z}$, then there exist analytic $B : \mathbb{T}^1 \to PSL(2,\mathbb{R})$ and analytic $\kappa : \mathbb{T}^1 \to \mathbb{R}$ such that

$$B(x+\alpha)A(x)B(x)^{-1} = \begin{pmatrix} \pm 1 & \kappa(x) \\ 0 & \pm 1 \end{pmatrix}.$$

Now since $\beta(\alpha) = 0$, we can further conjugate the cocycle to a constant matrix by solving $\phi(x+\alpha) - \phi(x) = \kappa(x) - \int_0^1 \kappa(x)dx$ with $\int_0^1 \phi(x)dx = 0$. Letting $B'(x) = \begin{pmatrix} \pm 1 & -\phi(x) \\ 0 & \pm 1 \end{pmatrix} B(x)$, we have

$$B'(x+\alpha)A(x)B'(x)^{-1} = \begin{pmatrix} \pm 1 & \int_0^1 \kappa(x)dx \\ 0 & \pm 1 \end{pmatrix}.$$

Therefore, in any case, the cocycle (α, A) is reducible.

Case (ii): θ is ϵ_0-resonant.

We denote by $\{n_j\}_j$ the infinite set of ϵ_0-resonances of θ, and we can actually get the following local almost reducibility lemma. This lemma was presented in [1], and we give the proof here for completeness.

Lemma 4.4. *Given $r_0 \in (0,1), C_0 > 1$, let $v \in C^{\omega}_{r_0}(\mathbb{T}^1, \mathbb{R})$ with $\|v\|_{r_0} < c_0 r_0^3$ and take $\epsilon_0 = \epsilon_0(r_0) > 0$, $\epsilon_1 = \epsilon_1(r_0, C_0) \in (0, r_0)$, $C_1 = C_1(\alpha, r_0, C_0) > 1$ as in Theorem 4.2. Fix some $n = |n_j| < \infty$ and let $N = |n_{j+1}| < \infty$. Then for any $r_1 \in (0, \frac{\epsilon_1}{30\pi})$, there exists $n_* = n_*(\alpha, r_0, r_1, \epsilon_0, \epsilon_1, C_0, C_1)$, such that if $n \geq n_*$ then there is $W : \mathbb{T}^1 \to SL(2,\mathbb{R})$ analytic with $\|W\|_{r_1/2} \leq Ce^{o(N)}$ and*

$$\|W(x+\alpha)^{-1} S_{v,E}(x) W(x) - R_{\sigma\theta}\|_{r_1/2} \leq Ce^{-cN},$$

where $\sigma \in \{\pm 1\}$, C is a large constant and c is a small constant that both depend on $\alpha, r_0, r_1, \epsilon_0, \epsilon_1, C_0, C_1$, but not on E or θ.

Remark 4.5. *Same as in Theorem 4.2, if $v = 2\lambda\cos(2\pi x)$, then the conclusion holds for $0 < |\lambda| < 1$.*

Then in the ϵ_0-resonant case, by Lemma 4.4, for any $r_1 \in (0, \frac{\epsilon_1}{30\pi})$ there exists n_* only depending on $\alpha, r_0, r_1, \epsilon_0, \epsilon_1, C_0, C_1$, that for any $|n_j| \geq n_*$ there exists $\tilde{W}_j \in C^{\omega}_{r_1/2}(\mathbb{T}^1, SL(2,\mathbb{R}))$ with $\|\tilde{W}_j\|_{r_1/2} \leq Ce^{o(|n_{j+1}|)}$ and $\|\tilde{W}_j(x+\alpha)^{-1} S_{v,E}(x) \tilde{W}_j(x) - R_{\sigma\theta}\|_{r_1/2} \leq Ce^{-c|n_{j+1}|}$. Let $W_j(x) = B(x)^{-1}\tilde{W}_j(x)$, where $B(x)$ is defined in (4.2). Then we have $\|W_j\|_{r_1/2} \leq \|B\|_{r_1/2} \cdot \|\tilde{W}_j\|_{r_1/2} < Ce^{o(|n_{j+1}|)}$ for $|n_{j+1}|$ large enough. This completes the proof of Proposition 3.2 as long as the proof of Lemma 4.4 is finished.

4.3. **Proof of Lemma 4.4.** In this subsection, all the constants may depend on $\alpha, r_0, r_1, \epsilon_0, \epsilon_1, C_0, C_1$, but not on E, θ, n or N. In the following C, c represent big and small constant respectively, and $n_* \in \mathbb{N}$ is also a constant.

First by almost localization result of the dual operator, we have the following:

Lemma 4.6. *For any $m \in [4C_0(1+n), C_0^{-1}N]$, let $I = [\,-[\frac{m}{2}], m - [\frac{m}{2}]\,]$. Then*

$$S_{v,E}(x) U^I(x) = e^{2\pi i\theta} U^I(x+\alpha) + e^{2\pi i\theta} \begin{pmatrix} h(x) \\ 0 \end{pmatrix}, \tag{4.3}$$

with $\|h\|_{r_1} \leq Ce^{-\frac{\epsilon_1}{3}m}$, where $u^I(x) = \sum_{k\in I} \hat{u}_k e^{2\pi i k x}$, $U^I(x) = \begin{pmatrix} e^{2\pi i\theta} u^I(x) \\ u^I(x-\alpha) \end{pmatrix}$. Moreover, we have

$$\|U^I\|_{r'} \leq C(\epsilon_1, C_1, r') e^{2\pi C_0 r' n}, \tag{4.4}$$

for any $0 < r' \leq r_1$.

Proof. Since $\hat{H}_{v,\alpha,\theta}\hat{u} = E\hat{u}$, a direct computation shows that (4.3) holds, where the Fourier coefficients $(\hat{h}_k)_{k\in\mathbb{Z}}$ of h satisfy

$$\hat{h}_k = \chi_I(k)(\,E - 2\cos(2\pi(\theta + k\alpha))\,)\hat{u}_k - \sum_{l\in\mathbb{Z}} \chi_I(k-l)\hat{u}_{k-l}\hat{v}_l, \tag{4.5}$$

and χ_I is the characteristic function of I. Since $\hat{H}_{v,\alpha,\theta}\hat{u} = E\hat{u}$, we also have

$$\hat{h}_k = -\chi_{\mathbb{Z}\setminus I}(k)(\,E - 2\cos(2\pi(\theta + k\alpha))\,)\hat{u}_k + \sum_{l\in\mathbb{Z}} \chi_{\mathbb{Z}\setminus I}(k-l)\hat{u}_{k-l}\hat{v}_l. \tag{4.6}$$

Now we claim that for all $k \in \mathbb{Z}$,

$$|\hat{h}_k| \le C(C_1, r_0, \epsilon_1) e^{-\frac{\epsilon_1}{3}|k|} e^{-\frac{\epsilon_1}{3}m}.$$

Recall that by Theorem 4.2, for $\frac{m}{4} \le |k| \le m$ we have $|\hat{u}_k| \le C_1 e^{-\epsilon_1 |k|}$, and $|\hat{u}_k| \le 1$ for all $k \in \mathbb{Z}$. Moreover, the Fourier coefficients of v satisfy $|\hat{v}_k| \le \|v\|_{r_0} e^{-2\pi r_0 |k|} \le c_0 r_0^3 e^{-2\pi r_0 |k|}$.

For $k \in I$, we have $|k| \le \frac{m}{2} + 1$. Then by (4.6),

$$\begin{aligned} |\hat{h}_k| &= |\sum_{l \in \mathbb{Z}} \chi_{\mathbb{Z} \setminus I}(k-l) \hat{u}_{k-l} \hat{v}_l| \le \sum_{|k-l| \ge \frac{m}{2}} |\hat{u}_{k-l} \hat{v}_l| \\ &\le \sum_{\frac{m}{2} \le |k-l| \le m} C_1 c_0 r_0^3 e^{-\epsilon_1 |k-l|} e^{-2\pi r_0 |l|} + \sum_{|l| > \frac{m}{2} - 1} c_0 r_0^3 e^{-2\pi r_0 |l|} \\ &\le C_1 c_0 r_0^3 e^{-\frac{m}{2}\epsilon_1} \sum_{l \in \mathbb{Z}} e^{-2\pi r_0 |l|} + c_0 r_0^3 \sum_{|l| > \frac{m}{2} - 1} e^{-2\pi r_0 |l|} \\ &\le C(C_1, r_0) e^{-\frac{\epsilon_1}{2} m} \\ &\le C(C_1, r_0, \epsilon_1) e^{-\frac{\epsilon_1}{3} m} e^{-\frac{\epsilon_1}{3}|k|}. \end{aligned}$$

If $k \notin I$, then $|k| \ge \frac{m}{2}$. By (4.5),

$$\begin{aligned} |\hat{h}_k| &= |\sum_{l \in \mathbb{Z}} \chi_I(k-l) \hat{u}_{k-l} \hat{v}_l| \le \sum_{|k-l| \le \frac{m}{2} + 1} |\hat{u}_{k-l} \hat{v}_l| \\ &= \left(\sum_{|k-l| < \frac{m}{4}} + \sum_{\frac{m}{4} \le |k-l| \le \frac{m}{2} + 1} \right) |\hat{u}_{k-l} \hat{v}_l| \\ &\le \sum_{|l| > \frac{|k|}{2}} c_0 r_0^3 e^{-2\pi r_0 |l|} + \sum_{l \in \mathbb{Z}} C_1 c_0 r_0^3 e^{-\epsilon_1 |k-l|} e^{-2\pi r_0 |l|} \\ &\le C(r_0, C_1, \epsilon_1) e^{-\epsilon_1 |k|} \\ &\le C(r_0, C_1, \epsilon_1) e^{-\frac{\epsilon_1}{3} m} e^{-\frac{\epsilon_1}{3}|k|}. \end{aligned}$$

Therefore,

$$\|h\|_{r_1} \le \sum_{k \in \mathbb{Z}} |\hat{h}_k| e^{2\pi r_1 |k|} \le C e^{-\frac{\epsilon_1}{3} m} \sum_{k \in \mathbb{Z}} e^{-(\frac{\epsilon_1}{3} - 2\pi r_1)|k|} \le C e^{-\frac{\epsilon_1}{3} m}.$$

Now we can give the upper bound of $\|U^I(x)\|$ for $|\Im x| \le r'$: $\forall 0 < r' \le r_1$

$$\begin{aligned} \|u^I\|_{r'} &\le \sum_{k \in I} |\hat{u}_k| e^{2\pi r' |k|} \le \sum_{|k| < C_0(1+n)} e^{2\pi r' |k|} + \sum_{C_0(1+n) \le |k| \le \frac{m}{2} + 1} C_1 e^{-\epsilon_1 |k|} e^{2\pi r' |k|} \\ &\le \frac{2 e^{2\pi r'}}{e^{2\pi r'} - 1} e^{2\pi r' C_0 (1+n)} + \frac{2 C_1}{1 - e^{-(\epsilon_1 - 2\pi r')}} e^{-(\epsilon_1 - 2\pi r') C_0 (1+n)} \\ &\le C(\epsilon_1, C_1, r') e^{2\pi r' C_0 n}. \end{aligned}$$

□

Once we have this, we can fix $n = |n_j|, N = |n_{j+1}|$, $I_0 = [\ -[\frac{C_0^{-1}}{2} N], C_0^{-1} N - [\frac{C_0^{-1}}{2} N]\ ]$. Let $U^{I_0}(x) = \begin{pmatrix} e^{2\pi i \theta} u^{I_0}(x) \\ u^{I_0}(x - \alpha) \end{pmatrix}$, where $u^{I_0}(x) = \sum_{k \in I_0} \hat{u}_k e^{2\pi i k x}$. Let $B(x) =$

$(U^{I_0}(x),\ \overline{U^{I_0}(x)})$. Recall that $L^{-1} := L_j^{-1} = \|2\theta - n_j\alpha\|_{\mathbb{T}}$. Since θ is ϵ_0-resonant, then

$$e^{\epsilon_0 n} \le L \le e^{o(N)}$$

for n large enough by Lemma 4.1.

We need the following auxiliary lemma to prove Lemma 4.4, which is proved in [6].

Lemma 4.7 ([6]). *There exists $n_* \in \mathbb{N}$ such that if $n \ge n_*$, then*

$$\inf_{x\in\mathbb{T}^1} |\det B(x)| \ge cL^{-C_*}.$$

where C_* is a constant that depends on $\alpha, \epsilon_0, \epsilon_1, r_0, r_1, C_0, C_1$.

The following lemmas were also mentioned in [6].

Lemma 4.8 ([6]). *Let $x_0 \in \mathbb{T}^1$. Then there exists $n_* \in \mathbb{N}$ that if $n \ge n_*$, then*

$$\sup_{|\Im z| \le r_1/2} |\det B(z) - \det B(x_0)| \le e^{-cN}.$$

Lemma 4.9 ([6]). *There exists $n_* \in \mathbb{N}$ that for $n \ge n_*$ we have*

$$\inf_{|\Im z|\le r_1/2} |\Im \det B(z)| \ge cL^{-C_*},$$

where C_* is the same as in Lemma 4.7.

Proof. Note that $\Re \det B(x) = 0$ for all $x \in \mathbb{T}^1$. The result follows directly from Lemma 4.7, Lemma 4.8 and Lemma 4.1. □

Proof of Lemma 4.4. Take $S = \Re U^{I_0}(x), T = \Im U^{I_0}(x)$ on $\mathbb{T}^1$. Then we have $B(x) = (S(x), \pm T(x)) \begin{pmatrix} 1 & 1 \\ \pm i & \mp i \end{pmatrix}$, and thus $\det B(x) = \mp 2i \det(S(x), \pm T(x))$, which implies that $\det(S(x), \pm T(x))$ does not change the sign as $x \in \mathbb{T}^1$ changes by Lemma 4.7. Let $\sigma \in \{\pm 1\}$ be chosen such that $\det(S, \sigma T) > 0$ and we denote $\tilde{W}(x) := (S(x), \sigma T(x))$. Since

$$\begin{pmatrix} 1 & 1 \\ \sigma i & -\sigma i \end{pmatrix} \begin{pmatrix} e^{2\pi i\theta} & 0 \\ 0 & e^{-2\pi i\theta} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \sigma i & -\sigma i \end{pmatrix}^{-1} = R_{-\sigma\theta},$$

by Lemma 4.6 we have

$$\|S_{v,E}(x)\tilde{W}(x) - \tilde{W}(x+\alpha)R_{-\sigma\theta}\|_{r_1} \le Ce^{-\frac{\epsilon_1}{3C_0}N},$$

and the complex extension considered here is the holomorphic one. We define $W(x) = \frac{\tilde{W}(x)}{|\det B(x)/2|^{1/2}}$ on $\mathbb{T}^1$, so that $\det W(x) = \frac{\det \tilde{W}(x)}{|\det B(x)/2|} = 1$ (by Lemma 4.9, there is no problem with branching when extending $|\det B(x)|^{-1/2}$ to $|\Im x| \le r_1/2$, since $\Re \det B(z) = 0$ for all $z \in \mathbb{C}$). Then by Lemma 4.9, and (4.4) of Lemma 4.6, we have

$$\begin{aligned} \|W\|_{r_1/2} &= \left\| \frac{\tilde{W}(x)}{|\det B(x)/2|^{1/2}} \right\|_{r_1/2} \le \frac{2\|\tilde{W}\|_{r_1/2}}{\inf_{|\Im x|\le r_1/2} |\det B(x)|^{1/2}} \\ &\le CL^{C_*} e^{\pi C_0 r_1 n} \le CL^{C_* + \frac{\pi C_0 r_1}{\epsilon_0}} \le Ce^{o(N)}, \end{aligned}$$

and furthermore, we have

$$\begin{aligned}&\left\|S_{v,E}(x)W(x)-\left|\frac{\det B(x+\alpha)}{\det B(x)}\right|^{1/2}W(x+\alpha)R_{-\sigma\theta}\right\|_{\frac{r_1}{2}}\\ =&\quad\left\|\frac{S_{v,E}(x)\tilde W(x)-\tilde W(x+\alpha)R_{-\sigma\theta}}{|\det B(x)|^{1/2}}\right\|_{\frac{r_1}{2}}\le Ce^{-\frac{\epsilon_1}{3C_0}N}L^{\frac{C_*}{2}}\end{aligned}$$

for $n\ge n_*=n_*(\alpha,\epsilon_0,\epsilon_1,r_0,r_1,C_0,C_1)$. Moreover, by Lemma 4.8 and 4.9, we have

$$\left\|\left|\frac{\det B(x+\alpha)}{\det B(x)}\right|^{1/2}-1\right\|_{\frac{r_1}{2}}\le\frac{\|\det B(x+\alpha)-\det B(x)\|_{r_1/2}}{\inf_{|\Im x|\le r_1/2}|\det B(x)|}\le Ce^{-cN}L^{C_*}$$

for $n\ge n_*$. In conclusion, we obtain that

$$\begin{aligned}\|S_{v,E}(x)W(x)-W(x+\alpha)R_{-\sigma\theta}\|_{r_1/2}&\le Ce^{-\frac{\epsilon_1}{3C_0}N}L^{\frac{C_*}{2}}+Ce^{-cN}L^{C_*}e^{o(N)}\\ &\le Ce^{-cN}.\end{aligned}$$

That is,

$$W(x+\alpha)^{-1}S_{v,E}(x)W(x)=R_{-\sigma\theta}(I+G(x)),$$

with

$$\|G(x)\|_{r_1/2}\le Ce^{o(N)}e^{-cN}\le Ce^{-cN}$$

for $n\ge n_*(\alpha,\epsilon_0,\epsilon_1,C_0,C_1,r_0,r_1)$.

□

5. Quantitative almost reducibility for $\beta(\alpha)>0$: proof of Proposition 3.3

For the case $\beta(\alpha)>0$, we first embed the local cocycle to a continuous $sl(2,\mathbb{R})$ linear system [45] for technical convenience. Since the linear system is close to a constant system, then it allows us to perform the KAM scheme. To get the quantitative almost reducibility for continuous system, we divide the proof to several different cases in the KAM iteration, which we will explain explicitly later. Since the cocycle is almost reducible if and only if the embedded linear system is almost reducible, in the end, we obtain the quantitative almost reducibility of the cocycle.

As we will show below, any analytic almost reducible $SL(2,\mathbb{R})$-cocycle that is not uniformly hyperbolic can be reduced to a cocycle in the form $(\alpha,R_\varrho e^{\tilde G(\cdot)})$ with $\|\tilde G\|$ small enough (consult section 5.2 for details). Therefore, we consider the reduced cocycle as a perturbation of a rotation, and deal with the corresponding continuous quasi-periodic linear $sl(2,\mathbb{R})$ system by the following local embedding theorem.

Theorem 5.1 ([45])**.** *Suppose that $\alpha\in\mathbb{T}^d$, $h>0$, $G\in C_h^\omega(\mathbb{T}^d,sl(2,\mathbb{R}))$, $A=2\pi\varrho J$ where $\varrho\in\mathbb{T}$ and $J=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$. Then there exist $\epsilon_*=\epsilon_*(A,h,|\alpha|)>0$, $F\in C^\omega_{\frac{h}{1+|\alpha|}}(\mathbb{T}^{d+1},sl(2,\mathbb{R}))$ such that the cocycle $(\alpha,e^Ae^{G(\cdot)})$ is the Poincaré map of*

$$\begin{cases}\dot x=(A+F(\theta))x\\ \dot\theta=\omega=(1,\alpha)\end{cases}$$

with $\|F\|_{\frac{h}{1+|\alpha|}} \le \tilde{C}\|G\|_h$, *provided that* $\|G\|_h < \epsilon_*$, *where* $\tilde{C} = \pi e^{\frac{2\pi h(2|\varrho|+1)}{1+|\alpha|}}$.

Remark 5.2. *By the proof in* [45]*, we know that in Theorem 5.1, for* $A = 2\pi\varrho J$ *we only need* $\epsilon_* < \frac{1}{\tilde{C}^2}$.

For the local continuous quasi-periodic linear $sl(2,\mathbb{R})$ system, we have the following iterative lemma.

5.1. **Iterative Lemma.** Let $\mathcal{A} > 2$ and $\{Q_\iota\}_{\iota\in\mathbb{N}_0}$ be the selected sequence in Lemma 2.2. Let $c_0 = \frac{1}{2\cdot 48^2}$, $0 < h < \frac{1}{120}$, $h_+ = \frac{h}{64}$. If

$$c_0 Q_\iota^{\frac{1}{2\mathcal{A}}} h^2 \ge \mathcal{A}^4, \quad 24\ln(2Q_\iota) < Q_\iota^{\frac{1}{2\mathcal{A}}}, \tag{5.1}$$

then we can get the following iterative lemma with $C > 1$ being a large absolute constant:

Proposition 5.3 (Iterative Lemma)**.** *We consider the continuous* $sl(2,\mathbb{R})$ *linear differential equation*

$$\begin{cases} \dot{x} = (A + F(\theta))x \\ \dot{\theta} = \omega = (1,\alpha) \end{cases} \tag{5.2}$$

where $\|F\|_h \le \varepsilon \le 2e^{-c_0 Q_\iota^{\frac{1}{\mathcal{A}}} h^2}$, $A = 2\pi\varrho J$ *or* $\begin{pmatrix} 0 & c^* \\ 0 & 0 \end{pmatrix}$ *with* $e^{-\frac{1}{6}Q_\iota h^2} \le |c^*| \le e^{-\frac{9}{4}Q_\iota h^4}$. *Then there exists a conjugation map* $B \in C^\omega_{h_+}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ *that conjugates (5.2) to*

$$\begin{cases} \dot{x} = (A_+ + F_+(\theta))x \\ \dot{\theta} = \omega = (1,\alpha) \end{cases} \tag{5.3}$$

where

(1) *either* $A_+ = 2\pi\varrho_+ J$ *with one of the following estimates:*

(i)

$$\begin{cases} \|B\|_{C^1} \le e^{C\tilde{Q}_{\iota+1}h_+^2}, \\ \|F_+\|_{h_+} \le e^{-\tilde{Q}_{\iota+1}h_+} \le e^{-Q_{\iota+1}^{\frac{1}{\mathcal{A}}}h_+} \end{cases}$$

(ii)

$$\begin{cases} \|B\|_{C^1} \le e^{C\tilde{Q}_{\iota+1}h_+^4}, \\ \|F_+\|_{h_+} \le 2e^{-c_0\tilde{Q}_{\iota+1}h_+^2} \le 2e^{-c_0 Q_{\iota+1}^{\frac{1}{\mathcal{A}}}h_+^2} \end{cases}$$

where $\tilde{Q}_{\iota+1}$ *equals either* $\overline{Q}_\iota$ *or* $Q_{\iota+1}$ *according to different situations, and* $\tilde{Q}_{\iota+1} \ge Q_\iota^{\mathcal{A}}$ *in both cases.*

(2) *or* $A_+ = \begin{pmatrix} 0 & c_+^* \\ 0 & 0 \end{pmatrix}$ *with*

$$\begin{cases} \|B\|_{C^1} \le e^{CQ_{\iota+1}h_+^4} \\ \|F_+\|_{h_+} \le 9e^{-\frac{9}{4}Q_{\iota+1}h_+^2} \\ e^{-\frac{1}{6}Q_{\iota+1}h_+^2} \le |c_+^*| \le e^{-\frac{9}{4}Q_{\iota+1}h_+^4} \end{cases}.$$

Remark 5.4. *As we can see, we have uniform control of the size of the conjugation as* $\|B\|_{C^1} \le e^{CQ_{\iota+1}h_+^2}$.

The iterative lemma will be proved in different situations. We summarize different cases and state them in several propositions, leaving the proof in section 5.4. We first introduce some notions. For any $h>0$, denote by $\mathcal{B}_h$ the set of $G\in C_h^\omega(\mathbb{T}^2, sl(2,\mathbb{R}))$ satisfying

$$|G|_h := \sum_{k\in\mathbb{Z}^2} \|\hat{G}(k)\| e^{2\pi|k|h} < \infty.$$

It is direct to check, for any $G\in C_h^\omega(\mathbb{T}^2, sl(2,\mathbb{R}))$ and $0<h'<h$, we have

$$|G|_{h'} \leq \frac{36}{\min\{1,(h-h')^2\}} \|G\|_h.$$

In the following propositions, let $0<h'<\frac{1}{160}$, $\tilde{h}=\frac{h'}{6}$, and we suppose that q_n is large enough such that

$$2c_0 q_n^{\frac{1}{2\mathcal{A}}} h'^2 \geq \mathcal{A}^4, \quad 24\ln(2q_n) < q_n^{\frac{1}{2\mathcal{A}}}. \tag{5.4}$$

Now according to the types of the constant matrix A(elliptic or parabolic), and the arithmetic property of the frequency α (Diophantine or Liouvillean), we divide the full proof into four different cases:

If A is elliptic, and the frequency α is relatively Diophantine (i.e. $q_{n+1}\leq q_n^{\mathcal{A}}$), then we have the following:

Proposition 5.5. *Consider the system (5.2) with $A=2\pi\varrho J$ and $|F|_{h'}\leq\epsilon\leq 2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2}$. Suppose that $q_{n+l}\leq q_n^{\mathcal{A}^4}$ for some $l\in\mathbb{N}$. Let $\Lambda_2 := \{k\in\mathbb{Z}^2 : |2\varrho\pm\langle k,\omega\rangle|\geq\epsilon^{\frac{1}{4}}\}$. Then there exists $B\in C_{\tilde{h}}^\omega(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that conjugates (5.2) to*

$$\begin{cases} \dot{x} = (\tilde{A}_1 + \tilde{F}_1(\theta))x \\ \dot{\theta} = \omega = (1,\alpha) \end{cases}, \tag{5.5}$$

such that

(a) if $\Lambda_2^c\cap\{k\in\mathbb{Z}^2 : |k|<\frac{q_{n+l}}{2}\}=\emptyset$, then $\tilde{A}_1=2\pi\bar{\varrho}J$ for some $\bar{\varrho}\in\mathbb{R}$, $\|B\|_{C^1}<2$, and $\|\tilde{F}_1\|_{\tilde{h}}\leq\epsilon e^{-q_{n+l}h'}$;

(b) if $\Lambda_2^c\cap\{k\in\mathbb{Z}^2 : |k|<\frac{q_{n+l}}{2}\}\neq\emptyset$, then we have the following:

(i) either $\tilde{A}_1=2\pi\bar{\varrho}J$ for some $\bar{\varrho}\in\mathbb{R}$, with

- *either $\|B\|_{C^1}\leq e^{2q_{n+l}h'^2}$, $\|\tilde{F}_1\|_{\tilde{h}}\leq 4e^{-\frac{1}{2}q_{n+l}h'}$,*
- *or $\|B\|_{C^1}\leq 8\pi q_{n+l}$, $\|\tilde{F}_1\|_{\tilde{h}}\leq 2e^{-\frac{c_0}{3}q_{n+l}h'^2}$;*

(ii) or $\tilde{A}_1=\begin{pmatrix}0 & \bar{c}\\ 0 & 0\end{pmatrix}$ with $\bar{c}\in\mathbb{R}$, $\|B\|_{C^1}\leq e^{q_{n+l}h'^4}$, $\|F_1\|_{\tilde{h}}\leq 9e^{-\frac{3}{4}q_{n+l}h'^2}$, $e^{-\frac{c_0}{3}q_{n+l}h'^2}\leq|\bar{c}|\leq 4e^{-\frac{3}{4}q_{n+l}h'^4}$.

If A is elliptic, and the frequency α is relatively Liouvillean (i.e. $q_{n+1}>q_n^{\mathcal{A}}$), then we have the following:

Proposition 5.6. *Consider the system (5.2) with $A=2\pi\varrho J$ and $|F|_{h'}\leq\epsilon\leq 2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2}$. Suppose that $q_{n+1}>q_n^{\mathcal{A}}$. Let $\Lambda_2 := \{k\in\mathbb{Z}^2 : |2\varrho\pm\langle k,\omega\rangle|\geq\epsilon^{\frac{1}{4}}\}$. Then there exists*

$B\in C^{\omega}_{\tilde h}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ *that conjugates* (5.2) *to*

$$\begin{cases} \dot x = (\tilde A_2 + \tilde F_2(\theta))x \\ \dot\theta = \omega = (1,\alpha) \end{cases}, \tag{5.6}$$

such that

(a) if $\Lambda_2^c \cap \{k\in\mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} = \emptyset$*, then* $\tilde A_2 = 2\pi\bar\varrho J$ *for some* $\bar\varrho\in\mathbb{R}$*, and* $\|B\|_{C^1} < e^{q_{n+1}h'^4}$, $\|\tilde F_2\|_{\frac{h'}{2}} \le \epsilon e^{-\frac{q_{n+1}h'}{2}}$*;*

(b) if $\Lambda_2^c \cap \{k\in\mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} \neq \emptyset$*, then we have the following:*

(i) either $\tilde A_2 = 2\pi\bar\varrho J$ *for some* $\bar\varrho\in\mathbb{R}$*, with*

- *either* $\|B\|_{C^1} \le e^{2q_{n+1}\tilde h^2}$, $\|\tilde F_2\|_{\tilde h} \le 4e^{-\frac{\tilde h}{2}q_{n+1}}$,
- *or* $\|B\|_{C^1} \le e^{q_{n+1}\tilde h^4}$, $\|\tilde F_2\|_{\tilde h} \le 2e^{-\frac{c_0}{3}q_{n+1}\tilde h^2}$*;*

(ii) or $\tilde A_2 = \begin{pmatrix} 0 & \bar c \\ 0 & 0 \end{pmatrix}$ *with* $\bar c\in\mathbb{R}$, $\|B\|_{C^1} \le e^{q_{n+1}\tilde h^4}$, $\|\tilde F_2\|_{\tilde h} \le 9e^{-\frac34 q_{n+1}\tilde h^2}$, $e^{-\frac{c_0}{3}q_{n+1}\tilde h^2} \le |\bar c| \le 4e^{-\frac34 q_{n+1}\tilde h^4}$.

If A is parabolic, and the frequency α is relatively Diophantine (i.e. $q_{n+1}\le q_n^{\mathcal{A}}$), then we have the following:

Proposition 5.7. *Consider the system (5.2) with* $A = \begin{pmatrix} 0 & c^* \\ 0 & 0 \end{pmatrix}$*,* $|c^*|\le 1$*, and* $|F|_{h'} \le \epsilon \le 2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2}$*. Suppose that* $q_{n+l} \le q_n^{\mathcal{A}^4}$ *for some* $l\in\mathbb{N}$*. Then there exists* $B\in C^{\omega}_{\tilde h}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ *that conjugates* (5.2) *to*

$$\begin{cases} \dot x = (\tilde A_3 + \tilde F_3(\theta))x \\ \dot\theta = \omega = (1,\alpha) \end{cases}, \tag{5.7}$$

such that

(i) either $\tilde A_3 = 2\pi\bar\varrho J$ *for some* $\bar\varrho\in\mathbb{R}$*, with*

- *either* $\|B\|_{C^1} \le 4e^{q_{n+l}h'^2}$, $\|\tilde F_3\|_{\tilde h} \le 4e^{-\frac12 q_{n+l}h'}$,
- *or* $\|B\|_{C^1} \le 4$, $\|\tilde F_3\|_{\tilde h} \le 2e^{-\frac{c_0}{3}q_{n+l}h'^2}$*;*

(ii) or $A_3 = \begin{pmatrix} 0 & \bar c \\ 0 & 0 \end{pmatrix}$ *with* $\bar c\in\mathbb{R}$, $\|B\|_{C^1} \le 4e^{\frac12 q_{n+l}h'^4}$, $\|\tilde F_3\|_{\tilde h} \le 9e^{-\frac34 q_{n+l}h'^2}$, $e^{-\frac{c_0}{3}q_{n+l}h'^2} \le |\bar c| \le 4e^{-\frac34 q_{n+l}h'^4}$.

If A is parabolic, and the frequency α is relatively Liouvillean (i.e. $q_{n+1} > q_n^{\mathcal{A}}$), then we have the following:

Proposition 5.8. *Consider the system (5.2) with* $A = \begin{pmatrix} 0 & c^* \\ 0 & 0 \end{pmatrix}$*,* $|c^*| \le e^{-\frac94 q_n h'^4}$*, and* $|F|_{h'} \le \epsilon \le 2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2}$*. If* $q_{n+1} > q_n^{\mathcal{A}}$*, then there exists* $B\in C^{\omega}_{\tilde h}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ *that conjugates* (5.2) *to*

$$\begin{cases} \dot x = (\tilde A_4 + \tilde F_4(\theta))x \\ \dot\theta = \omega = (1,\alpha) \end{cases}, \tag{5.8}$$

such that

(i) *either* $\tilde{A}_4 = 2\pi\bar{\varrho}J$ *for some* $\bar{\varrho} \in \mathbb{R}$, *with*
- *either* $\|B\|_{C^1} \le e^{2q_{n+1}\tilde{h}^2}$, $\|\tilde{F}_4\|_{\tilde{h}} \le 4e^{-\frac{1}{2}q_{n+1}\tilde{h}}$,
- *or* $\|B\|_{C^1} \le e^{q_{n+1}\tilde{h}^4}$, $\|\tilde{F}_4\|_{\tilde{h}} \le 2e^{-\frac{c_0}{3}q_{n+1}\tilde{h}^2}$;

(ii) *or* $\tilde{A}_4 = \begin{pmatrix} 0 & \bar{c} \\ 0 & 0 \end{pmatrix}$ *with* $\bar{c} \in \mathbb{R}$, $\|B\|_{C^1} \le e^{q_{n+1}\tilde{h}^4}$, $\|\tilde{F}_4\|_{\tilde{h}} \le 9e^{-\frac{3}{4}q_{n+1}\tilde{h}^2}$, $e^{-\frac{c_0}{3}q_{n+1}\tilde{h}^2} \le |\bar{c}| \le 4e^{-\frac{3}{4}q_{n+1}\tilde{h}^4}$.

Proof of Propositon 5.3. Once we have these preparing propostions, we can now finish the proof of Propositon 5.3. Let $h' = \frac{3}{4}h, \tilde{h} = \frac{h'}{6}, h_+ = \frac{\tilde{h}}{8}$. By the definition of $\|\cdot\|_h$ and $|\cdot|_h$, we have

$$|F|_{h'} \le \frac{36}{(h/4)^2}\|F\|_h \overset{(5.1)}{\le} \varepsilon^{\frac{4}{5}} < 2e^{-c_0 Q_\iota^{\frac{1}{\mathcal{A}}} h'^2}.$$

By (5.1), we have Q_ι and h' satisfy (5.4) with q_n replaced by Q_ι.

Now we divide the proof to several cases.

Case 1: $(Q_\iota, Q_{\iota+1})$ *is a* $CD(\mathcal{A}, \mathcal{A}, \mathcal{A}^3)$ *bridge.* Then $Q_{\iota+1} \le Q_\iota^{\mathcal{A}^3} < Q_\iota^{\mathcal{A}^4}$. According to the types of A (elliptic or parabolic) , we apply Proposition 5.5 or Proposition 5.7 respectively. Then there exists $B \in C^\omega_{\tilde{h}}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that conjugates (5.2) to (5.3) with desired estimates.

Case 2: $\overline{Q}_\iota > Q_\iota^{\mathcal{A}}$. In this case, Proposition 5.6 or Proposition 5.8 is applied respectively according to whether A is elliptic or parabolic. Then there exists $B_1 \in C^\omega_{\tilde{h}}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that conjugates (5.2) to

$$\begin{cases} \dot{x} = (\tilde{A} + \tilde{F}(\theta))x \\ \dot{\theta} = \omega = (1, \alpha) \end{cases}, \tag{5.9}$$

where

- either $\tilde{A} = 2\pi\bar{\varrho}J$ with
  - either $\begin{cases} \|B_1\|_{C^1} \le e^{2\overline{Q}_\iota \tilde{h}^2} \\ \|\tilde{F}\|_{\tilde{h}} \le 4e^{-\frac{1}{2}\overline{Q}_\iota \tilde{h}} \end{cases}$,
  - or $\begin{cases} \|B_1\|_{C^1} \le e^{\overline{Q}_\iota \tilde{h}^4} \\ \|\tilde{F}\|_{\tilde{h}} \le 2e^{-\frac{c_0}{3}\overline{Q}_\iota \tilde{h}^2} \end{cases}$;
- or $\tilde{A} = \begin{pmatrix} 0 & \bar{c} \\ 0 & 0 \end{pmatrix}$ with

$$\begin{cases} \|B_1\|_{C^1} \le e^{\overline{Q}_\iota \tilde{h}^4} \\ \|\tilde{F}\|_{\tilde{h}} \le 9e^{-\frac{3}{4}\overline{Q}_\iota \tilde{h}^2} \\ e^{-\frac{c_0}{3}\overline{Q}_\iota \tilde{h}^2} \le |\bar{c}| \le 4e^{-\frac{3}{4}\overline{Q}_\iota \tilde{h}^4} \end{cases}.$$

In this case, we can check that $\overline{Q}_\iota$ and $\frac{3}{4}\tilde{h}$ satisfy (5.4) with $\overline{Q}_\iota$ and $\frac{3}{4}\tilde{h}$ in place of q_n and h' respectively. Moreover, we have

$$|\tilde{F}|_{\frac{3}{4}\tilde{h}} \le \frac{36}{(\tilde{h}/4)^2}\|\tilde{F}\|_{\tilde{h}} \le 2e^{-\frac{4c_0}{15}\overline{Q}_\iota \tilde{h}^2} < 2e^{-c_0 \overline{Q}_\iota^{\frac{1}{2\mathcal{A}}} \tilde{h}^2}.$$

If $\tilde{A}$ is elliptic and $Q_{\iota+1} \le \overline{Q}_\iota^{\mathcal{A}}$, then we let $A_+ := \tilde{A}$, $F_+ := \tilde{F}$, $B := B_1$, and finish one step of iteration.

If $\tilde{A}$ is elliptic and $Q_{\iota+1} > \overline{Q}_\iota^{\mathcal{A}}$, then since $Q_{\iota+1} \le \overline{Q}_\iota^{\mathcal{A}^4}$ by Lemma 2.2, we can apply Proposition 5.5. There exists $B_2 \in C^\omega_{h_+}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that conjugates (5.9) to (5.3). Let $B := B_1 B_2$. Then we get the result.

If $\tilde{A}$ is parabolic, and $Q_{\iota+1} = \overline{Q}_\iota$, then we let $B := B_1, A_+ := \tilde{A}, F_+ := \tilde{F}$, and get the result.

Otherwise for $\tilde{A}$ being parabolic, since $Q_{\iota+1} \le \overline{Q}_\iota^{\mathcal{A}^4}$, then by Proposition 5.7 there exists $B_3 \in C^\omega_{h_+}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that conjugates (5.9) to (5.3). Let $B := B_1 B_3$. Then we finish the proof. □

5.2. **Proof of Proposition 3.3.** Suppose $A \in C^\omega_h(\mathbb{T}^1, SL(2,\mathbb{R}))$ for some $h > 0$. Since (α, A) is C^ω-almost reducible, then in the closure of its analytic conjugate class, there exists a constant cocycle (α, D) and the convergence is uniform in some fixed strip $\{|\Im z| < h'\}$, where $0 < h' < h$.

Let $\mathcal{A} > 2$ and $\{Q_\iota\}_{\iota\in\mathbb{N}}$ be the selected sequence in Lemma 2.2. Without loss of generality, we assume that $0 < h' < \frac{1}{120}$. Let $h_0 = \frac{h'}{1+|\alpha|}$, $c_0 = \frac{1}{2\cdot 48^2}$. Let $n_* \in \mathbb{N}$ be large enough that $c_0 q_{n_*}^{\frac{1}{2\mathcal{A}}} h_0^2 \ge \mathcal{A}^4$, $q_{n_*} \in \mathfrak{Q} \cap \{Q_\iota\}_\iota$ (this is possible by Remark 2.3), and for any $n \ge n_*$, q_n satisfies

$$24 \ln(2q_n) < q_n^{\frac{1}{2\mathcal{A}}}.$$

Assume $q_{n_*} = Q_{\iota_*}$. Let $h_{\iota+1} = \frac{h_\iota}{64}$. Then by the selection of $\{Q_\iota\}_\iota$, we have for any $\iota \in \mathbb{N}_0$,

$$c_0 Q_{\iota_*+\iota}^{\frac{1}{2\mathcal{A}}} h_\iota^2 \ge \mathcal{A}^4, \quad 24 \ln(2Q_{\iota_*+\iota}) < Q_{\iota_*+\iota}^{\frac{1}{2\mathcal{A}}}.$$

As for $D \in SL(2,\mathbb{R})$, there exists $P \in SL(2,\mathbb{R})$ that $\tilde{L}_* := P^{-1}DP$ is in the normal form R_ϱ, $\begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}$, or $\begin{pmatrix} 1 & \tilde{c} \\ 0 & 1 \end{pmatrix}$, where $\lambda, \varrho, \tilde{c} \in \mathbb{R}$ depend on D. Then for any $0 < \varepsilon < \|P\|^{-8}$, there exists $\tilde{B}_1 \in C^\omega_{h'}(\mathbb{T}^1, PSL(2,\mathbb{R}))$ such that

$$\tilde{B}_1(\cdot + \alpha)^{-1} A(\cdot) \tilde{B}_1(\cdot) = D + \tilde{G}_1(\cdot),$$

with $\|\tilde{G}_1\|_{h'} < \varepsilon$. Let $\tilde{B}_2 = \tilde{B}_1 P$. Then

$$\tilde{B}_2(\cdot + \alpha)^{-1} A(\cdot) \tilde{B}_2(\cdot) = \tilde{L}_* + \tilde{G}_2(\cdot),$$

with $\|\tilde{G}_2\|_{h'} = \|P^{-1}\tilde{G}_1 P\|_{h'} < \varepsilon^{\frac{3}{4}}$. If $\tilde{L}_* = \begin{pmatrix} 1 & \tilde{c} \\ 0 & 1 \end{pmatrix}$, then we let $K = \begin{pmatrix} \varepsilon^{-\frac{3}{16}} & 0 \\ 0 & \varepsilon^{\frac{3}{16}} \end{pmatrix}$, and get $K^{-1}\tilde{L}_* K = \begin{pmatrix} 1 & \tilde{c}\varepsilon^{\frac{3}{8}} \\ 0 & 1 \end{pmatrix}$. Now if $\tilde{L}_*$ is parabolic, we let $\tilde{B} = \tilde{B}_2 K$ and $\bar{G} = K^{-1}\tilde{G}_2 K + \begin{pmatrix} 0 & \tilde{c}\varepsilon^{\frac{3}{8}} \\ 0 & 0 \end{pmatrix}$; otherwise, let $\tilde{B} = \tilde{B}_2$ and $\bar{G} = \tilde{G}_2$. Then we have

$$\tilde{B}(\cdot + \alpha)^{-1} A(\cdot) \tilde{B}(\cdot) = L_* + \bar{G}(\cdot),$$

with $\|\bar{G}\|_{h'} < C(D)\varepsilon^{\frac{3}{8}}$, and L_* being a rotation or diagonal matrix.

If L_* is diagonal, i.e., $L_* = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, without loss of generality, we assume that $|\lambda| > 1$. If $|\lambda| - 1 \ge \|\bar{G}\|_{h'}^{1/2}$, then (α, A) is uniformly hyperbolic by the usual cone-field criterion [44], which contradicts with our hypothesis. Therefore in this case $|\lambda|-1\leq \|\bar G\|_{h'}^{1/2}$, and if $\lambda>1$, we have

$$L_*+\bar G=R_0(I+L_*+\bar G-I)$$

with $\|L_*+\bar G-I\|_{h'}\leq 2\|\bar G\|_{h'}^{1/2}$, and if $\lambda<-1$, we have

$$L_*+\bar G=R_{\frac12}(I-(L_*+\bar G+I))$$

with $\|L_*+\bar G+I\|_{h'}\leq 2\|\bar G\|_{h'}^{1/2}$. If L_* is a rotation, denoting $L_*=R_{-\varrho_0}$, then $L_*+\bar G=R_{-\varrho_0}(I+R_{\varrho_0}\bar G)$. Thus in both cases for L_*, by implicit function theorem, for $\|\bar G\|_{h'}$ small enough, i.e., ε small enough, there exists $\tilde G\in C^\omega_{h'}(\mathbb{T}^1,sl(2,\mathbb{R}))$ such that $L_*+\bar G=R_{-\varrho_0}e^{\tilde G}$ with $\varrho_0\in\mathbb{T}^1$, $\|\tilde G\|_{h'}\leq 2\|\bar G\|_{h'}^{\frac12}<\varepsilon^{\frac18}<e^{-Q_{\iota_*}h_0^2}$.

Then by Theorem 5.1, cocycle $(\alpha,R_{-\varrho_0}e^{\tilde G})$ can be embedded into the continuous linear differential equation

$$\begin{cases}\dot x=(\tilde A+\tilde F(\theta))x\\ \dot\theta=\omega=(1,\alpha)\end{cases}\tag{5.10}$$

with $\tilde A=2\pi\varrho_0 J$, $\|\tilde F\|_{h_0}<e^{-c_0Q_{\iota_*}h_0^2}$, and $\Phi^1(0,\cdot)=R_{-\varrho_0}e^{\tilde G(\cdot)}$, where $\Phi^t(\theta)$ is the fundamental solution matrix of (5.10). Let $A_0:=\tilde A, F_0:=\tilde F$. By Proposition 5.3, for $\iota\geq1$ there exist $B_\iota\in C^\omega(\mathbb{T}^2,PSL(2,\mathbb{R}))$, $A_\iota\in sl(2,\mathbb{R})$, and $F_\iota\in C^\omega(\mathbb{T}^2,sl(2,\mathbb{R}))$, such that B_ι conjugates

$$\begin{cases}\dot x=(A_{\iota-1}+F_{\iota-1}(\theta))x\\ \dot\theta=\omega=(1,\alpha)\end{cases}$$

to

$$\begin{cases}\dot x=(A_\iota+F_\iota(\theta))x\\ \dot\theta=\omega=(1,\alpha)\end{cases}\tag{5.11}$$

for either A_ι elliptic or parabolic, with corresponding estimates as in Proposition 5.3. Let $\tilde W_\iota=B_1B_2\cdots B_\iota$. By Remark 5.4, we have for any $\iota\in\mathbb{N}$

$$\|B_\iota\|_{C^1}\leq e^{CQ_{\iota_*+\iota}h_\iota^2}.\tag{5.12}$$

If there exists an infinite sequence $\{\iota_j\}_{j\in\mathbb{N}}$ such that A_{ι_j} is elliptic, then $\tilde W_{\iota_j}$ conjugates (5.10) to (5.11) with following two cases:

Case 1: $\|F_{\iota_j}\|_{C^0}\leq e^{-\tilde Q_{\iota_*+\iota_j}h_{\iota_j}}$, $\|B_{\iota_j}\|_{C^1}\leq e^{C\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^2}$. In this case, by (5.2), (5.12) and the selection of $\{Q_\iota\}_\iota$, we have

$$\|\tilde W_{\iota_j}\|_{C^1}\leq \iota_j e^{C\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^2}e^{\sum_{\iota=1}^{\iota_j-1}CQ_{\iota_*+\iota}h_\iota^2}<e^{2C\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^2}$$

since $\tilde Q_{\iota_*+\iota_j}\geq Q^{\mathcal{A}}_{\iota_*+\iota_j-1}$, and $Q_{\iota+1}\geq Q_\iota^{\mathcal{A}}$ for any $\iota\in\mathbb{N}_0$.

Case 2: $\|F_{\iota_j}\|_{C^0}\leq 2e^{-c_0\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^2}$, $\|B_{\iota_j}\|_{C^1}\leq e^{C\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^4}$. Similarly as Case 1, one can estimate

$$\|\tilde W_{\iota_j}\|_{C^1}\leq \iota_j e^{C\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^4}e^{\sum_{\iota=1}^{\iota_j-1}CQ_{\iota_*+\iota}h_\iota^2}<e^{2C\tilde Q_{\iota_*+\iota_j}h_{\iota_j}^4}.$$

Otherwise, there exists $N_* \in \mathbb{N}$ such that for $\iota \geq N_*$, we have that A_ι is parabolic with $\|F_\iota\|_{C^0} \leq 9e^{-\frac{9}{4}Q_{\iota_*+\iota}h_\iota^2}$, $\|B_\iota\|_{C^1} \leq e^{CQ_{\iota_*+\iota}h_\iota^4}$, and $e^{-\frac{1}{6}Q_{\iota_*+\iota}h_\iota^2} \leq |c_\iota^*| \leq e^{-\frac{9}{4}Q_{\iota_*+\iota}h_\iota^4}$. For $\iota \geq N_*$, combined with estimate (5.12), we obtain that

$$\|\tilde{W}_\iota\|_{C^1} \leq \iota e^{CQ_{\iota_*+\iota}h_\iota^4} e^{\sum_{k=1}^{\iota-1} CQ_{\iota_*+k}h_k^2} < e^{2CQ_{\iota_*+\iota}h_\iota^4}.$$

Now we assume that $\Phi_\iota^t(\theta)$ is the fundamental solution matrix of (5.11). Then

$$\Phi_\iota^t(\theta) = e^{A_\iota t}\left(I + \int_0^t e^{-A_\iota s}F_\iota(\theta + s\omega)\Phi_\iota^s(\theta)ds\right). \tag{5.13}$$

Let $\tilde{G}_\iota^t(\theta) = e^{-A_\iota t}\Phi_\iota^t(\theta)$ and $g_\iota(t) = \|\tilde{G}_\iota^t\|_{C^0}$. Then,

$$\tilde{G}_\iota^t(\theta) = I + \int_0^t e^{-A_\iota s}F_\iota(\theta + \omega s)e^{A_\iota s}\tilde{G}_\iota^s(\theta)ds,$$

and thus $g_\iota(t) \leq 1 + \int_0^t (1+s)^2\|F_\iota\|_{C^0} g_\iota(s)ds$ for A_ι both elliptic and parabolic with $|c_\iota^*| < 1$. By Gronwall's inequality, we have

$$g_\iota(t) \leq e^{\|F_\iota\|_{C^0}\int_0^t (1+s)^2 ds} \leq e^{t(1+t+\frac{t^2}{3})\|F_\iota\|_{C^0}}.$$

Let $t = 1$ in (5.13). Then $\Phi_\iota^1(0, \tilde{\theta}) = e^{A_\iota}(I + G_\iota(\tilde{\theta}))$, with

$$G_\iota(\tilde{\theta}) = \int_0^1 e^{-A_\iota s}F_\iota(s, \tilde{\theta} + s\alpha)e^{A_\iota s}\tilde{G}_\iota^s(0, \tilde{\theta})ds.$$

We get

$$\|G_\iota\|_{C^0} \leq \int_0^1 (1+s)^2\|F_\iota\|_{C^0} g_\iota(s)ds < 8\|F_\iota\|_{C^0}.$$

Since $\tilde{W}_\iota$ conjugates (5.10) to (5.11), then

$$\Phi^t(0, \tilde{\theta}) = \tilde{W}_\iota(t, \tilde{\theta} + t\alpha)\Phi_\iota^t(0, \tilde{\theta})\tilde{W}_\iota^{-1}(0, \tilde{\theta}).$$

Let $\bar{W}_\iota(\tilde{\theta}) = \tilde{W}_\iota(0, \tilde{\theta})$ and $t = 1$. Then $\bar{W}_\iota \in C^\omega(\mathbb{T}^1, PSL(2, \mathbb{R}))$ and we have

$$\bar{W}_\iota(\cdot + \alpha)^{-1}R_{-\varrho_0}e^{\tilde{G}(\cdot)}\bar{W}_\iota(\cdot) = e^{A_\iota}(I + G_\iota(\cdot)).$$

Let $W_\iota = \tilde{B}\bar{W}_\iota$. Then we get the desired result.

5.3. **Modified KAM scheme.** The proof of the Proposition 5.5 - Proposition 5.8 is based on classical KAM scheme [17] and non-standard KAM scheme developed in [23]. There are four steps altogether in the proofs of the KAM iteration. In this subsection, we will first revisit the KAM scheme, and provide necessary ingredients for the proof, and leave the full proof of Proposition 5.5 - Proposition 5.8 to the next subsection.

We will start from the system

$$\begin{cases} \dot{x} = (A + F(\theta))x \\ \dot{\theta} = \omega = (1, \alpha) \end{cases} \tag{5.14}$$

where $|F|_{h'} \leq \epsilon$ for some $0 < h' < \frac{1}{160}$ with $0 < \epsilon < \min\{10^{-8}, 2e^{-c_0 q_n^{\frac{1}{A}} h'^2}\}$, and q_n, h' satisfy (5.4).

Step 1: Elimination of the non-resonant terms.

In this step, we will try to eliminate the non-resonant terms of the perturbation $F(\theta)$ in (5.14). First we give the basic settings. For any $\omega \in \mathbb{R}^2, A \in sl(2, \mathbb{R}), \eta > 0$, we decompose $\mathcal{B}_{h'} = \mathcal{B}_{h'}^{(nre)}(\eta) \oplus \mathcal{B}_{h'}^{(re)}(\eta)$ (depending on A, ω, η) in such a way that for any $Y \in \mathcal{B}_{h'}^{(nre)}(\eta)$,

$$\partial_\omega Y,\ [A, Y] \in \mathcal{B}_{h'}^{(nre)}(\eta), \quad |\partial_\omega Y - [A, Y]|_{h'} \geq \eta |Y|_{h'}.$$

Lemma 5.9 ([23], Lemma 3.1)**.** *Let $\epsilon \in (0, (\frac{1}{10})^8)$ and $\epsilon^{1/4} \leq \eta < 1$. Then for any $F \in \mathcal{B}_{h'}$ satisfying $|F|_{h'} \leq \epsilon$, there exist $Y \in \mathcal{B}_{h'}$ and $F^{(re)} \in \mathcal{B}_{h'}^{(re)}(\eta)$ such that the system*

$$\begin{cases} \dot{x} = (A + F(\theta))x \\ \dot{\theta} = \omega = (1, \alpha) \end{cases}$$

can be conjugate to the system

$$\begin{cases} \dot{x} = (A + F^{(re)}(\theta))x \\ \dot{\theta} = \omega = (1, \alpha) \end{cases} \tag{5.15}$$

by the conjugation map e^Y, with $|Y|_{h'} \leq \epsilon^{1/2}, |F^{(re)}|_{h'} \leq 2\epsilon$.

For $\eta > 0, \tilde{\varrho} \in \mathbb{R}$, let $\Lambda_1(\eta), \Lambda_2(\tilde{\varrho}, \eta)$ be the subsets of $\mathbb{Z}^2$ with $\Lambda_1(\eta) = -\Lambda_1(\eta)$, $\Lambda_2(\tilde{\varrho}, \eta) = -\Lambda_2(\tilde{\varrho}, \eta)$, such that

$$k \in \Lambda_1(\eta) \Rightarrow |\langle k, \omega \rangle| \geq \eta, \quad k \in \Lambda_2(\tilde{\varrho}, \eta) \Rightarrow |2\tilde{\varrho} \pm \langle k, \omega \rangle| \geq \eta.$$

Furthermore, we let

$$\Lambda_{2,1}(\tilde{\varrho}, \eta) := \{k \in \mathbb{Z}^2 \ : \ |2\tilde{\varrho} - \langle k, \omega \rangle| \geq \eta\}, \quad \Lambda_{2,2}(\tilde{\varrho}, \eta) := \{k \in \mathbb{Z}^2 \ : \ |2\tilde{\varrho} + \langle k, \omega \rangle| \geq \eta\}.$$

Then, $\Lambda_{2,2}(\tilde{\varrho}, \eta) = -\Lambda_{2,1}(\tilde{\varrho}, \eta)$. These sets are very useful when one tries to analyze the structure of $F^{(re)}$. In the following, we will analyze this.

Obviously, the structure of $F^{(re)}$ depends on the constant matrix A. If $A = 2\pi\tilde{\varrho}J$, then in order to explore the structure clearly, it is better to state the result in $su(1,1)$. Recall that $su(1,1)$ is the space consisting of matrices of the form $\begin{pmatrix} \mathrm{i}t & v \\ \overline{v} & -\mathrm{i}t \end{pmatrix}$ ($t \in \mathbb{R}$, $v \in \mathbb{C}$), and $sl(2,\mathbb{R})$ is isomorphic to $su(1,1)$ by the rule $A \mapsto MAM^{-1}$, where $M = \frac{1}{1+\mathrm{i}} \begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix}$, since a simple calculation yields

$$M \begin{pmatrix} x & y+z \\ y-z & -x \end{pmatrix} M^{-1} = \begin{pmatrix} \mathrm{i}z & x - \mathrm{i}y \\ x + \mathrm{i}y & -\mathrm{i}z \end{pmatrix}, x, y, z \in \mathbb{R}.$$

Within this context, we have the following:

Corollary 5.10 ([23], Corollary 3.2)**.** *Let $A = 2\pi\tilde{\varrho}J$. Then the conclusion of Lemma 5.9 holds with $F^{(re)}$ in the form*

$$\begin{aligned} MF^{(re)}M^{-1} &= \sum_{k \in \Lambda_1^c} \begin{pmatrix} i\hat{F}_-^{(re)}(k) & 0 \\ 0 & -i\hat{F}_-^{(re)}(k) \end{pmatrix} e^{2\pi i \langle k, \theta \rangle} \\ &+ \sum_{k \in \Lambda_{2,1}^c} \hat{F}_{+,1}^{(re)}(k) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} e^{2\pi i \langle k, \theta \rangle} + \sum_{k \in \Lambda_{2,2}^c} \hat{F}_{+,2}^{(re)}(k) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} e^{2\pi i \langle k, \theta \rangle}, \end{aligned}$$

where $\hat{F}_\pm^{(re)}(k) = \frac{\hat{F}_{12}^{(re)}(k) \pm \hat{F}_{21}^{(re)}(k)}{2}$, $\hat{F}_{+,1}^{(re)}(k) = \hat{F}_{11}^{(re)}(k) - i\hat{F}_+^{(re)}(k)$, $\hat{F}_{+,2}^{(re)}(k) = \hat{F}_{11}^{(re)}(k) + i\hat{F}_+^{(re)}(k)$, and $\Lambda_1 := \Lambda_1(\eta), \Lambda_{2,j} := \Lambda_{2,j}(\tilde{\varrho}, \eta)$ with $j = 1, 2$.

If A is parabolic, then we have the following:

Corollary 5.11. *Let* $A = \begin{pmatrix} 0 & c^* \\ 0 & 0 \end{pmatrix}$ *with* $|c^*| \leq 1$. *Then the conclusion of Lemma 5.9 holds with* $F^{(re)}$ *in the form*

$$F^{(re)} = \sum_{k \in \Lambda_1^c} \hat{F}^{(re)}(k) e^{2\pi i \langle k, \theta \rangle},$$

where $\Lambda_1 := \Lambda_1(\eta^{1/3})$.

Proof. Denote $\tilde{\mathcal{B}}_{h'}^{(nre)} = \{Y \in \mathcal{B}_{h'} \mid Y = \sum_{k \in \Lambda_1} \hat{Y}(k) e^{2\pi i \langle k, \theta \rangle}\}$. Then for any $Y \in \tilde{\mathcal{B}}_{h'}^{(nre)}$, we have $\partial_\omega Y = \sum_{k \in \Lambda_1} 2\pi i \langle k, \omega \rangle \hat{Y}(k) e^{2\pi i \langle k, \theta \rangle} \in \tilde{\mathcal{B}}_{h'}^{(nre)}$,

$$[A, Y] = \begin{pmatrix} c^* Y_{21} & -2c^* Y_{11} \\ 0 & -c^* Y_{21} \end{pmatrix} \in \tilde{\mathcal{B}}_{h'}^{(nre)},$$

and then

$$\partial_\omega Y - [A, Y] = \begin{pmatrix} \partial_\omega Y_{11} - c^* Y_{21} & \partial_\omega Y_{12} + 2c^* Y_{11} \\ \partial_\omega Y_{21} & -\partial_\omega Y_{11} + c^* Y_{21} \end{pmatrix}.$$

Then one can check that $|\partial_\omega Y - [A, Y]|_{h'} \geq \eta |Y|_{h'}$, which means $\tilde{\mathcal{B}}_{h'}^{(nre)} \subseteq \mathcal{B}_{h'}^{(nre)}(\eta)$, and the result follows. □

Therefore, by Corollary 5.10 and Corollary 5.11, in order to analyze the structure of $F^{(re)}$, one only need to analyze the structure of $\Lambda_1^c(\eta)$ and $\Lambda_{2,j}^c(\tilde{\varrho}, \eta)$ $(j = 1, 2)$.

Lemma 5.12. *Let* $\tilde{\varrho} \in \mathbb{R}, 0 < \eta \leq \frac{1}{4 q_n^{\mathcal{A}^4}}$. *If* $q_{n+l} \leq q_n^{\mathcal{A}^4}$ *for some* $l \in \mathbb{N}$, *then*

$$\Lambda_1^c(\eta) \cap \{k \in \mathbb{Z}^2 : |k| < q_{n+l}\} = \{0\}, \tag{5.16}$$

$$\# \left(\Lambda_{2,j}^c(\tilde{\varrho}, \eta) \cap \left\{ k \in \mathbb{Z}^2 : |k| < \frac{q_{n+l}}{2} \right\} \right) \leq 1, \quad (j = 1, 2). \tag{5.17}$$

Proof. For any $k \in \mathbb{Z}^2$ with $0 < |k| < q_{n+l}$, we have

$$|\langle k, \omega \rangle| \geq \frac{1}{2 q_{n+l}} \geq \frac{1}{2 q_n^{\mathcal{A}^4}} > \eta,$$

which implies (5.16).

Moreover, if there exist distinct $k, k' \in \Lambda_{2,1}^c(\tilde{\varrho}, \eta)$ with $|k|, |k'| < \frac{q_{n+l}}{2}$, then we have $|\langle k - k', \omega \rangle| < 2\eta$, which contradicts with the fact

$$|\langle k - k', \omega \rangle| \geq \frac{1}{2 q_{n+l}} \geq \frac{1}{2 q_n^{\mathcal{A}^4}} \geq 2\eta.$$

This implies (5.17) with $j = 1$. Since $\Lambda_{2,2}(\tilde{\varrho}, \eta) = -\Lambda_{2,1}(\tilde{\varrho}, \eta)$, we also obtain (5.17) with $j = 2$. □

In order to analyze the structure of $\Lambda_1^c(\eta)$ and $\Lambda_{2,j}^c(\tilde{\varrho}, \eta)$ $(j = 1, 2)$ for the case $q_{n+1} > q_n^{\mathcal{A}}$, we will need the following lemma.

Lemma 5.13 ([23], Lemma 4.1)**.** *For any* $k = (k_1, k_2) \in \mathbb{Z}^2$ *satisfying*

- $|k| := |k_1| + |k_2| \leq \frac{q_{n+1}}{6}$, *and*
- $k \neq l(p_n, -q_n)$, $l \in \mathbb{Z}$,

we have $|\langle k,\omega\rangle| \geq \frac{1}{7q_n}$.

Remark 5.14. *In fact, in the case* $q_{n+1} > q_n^{\mathcal{A}}$ *with* n *sufficiently large, we have* $|\langle k,\omega\rangle| \geq \frac{1}{7q_n}$ *for any* $|k| \leq \frac{q_{n+1}}{3}$ *and* $k \neq l(p_n, -q_n), l \in \mathbb{Z}$.

Remark 5.15. *Let* $\eta = \epsilon^{1/4}$. *If* $\eta^{1/3} \leq \frac{1}{7q_n}$, *by the above lemma, it is obvious that*

$$\Lambda_1^c(\eta^{1/3}) \cap \left\{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\right\} \subseteq \left\{k = l(p_n, -q_n) : l \in \mathbb{Z}, |k| < \frac{q_{n+1}}{6}\right\}.$$

Moreover, if $|\tilde{\varrho}| < \frac{1}{2}\epsilon^{1/4}$, *then we have*

$$\Lambda_2^c(\tilde{\varrho},\eta) \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} \subseteq \left\{k = l(p_n, -q_n) : l \in \mathbb{Z}, |k| < \frac{q_{n+1}}{6}\right\}.$$

Step 2: Rotation.

For the Diophantine case ($q_{n+1} \leq q_n^{\mathcal{A}}$), we follow the ideas in [17], taking a rotation to eliminate the resonant term. Otherwise, we follow the ideas in [23]: In general, if we don't have the assumption $|\tilde{\varrho}| < \frac{1}{2}\epsilon^{1/4}$, then we have

$$\Lambda_{2,1}^c(\tilde{\varrho},\epsilon^{1/4}) \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} \subseteq \left\{k = k_* + l(p_n, -q_n) : l \in \mathbb{Z}, |k| < \frac{q_{n+1}}{6}\right\},$$

for some $k_* \in \Lambda_{2,1}^c(\tilde{\varrho},\epsilon^{1/4}) \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\}$. In this case, we can take a rotation to make the truncated system a periodic system:

Lemma 5.16 ([23], Lemma 5.2)**.** *Let* $A = 2\pi\tilde{\varrho}J$ *and* $\Lambda_2 := \Lambda_2(\tilde{\varrho},\epsilon^{1/4})$, $\Lambda_{2,j} := \Lambda_{2,j}(\tilde{\varrho},\epsilon^{1/4})$ *(j=1,2). If* $\Lambda_2^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} \neq \emptyset$, *then there exists* $\tilde{\mathcal{Q}} \in C^\omega_{\frac{h'}{3}}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ *that conjugates (5.15) to*

$$\begin{cases} \dot{x} = (A' + F'(\theta))x \\ \dot{\theta} = \omega \end{cases} \tag{5.18}$$

where

$$A' = 2\pi(\tilde{\varrho} - \frac{\langle k_*,\omega\rangle}{2})J, \qquad \mathcal{T}_{\frac{q_{n+1}}{6}}F' = \sum_{\substack{k=l(p_n,-q_n)\\ |k|<\frac{q_{n+1}}{6}}} \hat{F}'(k)e^{2\pi i\langle k,\theta\rangle},$$

with estimate $|F'|_{h'/3} \leq 2\epsilon^{3/4}$. *Moreover,* $\tilde{\mathcal{Q}}(\theta) = \mathcal{Q}(\theta)e^{Y(\theta)}$ *takes the form*

$$\mathcal{Q}(\theta) = \begin{pmatrix} \cos(\pi\langle k_*,\theta\rangle) & \sin(\pi\langle k_*,\theta\rangle) \\ -\sin(\pi\langle k_*,\theta\rangle) & \cos(\pi\langle k_*,\theta\rangle) \end{pmatrix},$$

where $k_* \in \Lambda_{2,1}^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\}$. *Furthermore, we have the following estimates*

$$|Y|_{\frac{h'}{3}} \leq \epsilon^{\frac{3}{8}}, \qquad \|\mathcal{Q}\|_{C^1} \leq 2\pi|k_*| < \frac{\pi}{3}q_{n+1}.$$

Step 3: Floquet's theorem.

By Lemma 5.16, one can observe that due to the special form of the perturbation, the truncated system

$$\begin{cases} \dot{x} = (A' + \mathcal{T}_{\frac{q_{n+1}}{6}}F'(\theta))x = (A' + \sum_{\substack{k=l(p_n,-q_n)\\ |k|<\frac{q_{n+1}}{6}}} \hat{F}'(k)e^{2\pi i\langle k,\theta\rangle})x \\ \dot{\theta} = \omega = (1,\alpha) \end{cases} \tag{5.19}$$

is in fact a periodic system, which can be conjugate to the constant by the famous Floquet's theorem. The following lemma will provide estimates that are sufficient for us.

Lemma 5.17 ([23], Lemma 7.1)**.** *The system*

$$\begin{cases} \dot{x} = G(\theta)x \\ \dot{\theta} = \omega = (1, \alpha) \end{cases}$$

where $G \in \mathcal{B}_{h''}$ for some $h'' > 0$, $|G|_{h''} < \varepsilon$ and is of the form

$$G(\theta_1, \theta_2) = \sum_{l \in \mathbb{Z}} \hat{G}(lp, -lq) e^{2\pi i l(p\theta_1 - q\theta_2)},$$

with $(p, -q) \in \mathbb{Z}^2$ fixed, can be conjugate to some constant system

$$\begin{cases} \dot{x} = D \\ \dot{\theta} = \omega \end{cases}$$

by a conjugation map $B \in C^\omega(\mathbb{T}^2, PSL(2, \mathbb{R}))$, with $D \in sl(2, \mathbb{R})$, and the estimations

$$\|B\|_{h''} \le e^{2\frac{|G|_{h''}}{|\tau|}(2+(|p|+|q|)h'')}, \quad \text{and} \quad \|D\| \le |\tau| e^{\frac{2|G|_{h''}}{|\tau|}},$$

where $\tau = p - q\alpha$.

Step 4: Normalization of the constant matrix.

By Floquet's theorem (Lermma 5.17), the truncated system (5.19) can be conjugate to system with constant matrix $\tilde{A}$, and consequently, the non-truncated system (5.18) is conjugate to

$$\begin{cases} \dot{x} = (\tilde{A} + \tilde{F}(\theta))x \\ \dot{\theta} = \omega \end{cases} . \tag{5.20}$$

However, the constant matrix $\tilde{A}$ may be out of control. Before finishing one step of KAM iteration, we will try to normalize the constant matrix. If the constant matrix $\tilde{A}$ is elliptic or hyperbolic, then we need the following:

Lemma 5.18 ([23], Lemma 8.1)**.** *Let $A \in sl(2, \mathbb{R})$ satisfy $spec(A) = \{i\varrho, -i\varrho\}$ with $0 \ne \varrho \in \mathbb{R}$. There exists $P \in SL(2, \mathbb{R})$ such that $\|P\| \le 2(\frac{\|A\|}{|\varrho|})^{1/2}$ and $P^{-1}AP = \begin{pmatrix} 0 & \varrho \\ -\varrho & 0 \end{pmatrix}$.*

Lemma 5.19 ([38], Proposition 18)**.** *Let $A \in sl(2, \mathbb{R})$ satisfy $spec(A) = \{\lambda, -\lambda\}$ with $0 \ne \lambda \in \mathbb{R}$. There exists $P \in SL(2, \mathbb{R})$ such that $\|P\| \le (\frac{\|A\|}{|\lambda|})^{1/2}$ and $P^{-1}AP = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$.*

If the initial system (5.14) is not uniformly hyperbolic, then the conjugated system (5.20) is also not uniformly hyperbolic, since uniformly hyperbolic is a conjugate invariant. This will provide us additional information, when we try to normalize the constant matrix.

Lemma 5.20. *Let $\mathcal{K} > 0$, $0 < h' < \frac{1}{160}$, and $\mathcal{K}h'^4 > 4$. For the system (5.20), with estimates $\|\tilde{A}\| \le e^{\frac{\mathcal{K}h'^4}{4}}$ and $\|\tilde{F}\|_{h''} \le e^{-\mathcal{K}h'}$ for some $h'' > 0$. If (5.20) is not uniformly hyperbolic, then there exists $P \in SL(2,\mathbb{R})$ that conjugates (5.20) to*

$$\left\{ \begin{array}{l} \dot{x} = (\bar{A} + \bar{F}(\theta))x \\ \dot{\theta} = \omega \end{array} \right. \tag{5.21}$$

where

(1) either $\bar{A} = 2\pi\bar{\varrho}J$ for some $\bar{\varrho} \in \mathbb{R}$ with

(i) either $\left\{ \begin{array}{l} \|P\| \le 2e^{\mathcal{K}h'^2} \\ \|\bar{F}\|_{h''} \le 4e^{-\frac{\mathcal{K}h'}{2}} \end{array} \right.$,

(ii) or $\left\{ \begin{array}{l} \|P\| \le 2 \\ \|\bar{F}\|_{h''} \le 2e^{-\frac{c_0\mathcal{K}h'^2}{3}} \end{array} \right.$;

(2) or $\bar{A} = \begin{pmatrix} 0 & \bar{c} \\ 0 & 0 \end{pmatrix}$ *where $\bar{c} \in \mathbb{R}$ with*

$$\left\{ \begin{array}{l} \|P\| \le 2e^{\frac{\mathcal{K}h'^4}{2}} \\ \|\bar{F}\|_{h''} \le 9e^{-\frac{3}{4}\mathcal{K}h'^2} \\ e^{-\frac{c_0}{3}\mathcal{K}h'^2} \le |\bar{c}| \le 4e^{-\frac{3}{4}\mathcal{K}h'^4} \end{array} \right. .$$

Proof. We denote $\tilde{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{pmatrix}$, and divide the proof into several different cases.

Case 1: $\det \tilde{A} = (2\pi\tilde{\varrho})^2 > 0$. That is, $\tilde{A}$ is elliptic.

Case 1.1: $|2\pi\tilde{\varrho}| \ge e^{-\mathcal{K}h'^2}$. In this case, by Lemma 5.18, there exists $P_1 \in SL(2,\mathbb{R})$ with $\|P_1\| \le 2(\frac{\|\tilde{A}\|}{2\pi\tilde{\varrho}})^{1/2} < 2e^{\mathcal{K}h'^2}$, such that $P_1^{-1}\tilde{A}P_1 = 2\pi\tilde{\varrho}J$. We let $\bar{A} := 2\pi\tilde{\varrho}J$, $P := P_1$ and $\bar{F} := P_1^{-1}\tilde{F}P_1$. Then

$$\|\bar{F}\|_{h''} \le \|P_1\|^2\|\tilde{F}\|_{h''} < 4e^{-\frac{\mathcal{K}h'}{2}}.$$

Case 1.2: $|2\pi\tilde{\rho}| < e^{-\mathcal{K}h'^2}$. In this case, $\det \tilde{A} = -a_{11}^2 - a_{12}a_{21} = (2\pi\tilde{\varrho})^2 > 0$, which implies that

$$|a_{12}a_{21}| = a_{11}^2 + (2\pi\tilde{\varrho})^2 \ge (2\pi\tilde{\varrho})^2 > 0.$$

Hence, we have $\max\{|a_{12}|, |a_{21}|\} \ge |2\pi\tilde{\varrho}| > 0$. Without loss of generality, we assume that $|a_{12}| \ge |2\pi\tilde{\varrho}|$. Then the system (5.20) can be rewritten as

$$\left\{ \begin{array}{l} \dot{x} = (\tilde{A}^{(1)} + \tilde{F}^{(1)}(\theta))x \\ \dot{\theta} = \omega \end{array} \right.$$

where

$$\tilde{A}^{(1)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + \frac{(2\pi\tilde{\varrho})^2}{a_{12}} & -a_{11} \end{pmatrix}, \qquad \tilde{F}^{(1)}(\theta) = \tilde{F}(\theta) + \begin{pmatrix} 0 & 0 \\ -\frac{(2\pi\tilde{\varrho})^2}{a_{12}} & 0 \end{pmatrix},$$

with $\det \tilde{A}^{(1)} = 0$, and

$$\|\tilde{A}^{(1)}\| \le \|\tilde{A}\| + 2\pi|\tilde{\rho}| < 2e^{\frac{\mathcal{K}h'^4}{4}},$$
$$\|\tilde{F}^{(1)}\|_{h''} \le \|\tilde{F}\|_{h''} + 2\pi|\tilde{\varrho}| < 2e^{-\mathcal{K}h'^2}.$$

Then we will reduce it to the following case:

Case 2: $\det \tilde{A} = -\lambda^2 \le 0$. That is, $\tilde{A}$ is hyperbolic or parabolic. First we claim that we only need to consider the case $|\lambda| < (\|\tilde{F}\|_{h''}\|\tilde{A}\|)^{1/3}$, since we assume the system (5.20) is not uniformly hyperbolic. Otherwise, by Lemma 5.19, there exists $P_1 \in SL(2,\mathbb{R})$ such that $\|P_1\| \le (\frac{\|\tilde{A}\|}{|\lambda|})^{1/2}$ and $P_1^{-1}\tilde{A}P_1 = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$. Then (5.20) is conjugate to

$$\begin{cases} \dot{x} = (\begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} + P_1^{-1}\tilde{F}(\theta)P_1)x \\ \dot{\theta} = \omega \end{cases} \tag{5.22}$$

with $\|P_1^{-1}\tilde{F}(\theta)P_1\|_{h''} \le \frac{\|\tilde{A}\|\|\tilde{F}\|_{h''}}{|\lambda|}$. Therefore, if $|\lambda| \ge (\frac{\|\tilde{A}\|\|\tilde{F}\|_{h''}}{|\lambda|})^{1/2}$, i.e., $|\lambda| \ge (\|\tilde{F}\|_{h''}\|\tilde{A}\|)^{1/3}$, the system (5.22) is uniformly hyperbolic by the usual cone-field criterion [44], which contradicts with our assumption.

Then either in this case, or in Case 1.2, we can write the system (5.20) as

$$\begin{cases} \dot{x} = (\tilde{A}^{(2)} + \tilde{F}^{(2)}(\theta))x \\ \dot{\theta} = \omega \end{cases} \tag{5.23}$$

with $\det \tilde{A}^{(2)} = -\lambda^2 \le 0$, $|\lambda| < (\|\tilde{A}\|\|\tilde{F}\|_{h''})^{\frac{1}{3}} < e^{-\frac{\mathcal{K}h'}{4}}$, $\|\tilde{A}^{(2)}\| < 2e^{\frac{\mathcal{K}h'^4}{4}}$, $\|\tilde{F}^{(2)}\|_{h''} < 2e^{-\mathcal{K}h'^2}$. We assume that $\tilde{A}^{(2)} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & -\tilde{a}_{11} \end{pmatrix}$. Then there exists $P_2 \in SL(2,\mathbb{R})$ with $\|P_2\| \le 2$ such that $P_2^{-1}\tilde{A}^{(2)}P_2 = \begin{pmatrix} \lambda & \bar{a}_{12} \\ 0 & -\lambda \end{pmatrix}$ with $\bar{a}_{12} = \tilde{a}_{12} - \tilde{a}_{21}$. Under this conjugation map, (5.23) is conjugate to

$$\begin{cases} \dot{x} = (\tilde{A}^{(3)} + \tilde{F}^{(3)}(\theta))x \\ \dot{\theta} = \omega \end{cases} \tag{5.24}$$

where $\tilde{A}^{(3)} = \begin{pmatrix} 0 & \bar{a}_{12} \\ 0 & 0 \end{pmatrix}$, $\tilde{F}^{(3)} = P_2^{-1}\tilde{F}^{(2)}P_2 + \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$, with

$$|\bar{a}_{12}| \le 4e^{\frac{\mathcal{K}h'^4}{4}}, \quad \|\tilde{F}^{(3)}\|_{h''} \le 9e^{-\mathcal{K}h'^2}.$$

Recall that $c_0 = \frac{1}{2\cdot 48^2}$. Now we consider the following three sub-cases according to the value of $|\bar{a}_{12}|$:

Case 2.1: $|\bar{a}_{12}| \le e^{-\frac{c_0}{3}\mathcal{K}h'^2}$. We let $\bar{A} := 0$, $\bar{F} := \tilde{F}^{(3)} + \tilde{A}^{(3)}$, $P := P_2$. Then $\|P\| \le 2$ and

$$\|\bar{F}\|_{h''} \le \|\tilde{F}^{(3)}\|_{h''} + |\bar{a}_{12}| < 2e^{-\frac{c_0}{3}\mathcal{K}h'^2}.$$

Case 2.2: $e^{-\frac{c_0}{3}\mathcal{K}h'^2} < |\bar{a}_{12}| \le e^{-\frac{3}{4}\mathcal{K}h'^4}$. Let $\bar{A} := \tilde{A}^{(3)}$ with $\bar{c} = \bar{a}_{12}$, $\bar{F} := \tilde{F}^{(3)}$, $P := P_2$. Then $\|P\| \le 2$, $e^{-\frac{c_0}{3}\mathcal{K}h'^2} < |\bar{c}| \le e^{-\frac{3}{4}\mathcal{K}h'^4}$, and

$$\|\bar{F}\|_{h''} = \|\tilde{F}^{(3)}\|_{h''} \le 9e^{-\mathcal{K}h'^2}.$$

Case 2.3: $|\bar{a}_{12}| > e^{-\frac{3}{4}\mathcal{K}h'^4}$. Let $H = \begin{pmatrix} e^{\frac{\mathcal{K}h'^4}{2}} & 0 \\ 0 & e^{-\frac{\mathcal{K}h'^4}{2}} \end{pmatrix}$. Then H conjugates the system (5.24) to (5.21) where $\bar{A} = \begin{pmatrix} 0 & \bar{c} \\ 0 & 0 \end{pmatrix}$ with $\bar{c} = e^{-\mathcal{K}h'^4}\bar{a}_{12}$, and $\bar{F} = H^{-1}\tilde{F}^{(3)}H$.

Let $P := P_2 H$. Then $\|P\| \le 2e^{\frac{\mathcal{K}h'^4}{2}}$, and

$$\|\bar{F}\|_{h''} \le \|H\|^2 \|\tilde{F}^{(3)}\|_{h''} \le 9e^{\mathcal{K}h'^4} e^{-\mathcal{K}h'^2} < 9e^{-\frac{3}{4}\mathcal{K}h'^2}.$$

Moreover, we have $|\bar{c}| > e^{-\frac{7}{4}\mathcal{K}h'^4} > e^{-\frac{c_0}{3}\mathcal{K}h'^2}$, and

$$|\bar{c}| \le 4e^{-\mathcal{K}h'^4} e^{\frac{\mathcal{K}h'^4}{4}} = 4e^{-\frac{3}{4}\mathcal{K}h'^4}.$$

□

5.4. **Proof of Proposition 5.5-5.8.** With the above preparing lemmas, now we can finish the whole proof.

5.4.1. *Proof of Proposition 5.5.* By (5.4), we can check that $2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2} < 10^{-8}$. Then by Lemma 5.9, there exist $Y_1 \in \mathcal{B}_{h'}$ and $F^{(re)} \in \mathcal{B}_{h'}^{(re)}(\epsilon^{\frac{1}{4}})$ such that (5.2) can be conjugate to (5.15) by the conjugation map e^{Y_1}, and $|Y_1|_{h'} \le \epsilon^{1/2}$, $|F^{(re)}|_{h'} \le 2\epsilon$.

(a) If $\Lambda_2^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+l}}{2}\} = \emptyset$, then by Corollary 5.10 and Lemma 5.12, $F^{(re)}$ has the form

$$\mathcal{T}_{\frac{q_{n+l}}{2}} F^{(re)} = \begin{pmatrix} 0 & \hat{F}_-^{(re)}(0) \\ -\hat{F}_-^{(re)}(0) & 0 \end{pmatrix},$$

since $\epsilon^{\frac{1}{4}} < \frac{1}{4q_n^{\mathcal{A}^4}}$ by (5.4). Let $B = e^{Y_1}$, $\tilde{A}_1 = A + \hat{F}_-^{(re)}(0)J =: 2\pi\bar{\varrho}J$, $\tilde{F}_1 = \mathcal{R}_{\frac{q_{n+l}}{2}} F^{(re)}$. Then we have

$$\|B\|_{C^1} = \|e^{Y_1}\|_{C^1} < 2,$$

and

$$\begin{aligned} \|\tilde{F}_1\|_{h'/2} &\le \sum_{|k| \ge q_{n+l}/2} |\hat{F}^{(re)}(k)| e^{\pi|k|h'} \\ &= \sum_{|k| \ge q_{n+l}/2} |\hat{F}^{(re)}(k)| e^{2\pi|k|h'} e^{-\pi|k|h'} \\ &< \epsilon e^{-q_{n+l}h'}. \end{aligned}$$

(b) If $\Lambda_2^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+l}}{2}\} \ne \emptyset$, then by Lemma 5.12, there exists a unique $k_* \in \Lambda_{2,1}^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+l}}{2}\}$. Furthermore, by Corollary 5.10, $F^{(re)}$ has the form

$$M\mathcal{T}_{\frac{q_{n+l}}{2}} F^{(re)} M^{-1} = \begin{pmatrix} i\hat{F}_-^{(re)}(0) & 0 \\ 0 & -i\hat{F}_-^{(re)}(0) \end{pmatrix} + \begin{pmatrix} 0 & \hat{F}_{+,1}^{(re)}(k_*) e^{2\pi i\langle k_*, \theta\rangle} \\ \overline{\hat{F}_{+,1}^{(re)}(k_*)} e^{-2\pi i\langle k_*, \theta\rangle} & 0 \end{pmatrix}.$$

Let $\mathcal{Q}(\theta) = \begin{pmatrix} \cos(\pi\langle k_*, \theta\rangle) & \sin(\pi\langle k_*, \theta\rangle) \\ -\sin(\pi\langle k_*, \theta\rangle) & \cos(\pi\langle k_*, \theta\rangle) \end{pmatrix}$. Then direct calculation shows that $\mathcal{Q}$ conjugates (5.15) to

$$\begin{cases} \dot{x} = (\tilde{A} + \tilde{F}(\theta))x \\ \dot{\theta} = \omega, \end{cases} \tag{5.25}$$

where $\tilde{F} = \mathcal{Q}^{-1}\mathcal{R}_{\frac{q_{n+l}}{2}} F^{(re)}\mathcal{Q}$, and $\tilde{A}$ has the form

$$\tilde{A} = (2\pi\varrho - \pi\langle k_*, \omega\rangle + \hat{F}_-^{(re)}(0))J + M^{-1} \begin{pmatrix} 0 & \hat{F}_{+,1}^{(re)}(k_*) \\ \overline{\hat{F}_{+,1}^{(re)}(k_*)} & 0 \end{pmatrix} M.$$

Then one has

$$\begin{aligned}\|\mathcal{Q}\|_{C^1} &\le 2\pi|k_*| < \pi q_{n+l}\\ \|\tilde{A}\| &\le \pi|2\varrho - \langle k_*,\omega\rangle| + 4\epsilon < \pi\epsilon^{1/4} + 4\epsilon < 8\epsilon^{1/4},\end{aligned}$$

and furthermore, we have

$$\begin{aligned}\|\tilde{F}\|_{h'/6} &\le \|\mathcal{Q}\|^2_{h'/6}\sum_{|k|\ge\frac{q_{n+l}}{2}}|\hat{F}^{(re)}(k)|e^{2\pi|k|\frac{h'}{6}}\\ &\le 2\epsilon e^{\frac{\pi|k_*|h'}{3}}\sum_{|k|\ge\frac{q_{n+l}}{2}}e^{-2\pi|k|h'}e^{2\pi|k|\frac{h'}{6}} < \epsilon e^{-q_{n+l}h'}.\end{aligned}$$

Then by Lemma 5.20, there exists $P\in SL(2,\mathbb{R})$ that conjugates (5.25) to (5.5), where

- either $\tilde{A}_1 = 2\pi\bar{\varrho}J$ for some $\bar{\varrho}\in\mathbb{R}$, with
  - either $\begin{cases}\|P\|\le 2e^{q_{n+l}h'^2}\\ \|\tilde{F}_1\|_{\tilde{h}}\le 4e^{-\frac{1}{2}q_{n+l}h'^2}\end{cases}$,
  - or $\begin{cases}\|P\|\le 2\\ \|\tilde{F}_1\|_{\tilde{h}}\le 2e^{-\frac{c_0}{3}q_{n+l}h'^2}\end{cases}$;
- or $\tilde{A}_1 = \begin{pmatrix}0 & \bar{c}\\ 0 & 0\end{pmatrix}$ where $\bar{c}\in\mathbb{R}$, with

$$\begin{cases}\|P\|\le 2e^{\frac{1}{2}q_{n+l}h'^4}\\ \|\tilde{F}_1\|_{\tilde{h}}\le 9e^{-\frac{3}{4}q_{n+l}h'^2}\\ e^{-\frac{c_0}{3}q_{n+l}h'^2}\le|\bar{c}|\le 4e^{-\frac{3}{4}q_{n+l}h'^4}\end{cases}.$$

Let $B = e^{Y_1}\mathcal{Q}P$. Then we get the desired result. □

5.4.2. *Proof of Proposition 5.6.* Since $2e^{-c_0 q_n^{\frac{1}{A}}h'^2} < 10^{-8}$, then by Lemma 5.9, there exist $Y_1\in\mathcal{B}_{h'}$ and $F^{(re)}\in\mathcal{B}^{(re)}_{h'}(\epsilon^{\frac{1}{4}})$ such that (5.2) can be conjugate to (5.15) by the conjugation map e^{Y_1}, and $|Y_1|_{h'}\le\epsilon^{1/2}$, $|F^{(re)}|_{h'}\le 2\epsilon$.

(a) If $\Lambda_2^c\cap\{k\in\mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} = \emptyset$, then by Corollary 5.10, we have

$$\mathcal{T}_{\frac{q_{n+1}}{6}}F^{(re)} = \sum_{\substack{k\in\Lambda_1^c,\\ |k|<\frac{q_{n+1}}{6}}}\begin{pmatrix}0 & \hat{F}_-^{(re)}(k)\\ -\hat{F}_-^{(re)}(k) & 0\end{pmatrix}e^{2\pi i\langle k,\theta\rangle}.$$

Let

$$E(\theta) = \sum_{\substack{k\in\Lambda_1^c,\\ 0<|k|<\frac{q_{n+1}}{6}}}\begin{pmatrix}0 & \hat{F}_-^{(re)}(k)\\ -\hat{F}_-^{(re)}(k) & 0\end{pmatrix}\frac{e^{2\pi i\langle k,\theta\rangle}}{2\pi i\langle k,\omega\rangle}.$$

Then $e^{E(\theta)}$ conjugates the system (5.15) to

$$\begin{cases}\dot{x} = (A + \hat{F}_-^{(re)}(0)J + e^{-E(\theta)}\mathcal{R}_{\frac{q_{n+1}}{6}}F^{(re)}(\theta)e^{E(\theta)}\ )x\\ \dot{\theta} = \omega\end{cases}.$$

Since for any $0<|k|<\frac{q_{n+1}}{6}$, we have $|\langle k,\omega\rangle|\ge\frac{1}{2q_{n+1}}$, then

$$|E|_{h'}\le\frac{2}{\pi}\epsilon q_{n+1}.$$

Let $B = e^{Y_1}e^E$, $\tilde{A}_2 = A + \hat{F}_-^{(re)}(0)J =: 2\pi\bar{\varrho}J$, $\tilde{F}_2 = e^{-E}\mathcal{R}_{\frac{q_{n+1}}{6}}F^{(re)}e^E$. Then we have

$$\|B\|_{C^1} \leq \frac{1}{h'}\|B\|_{h'} \leq \frac{1}{h'}e^{|Y_1|_{h'}}e^{|E|_{h'}} \leq \frac{2}{h'}e^{\frac{2}{\pi}\epsilon q_{n+1}} < e^{q_{n+1}h'^4},$$

$$\|\tilde{F}_2\|_{h'/2} \leq e^{4q_{n+1}\epsilon/\pi}\sum_{|k|\geq q_{n+1}/6}|\hat{F}^{(re)}(k)|e^{\pi|k|h'} < \epsilon e^{-\frac{q_{n+1}h'}{2}}.$$

(b) If $\Lambda_2^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\} \neq \emptyset$, then by Lemma 5.16, there exists $\tilde{\mathcal{Q}}_1 \in C^\omega_{\frac{h'}{3}}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that conjugates (5.15) to

$$\begin{cases} \dot{x} = (\tilde{A} + \tilde{F}(\theta))x \\ \dot{\theta} = \omega \end{cases}, \tag{5.26}$$

where

$$\tilde{A} = 2\pi(\varrho - \frac{\langle k_*, \omega\rangle}{2})J, \qquad \mathcal{T}_{\frac{q_{n+1}}{6}}\tilde{F} = \sum_{\substack{k=l(p_n,-q_n)\\ |k|<\frac{q_{n+1}}{6}}}\hat{\tilde{F}}(k)e^{2\pi i\langle k,\theta\rangle},$$

with estimate $|\tilde{F}|_{h'/3} \leq 2\epsilon^{3/4}$. Moreover, $\tilde{\mathcal{Q}}_1(\theta) = \mathcal{Q}(\theta)e^{Y_2}$ takes the form

$$\mathcal{Q}(\theta) = \begin{pmatrix} \cos(\pi\langle k_*, \theta\rangle) & \sin(\pi\langle k_*, \theta\rangle) \\ -\sin(\pi\langle k_*, \theta\rangle) & \cos(\pi\langle k_*, \theta\rangle) \end{pmatrix},$$

where $k_* \in \Lambda_{2,1}^c \cap \{k \in \mathbb{Z}^2 : |k| < \frac{q_{n+1}}{6}\}$. Furthermore, we have the following estimates

$$|Y_2|_{\frac{h'}{3}} \leq \epsilon^{\frac{3}{8}}, \qquad \|\mathcal{Q}\|_{C^1} \leq 2\pi|k_*| < \frac{\pi}{3}q_{n+1}.$$

Now we consider the system

$$\begin{cases} \dot{x} = (\tilde{A} + \mathcal{T}_{\frac{q_{n+1}}{6}}\tilde{F})x \\ \dot{\theta} = \omega \end{cases}. \tag{5.27}$$

By Lemma 5.17, there exist $D \in sl(2,\mathbb{R})$ and $L_1 \in C^\omega_{\frac{h'}{3}}(\mathbb{T}^2, PSL(2,\mathbb{R}))$ that L_1 conjugates (5.27) to

$$\begin{cases} \dot{x} = Dx \\ \dot{\theta} = \omega \end{cases}.$$

Since we have

$$|\tilde{A} + \mathcal{T}_{\frac{q_{n+1}}{6}}\tilde{F}|_{\frac{h'}{3}} \leq 4\epsilon^{\frac{1}{4}},$$

then one can compute $\|L_1\|_{\frac{h'}{3}} < e^{64\epsilon^{1/4}h'q_nq_{n+1}}$, and

$$\|D\| \leq \frac{1}{q_{n+1}}e^{16q_{n+1}\epsilon^{1/4}} \leq e^{\frac{q_{n+1}}{4}\tilde{h}^4}. \tag{5.28}$$

Then under the conjugation map L_1, the system (5.26) becomes

$$\begin{cases} \dot{x} = (D + L_1^{-1}\mathcal{R}_{\frac{q_{n+1}}{6}}\tilde{F}L_1)x \\ \dot{\theta} = \omega \end{cases}. \tag{5.29}$$

By (5.4), we have

$$\|L_1^{-1}\mathcal{R}_{\frac{q_{n+1}}{6}}\tilde{F}L_1\|_{\frac{h'}{6}} \leq e^{128\epsilon^{1/4}q_nh'q_{n+1}}\sum_{|k|\geq\frac{q_{n+1}}{6}}|\hat{\tilde{F}}(k)|e^{2\pi|k|\frac{h'}{6}} < \epsilon^{\frac{3}{4}}e^{-q_{n+1}\tilde{h}}.$$

Then by (5.28), one can apply Lemma 5.20, and there exists $P \in SL(2,\mathbb{R})$ that conjugates (5.29) to (5.6), where

- either $\tilde{A}_2 = 2\pi\bar{\varrho}J$ for some $\bar{\varrho} \in \mathbb{R}$ with
  - either $\begin{cases} \|P\| \le 2e^{q_{n+1}\tilde{h}^2} \\ \|\tilde{F}_2\|_{\tilde{h}} \le 4e^{-\frac{1}{2}q_{n+1}\tilde{h}} \end{cases}$,
  - or $\begin{cases} \|P\| \le 2 \\ \|\tilde{F}_2\|_{\tilde{h}} \le 2e^{-\frac{c_0}{3}q_{n+1}\tilde{h}^2} \end{cases}$;
- or $\tilde{A}_2 = \begin{pmatrix} 0 & \bar{c} \\ 0 & 0 \end{pmatrix}$, where $\bar{c} \in \mathbb{R}$, with

$$\begin{cases} \|P\| \le 2e^{\frac{q_{n+1}}{2}\tilde{h}^4} \\ \|\tilde{F}_2\|_{\tilde{h}} \le 9e^{-\frac{3}{4}q_{n+1}\tilde{h}^2} \\ e^{-\frac{c_0}{3}q_{n+1}\tilde{h}^2} \le |\bar{c}| \le 4e^{-\frac{3}{4}q_{n+1}\tilde{h}^4} \end{cases}.$$

Let $B = e^{Y_1}\tilde{\mathcal{Q}}_1 L_1 P$. Then we get the desired result. □

5.4.3. *Proof of Proposition 5.7.* Let $\eta = \epsilon^{\frac{1}{4}}$. Since $2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2} < 10^{-8}$, by Lemma 5.9, there exist $Y_1 \in \mathcal{B}_{h'}$ and $F^{(re)} \in \mathcal{B}_{h'}^{(re)}(\eta)$ such that (5.2) can be conjugate to (5.15) by the conjugation map e^{Y_1}, and $|Y_1|_{h'} \le \epsilon^{1/2}$, $|F^{(re)}|_{h'} \le 2\epsilon$. Then by Lemma 5.12, we have $\Lambda_1^c(\eta^{\frac{1}{3}}) \cap \{k \in \mathbb{Z}^2 : |k| < q_{n+l}\} = \{0\}$, since $\epsilon^{\frac{1}{12}} < \frac{1}{4q_n^{\mathcal{A}^4}}$ by (5.4). Furthermore, by Corollary 5.11, one has $F^{(re)} = \hat{F}^{(re)}(0) + \mathcal{R}_{q_{n+l}}F^{(re)}$. Rewrite the system (5.15) as

$$\begin{cases} \dot{x} = (\tilde{A} + \tilde{F}(\theta))x \\ \dot{\theta} = \omega \end{cases}, \tag{5.30}$$

where $\tilde{A} = A + \hat{F}^{(re)}(0)$, $\tilde{F} = \mathcal{R}_{q_{n+l}}F^{(re)}$. Then $\|\tilde{A}\| \le 2$, and

$$\|\tilde{F}\|_{\frac{h'}{6}} \le \sum_{|k| \ge q_{n+l}} |\hat{F}^{(re)}(k)| e^{2\pi|k|\frac{h'}{6}} \le |F^{(re)}|_{h'} e^{-\frac{5}{3}\pi q_{n+l}h'} < \epsilon e^{-q_{n+l}h'}.$$

By Lemma 5.20, there exists $P \in SL(2,\mathbb{R})$ that conjugates (5.30) to (5.7), and let $B = e^{Y_1}P$. Then we get the desired estimates. Since the estimates are similar as estimates in Proposition 5.5, we omit the details. □

5.4.4. *Proof of Proposition 5.8.* Let $\eta = \epsilon^{1/4}$. Since $2e^{-c_0 q_n^{\frac{1}{\mathcal{A}}} h'^2} < 10^{-8}$, then by Lemma 5.9, there exists $Y_1 \in \mathcal{B}_{h'}$ and $F^{(re)} \in \mathcal{B}_{h'}^{(re)}(\eta)$ such that (5.2) can be conjugate to (5.15) by the conjugation map e^{Y_1}, and $|Y_1|_{h'} \le \epsilon^{1/2}$, $|F^{(re)}|_{h'} \le 2\epsilon$.

On the other hand, note $\epsilon^{\frac{1}{12}} < \frac{1}{7q_n}$ by (5.4). Then by Corollary 5.11 and Remark 5.15, $F^{(re)}$ takes the form

$$\mathcal{T}_{\frac{q_{n+1}}{6}}F^{(re)}(\theta) = \sum_{\substack{k = l(p_n, -q_n) \\ |k| < \frac{q_{n+1}}{6}}} \hat{F}^{(re)}(k) e^{2\pi i\langle k, \theta\rangle}.$$

Now we consider the system

$$\begin{cases} \dot{x} = (A + \mathcal{T}_{\frac{q_{n+1}}{6}}F^{(re)}(\theta))x \\ \dot{\theta} = \omega \end{cases}, \tag{5.31}$$

with $|A+\mathcal{T}_{\frac{q_{n+1}}{6}}F^{(re)}|_{h'}\le|c^*|+|F^{(re)}|_{h'}\le 3\tilde{\epsilon}$ with $\tilde{\epsilon}=2e^{-c_0 q_n^{\frac{1}{A}}h'^2}$. Then by Lemma 5.17, there exist $D\in sl(2,\mathbb{R})$ and $L_2\in C^{\omega}_{h'}(\mathbb{T}^2,PSL(2,\mathbb{R}))$ that conjugates (5.31) to

$$\begin{cases}\dot{x}=Dx\\ \dot{\theta}=\omega\end{cases},$$

with $\|L_2\|_{h'}<e^{48\tilde{\epsilon}h' q_n q_{n+1}}$, and

$$\|D\|\le\frac{1}{q_{n+1}}e^{12\tilde{\epsilon}q_{n+1}}\le e^{\frac{q_{n+1}}{4}}\tilde{h}^4. \tag{5.32}$$

Then under the conjugation map L_2, the system (5.15) becomes

$$\begin{cases}\dot{x}=(D+L_2^{-1}\mathcal{R}_{\frac{q_{n+1}}{6}}F^{(re)}L_2)x\\ \dot{\theta}=\omega\end{cases}, \tag{5.33}$$

with estimate

$$\begin{aligned}\|L_2^{-1}\mathcal{R}_{\frac{q_{n+1}}{6}}F^{(re)}L_2\|_{\frac{h'}{6}} &\le e^{96\tilde{\epsilon}h' q_n q_{n+1}}\sum_{|k|\ge\frac{q_{n+1}}{6}}|\hat{F}^{(re)}(k)|e^{2\pi|k|\frac{h'}{6}}\\ &\le e^{96\tilde{\epsilon}h' q_n q_{n+1}}e^{-2\pi\frac{q_{n+1}}{6}\frac{5h'}{6}}|F^{(re)}|_{h'}<\epsilon e^{-q_{n+1}\tilde{h}}.\end{aligned}$$

Then by (5.32), one can apply Lemma 5.20, and there exists $P\in SL(2,\mathbb{R})$ that conjugates (5.33) to (5.8). Let $B=e^{Y_1}L_2P$. Then we get the desired estimates. Since the estimates are similar as estimates in Proposition 5.6, we omit the details. □

6. Measure complexity estimation for elliptic case: proof of Proposition 3.5

In this section, we will prove the measure complexity of a cocycle is sub-polynomial if the constant matrices in the conjugated cocycles are elliptic, and the estimates of the conjugations W_j and perturbations G_j satisfy $\|G_j\|_{C^0}^{\eta}\|W_j\|_{C^1}\to 0$ as $j\to\infty$ for any $\eta>0$.

The following lemma gives the difference of two distinct points under the projective action of a cocycle.

Lemma 6.1. *Let $a\in\mathbb{T}^1$, $B\in C^1(\mathbb{T}^1,PSL(2,\mathbb{R}))$. Then for any $(\theta,\varphi),(\tilde{\theta},\tilde{\varphi})\in\mathbb{T}^1\times\mathbb{RP}^1$, we have*

$$d(T_{(a,B)}(\theta,\varphi),T_{(a,B)}(\tilde{\theta},\tilde{\varphi}))\le C^*\|B\|_{C^1}^4 d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})),$$

where C^* is an absolute constant, and $d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})):=\max\{\|\theta-\tilde{\theta}\|,\|\varphi-\tilde{\varphi}\|\}$. In particular, if $B=R_\varrho$, then

$$d(T_{(a,B)}(\theta,\varphi),T_{(a,B)}(\tilde{\theta},\tilde{\varphi}))=d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})).$$

Proof. Suppose $B(\theta)=\begin{pmatrix}b_{11}(\theta)&b_{12}(\theta)\\ b_{21}(\theta)&b_{22}(\theta)\end{pmatrix}$. Denote $T_{(a,B)}(\theta,\varphi)=(\theta+a,f(\theta,\varphi))$, where $\tan(\hat{f}(\theta,\varphi))=\frac{b_{21}(\theta)+b_{22}(\theta)\tan\hat{\varphi}}{b_{11}(\theta)+b_{12}(\theta)\tan\hat{\varphi}}=\frac{b_{21}(\theta)\cot\hat{\varphi}+b_{22}(\theta)}{b_{11}(\theta)\cot\hat{\varphi}+b_{12}(\theta)}$, with $\hat{f}:=2\pi\gamma(f)$ and $\hat{\varphi}:=2\pi\gamma(\varphi)$, where $\gamma:\mathbb{RP}^1\to(-\frac{1}{4},\frac{1}{4}]$ is the lift of the identity map on $\mathbb{RP}^1$. Then, we get

$$\|f(\theta,\varphi)-f(\theta,\tilde{\varphi})\|\le|\frac{1}{2\pi}(\hat{f}(\theta,\varphi)-\hat{f}(\theta,\tilde{\varphi}))|\le\frac{1}{2\pi}\left\|\frac{\partial\hat{f}}{\partial\varphi}\right\|_{C^0}\|\varphi-\tilde{\varphi}\|.$$

Now we give the estimation of $\|\frac{\partial \hat{f}}{\partial \varphi}\|_{C^0}$: By a simple computation,

$$\frac{\partial \hat{f}}{\partial \varphi}(\theta,\varphi) = \frac{2\pi(1+(\tan\hat{\varphi})^2)}{(b_{11}(\theta)+b_{12}(\theta)\tan\hat{\varphi})^2+(b_{21}(\theta)+b_{22}(\theta)\tan\hat{\varphi})^2} \tag{6.1}$$

$$= \frac{2\pi(1+(\cot\hat{\varphi})^2)}{(b_{11}(\theta)\cot\hat{\varphi}+b_{12}(\theta))^2+(b_{21}(\theta)\cot\hat{\varphi}+b_{22}(\theta))^2}. \tag{6.2}$$

Moreover, for any $y\in\mathbb{R}$, we have

$$\begin{aligned}(b_{11}(\theta)+b_{12}(\theta)y)^2+(b_{21}(\theta)+b_{22}(\theta)y)^2 &\geq \frac{1}{b_{12}^2(\theta)+b_{22}^2(\theta)} \geq \frac{1}{2\|B\|_{C^1}^2},\\ (b_{11}(\theta)y+b_{12}(\theta))^2+(b_{21}(\theta)y+b_{22}(\theta))^2 &\geq \frac{1}{b_{11}^2(\theta)+b_{21}^2(\theta)} \geq \frac{1}{2\|B\|_{C^1}^2}.\end{aligned}$$

Therefore, if $|\tan\hat{\varphi}|\leq 1$, then by (6.1), we have $\left|\frac{\partial \hat{f}}{\partial \varphi}\right| \leq 8\pi\|B\|_{C^1}^2$. Otherwise, by (6.2), we also have $\left|\frac{\partial \hat{f}}{\partial \varphi}\right| \leq 8\pi\|B\|_{C^1}^2$. Hence, for any $\varphi,\tilde{\varphi}\in\mathbb{RP}^1$,

$$\|f(\theta,\varphi)-f(\theta,\tilde{\varphi})\| \leq 4\|B\|_{C^1}^2\|\varphi-\tilde{\varphi}\|.$$

Similarly, we get

$$\|f(\theta,\varphi)-f(\tilde{\theta},\varphi)\| \leq |\frac{1}{2\pi}(\hat{f}(\theta,\varphi)-\hat{f}(\tilde{\theta},\varphi))| \leq \frac{1}{2\pi}\left\|\frac{\partial \hat{f}}{\partial \theta}\right\|_{C^0}\|\theta-\tilde{\theta}\|,$$

and

$$\left|\frac{\partial \hat{f}}{\partial \theta}(\theta,\varphi)\right| \leq 8\|B\|_{C^1}^2\cdot\frac{1}{2\pi}\left|\frac{\partial \hat{f}}{\partial \varphi}(\theta,\varphi)\right| \leq 32\|B\|_{C^1}^4, \quad \forall(\theta,\varphi)\in\mathbb{T}^1\times\mathbb{RP}^1.$$

Therefore,

$$\begin{aligned}d(T_{(a,B)}(\theta,\varphi),T_{(a,B)}(\tilde{\theta},\tilde{\varphi})) &= \max\{\|f(\theta,\varphi)-f(\tilde{\theta},\tilde{\varphi})\|,\|\theta-\tilde{\theta}\|\}\\ &\leq \max\{\|f(\theta,\varphi)-f(\theta,\tilde{\varphi})\|+\|f(\theta,\tilde{\varphi})-f(\tilde{\theta},\tilde{\varphi})\|,\|\theta-\tilde{\theta}\|\}\\ &\leq C^*\|B\|_{C^1}^4 d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})).\end{aligned}$$

If $B=R_\varrho$, then we have $T_{(a,B)}(\theta,\varphi)=(\theta+a,\varphi+\varrho)$ for any $(\theta,\varphi)\in\mathbb{T}^1\times\mathbb{RP}^1$, which implies

$$d(T_{(a,B)}(\theta,\varphi),T_{(a,B)}(\tilde{\theta},\tilde{\varphi})) = d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})).$$

□

Lemma 6.2. *Let $a\in\mathbb{T}^1$. For any $A\in C^1(\mathbb{T}^1,PSL(2,\mathbb{R}))$, $F\in C^0(\mathbb{T}^1,SL(2,\mathbb{R}))$, if $\|F-I\|_{C^0}=\varepsilon\leq\frac{1}{6}$, then*

$$d(T_{(a,AF)}(\theta,\varphi),T_{(a,A)}(\theta,\varphi)) \leq 2C^*\|A\|_{C^1}^4\|F-I\|_{C^0}, \quad \forall\ (\theta,\varphi)\in\mathbb{T}^1\times\mathbb{RP}^1,$$

where C^* is the global constant in Lemma 6.1.

Proof. Since $(a,AF)=(a,A)\circ(0,F)$, we have $T_{(a,AF)}=T_{(a,A)}\circ T_{(0,F)}$. Then by Lemma 6.1, for any $(\theta,\varphi)\in\mathbb{T}^1\times\mathbb{RP}^1$,

$$\begin{aligned}d(T_{(a,AF)}(\theta,\varphi),T_{(a,A)}(\theta,\varphi)) &= d(T_{(a,A)}(T_{(0,F)}(\theta,\varphi)),T_{(a,A)}(\theta,\varphi))\\ &\leq C^*\|A\|_{C^1}^4 d(T_{(0,F)}(\theta,\varphi),(\theta,\varphi)).\end{aligned}$$

Denote $T_{(0,F)}(\theta,\varphi) =: (\theta,\varphi_1)$ and $F(\theta) = \begin{pmatrix} 1+f_1(\theta) & f_2(\theta) \\ f_3(\theta) & 1+f_4(\theta) \end{pmatrix}$. Let $\hat{\varphi} := 2\pi\gamma(\varphi)$ and $\hat{\varphi}_1 := 2\pi\gamma(\varphi_1)$, where $\gamma : \mathbb{RP}^1 \to (-\frac{1}{4}, \frac{1}{4}]$ is the lift of the identity map on $\mathbb{RP}^1$. If $|\sin\hat{\varphi}| \le \frac{\sqrt{2}}{2}$, then $|\cos\hat{\varphi}| \ge \frac{\sqrt{2}}{2}$ and

$$\begin{aligned}
\|\varphi - \varphi_1\| = \frac{1}{2\pi}|\hat{\varphi} - \hat{\varphi}_1| &\le \frac{1}{2\pi}|\tan\hat{\varphi} - \tan\hat{\varphi}_1| \\
&= \frac{1}{2\pi}\left|\frac{f_3\cos\hat{\varphi} + (1+f_4)\sin\hat{\varphi}}{(1+f_1)\cos\hat{\varphi} + f_2\sin\hat{\varphi}} - \frac{\sin\hat{\varphi}}{\cos\hat{\varphi}}\right| \\
&= \frac{1}{2\pi}\left|\frac{(f_4 - f_1)\cos\hat{\varphi}\sin\hat{\varphi} - f_2\sin^2\hat{\varphi} + f_3\cos^2\hat{\varphi}}{(1+f_1)\cos^2\hat{\varphi} + f_2\sin\hat{\varphi}\cos\hat{\varphi}}\right| \\
&\le \frac{12\varepsilon}{2\pi} < 2\varepsilon,
\end{aligned}$$

which implies that

$$d(T_{(0,F)}(\theta,\varphi),(\theta,\varphi)) = \|\varphi - \varphi_1\| < 2\varepsilon.$$

If $|\sin\hat{\varphi}| > \frac{\sqrt{2}}{2}$, then similarly,

$$\begin{aligned}
\|\varphi - \varphi_1\| = \frac{1}{2\pi}|\hat{\varphi} - \hat{\varphi}_1| &\le \frac{1}{2\pi}|\cot\hat{\varphi} - \cot\hat{\varphi}_1| \\
&= \frac{1}{2\pi}\left|\frac{(f_1 - f_4)\cos\hat{\varphi}\sin\hat{\varphi} + f_2\sin^2\hat{\varphi} - f_3\cos^2\hat{\varphi}}{(1+f_4)\sin^2\hat{\varphi} + f_3\sin\hat{\varphi}\cos\hat{\varphi}}\right| \\
&\le \frac{12\varepsilon}{2\pi} < 2\varepsilon,
\end{aligned}$$

which also implies that

$$d(T_{(0,F)}(\theta,\varphi),(\theta,\varphi)) = \|\varphi - \varphi_1\| < 2\varepsilon.$$

Therefore,

$$d(T_{(a,AF)}(\theta,\varphi), T_{(a,A)}(\theta,\varphi)) \le 2C^*\|A\|_{C^1}^4\|F - I\|_{C^0}.$$

$\square$

Lemma 6.3 ([5]). *For matrices $\{M_l\}_l \subseteq SL(2,\mathbb{R}), \{I+\xi_l\}_l \subseteq C(\mathbb{T}^1, SL(2,\mathbb{R}))$, we have that*

$$M_l(I+\xi_l)\dots M_1(I+\xi_1) = M^{(l)}(I+\xi^{(l)}),$$

with $M^{(l)} = M_l \dots M_1$ and $\xi^{(l)}$ satisfying

$$\|\xi^{(l)}\| \le e^{\sum_{k=1}^{l}\|M^{(k)}\|^2\|\xi_k\|} - 1.$$

By Lemma 6.3 and Lemma 6.2, the following holds:

Lemma 6.4. *Let $A = R_\varrho$ with $\varrho \in \mathbb{T}^1$ or $A = \begin{pmatrix} 1 & c_* \\ 0 & 1 \end{pmatrix}$ with $|c_*| \le 1$ and $\tilde{A}(\theta) = A(I + F(\theta)) \in C(\mathbb{T}^1, SL(2,\mathbb{R}))$. Then for $0 \le m \le \frac{1}{4\|F\|_{C^0}^{1/3}}$, we have*

$$d(T_{(\alpha,A)^m}(\theta,\varphi), T_{(\alpha,\tilde{A})^m}(\theta,\varphi)) \le 16C^*m^3\|A^m\|^4\|F\|_{C^0},\ \forall\ (\theta,\varphi) \in \mathbb{T}^1 \times \mathbb{RP}^1.$$

Proof. By Lemma 6.3, we get $\tilde{A}_m(\theta) = A^m(I+F^{(m)}(\theta))$ with $\|F^{(m)}(\theta)\| \le e^{\sum_{k=1}^m \|A^k\|^2\|F\|_{C^0}} - 1 < 8m^3\|F\|_{C^0} < 1/6$ for any $\theta \in \mathbb{T}^1$. Then by Lemma 6.2, we obtain that

$$\begin{aligned} d(T_{(\alpha,A)^m}(\theta,\varphi), T_{(\alpha,\tilde{A})^m}(\theta,\varphi)) &= d(T_{(m\alpha,A^m)}(\theta,\varphi), T_{(m\alpha,A^m(I+F^{(m)}))}(\theta,\varphi)) \\ &\le 2C^*\|A^m\|^4\|F^{(m)}\|_{C^0} \le 16C^*m^3\|A^m\|^4\|F\|_{C^0}. \end{aligned}$$

□

Proof of Proposition 3.5. Let $\tau, \epsilon > 0$. For any $(\theta,\varphi), (\tilde{\theta},\tilde{\varphi}) \in \mathbb{T}^1 \times \mathbb{RP}^1$, we denote $\varphi_n = \pi_2 \circ T_{(\alpha,A)^n}(\theta,\varphi)$, $\tilde{\varphi}_n = \pi_2 \circ T_{(\alpha,A)^n}(\tilde{\theta},\tilde{\varphi})$. For any $\varepsilon > 0$, there exists $j_* \in \mathbb{N}$ such that $\forall\ j \ge j_*$,

$$(6C^*)^{4+\tau}\|W_j\|_{C^1}^{16}\|G_j\|_{C^0}^{\frac{\tau}{4}} < \epsilon^2\varepsilon, \quad \text{and} \quad \|W_j\|_{C^1}^4\|G_j\|_{C^0}^{\frac{1}{4}} < \epsilon,$$

where $C^* \ge 1$ is the global constant in Lemma 6.1. Then for $j \ge j_*$, by Lemma 6.1, 6.2, and 6.4, we have

$$\begin{aligned} \|\varphi_n - \tilde{\varphi}_n\| &= \|\pi_2 \circ T_{(0,W_j)\circ(\alpha,R_{\varrho_j}(I+G_j))^n\circ(0,W_j)^{-1}}(\theta,\varphi) \\ &\qquad -\pi_2 \circ T_{(0,W_j)\circ(\alpha,R_{\varrho_j}(I+G_j))^n\circ(0,W_j)^{-1}}(\tilde{\theta},\tilde{\varphi})\| \\ &\le C^*\|W_j\|_{C^1}^4 d\Big(\ T_{(\alpha,R_{\varrho_j}(I+G_j))^n\circ(0,W_j)^{-1}}(\theta,\varphi), T_{(\alpha,R_{\varrho_j}(I+G_j))^n\circ(0,W_j)^{-1}}(\tilde{\theta},\tilde{\varphi})\ \Big) \\ &\le C^*\|W_j\|_{C^1}^4 \Big\{ d(\ T_{(\alpha,R_{\varrho_j}(I+G_j))^n}(T_{(0,W_j)^{-1}}(\theta,\varphi)), T_{(\alpha,R_{\varrho_j})^n}(T_{(0,W_j)^{-1}}(\theta,\varphi))\ ) \\ &\qquad + d(\ T_{(\alpha,R_{\varrho_j})^n\circ(0,W_j)^{-1}}(\theta,\varphi), T_{(\alpha,R_{\varrho_j})^n\circ(0,W_j)^{-1}}(\tilde{\theta},\tilde{\varphi})\ ) \\ &\qquad + d(\ T_{(\alpha,R_{\varrho_j}(I+G_j))^n}(T_{(0,W_j)^{-1}}(\tilde{\theta},\tilde{\varphi})), T_{(\alpha,R_{\varrho_j})^n}(T_{(0,W_j)^{-1}}(\tilde{\theta},\tilde{\varphi}))\ ) \Big\} \\ &\le C^*\|W_j\|_{C^1}^4 \Big(64C^*n^3\|G_j\|_{C^0} + C^*\|W_j\|_{C^1}^4 d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})) \Big) \end{aligned}$$

for $0 \le n \le \frac{1}{4\|G_j\|_{C^0}^{1/3}}$. Now let $M_j = \left\lceil \frac{1}{6C^*\|G_j\|_{C^0}^{1/4}} \right\rceil$. Then for $d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})) < \frac{\epsilon}{2C^{*2}\|W_j\|_{C^1}^8}$, we have

$$\bar{d}_{M_j}((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})) < \epsilon.$$

If we let $L_j := \left\{ (k_1 \cdot \frac{\epsilon}{2C^{*2}\|W_j\|_{C^1}^8}, k_2 \cdot \frac{\epsilon}{2C^{*2}\|W_j\|_{C^1}^8}) \ : \ 0 \le k_1, k_2 \le \left[\frac{2C^{*2}\|W_j\|_{C^1}^8}{\epsilon} \right] + 1 \right\}$, then for any $(\theta,\varphi) \in \mathbb{T}^1 \times \mathbb{RP}^1$, there exists $(\tilde{\theta},\tilde{\varphi}) \in L_j$ such that $d((\theta,\varphi),(\tilde{\theta},\tilde{\varphi})) < \frac{\epsilon}{2C^{*2}\|W_j\|_{C^1}^8}$, which implies that $S_{M_j}(d,\rho,\epsilon) < \frac{5C^{*4}\|W_j\|_{C^1}^{16}}{\epsilon^2}$ for any $\rho \in \mathcal{M}(\mathbb{T}^1 \times \mathbb{RP}^1)$. Hence, we have

$$\frac{S_{M_j}(d,\rho,\epsilon)}{M_j^\tau} < \frac{6^{\tau+1}C^{*(4+\tau)}}{\epsilon^2}\|G_j\|_{C^0}^{\frac{\tau}{4}}\|W_j\|_{C^1}^{16} < \varepsilon, \quad \text{as } j \ge j_*.$$

Therefore, for any $\rho \in \mathcal{M}(\mathbb{T}^1 \times \mathbb{RP}^1)$ and $\tau > 0$, we have

$$\liminf_{n\to\infty} \frac{S_n(d,\rho,\epsilon)}{n^\tau} = 0, \quad \forall \epsilon > 0,$$

which means that the complexity of $(\mathbb{T}^1 \times \mathbb{RP}^1, T_{(\alpha,A)}, \rho)$ is sub-polynomial. □

7. Möbius disjointness for the parabolic case: proof of Proposition 3.6

The main ideas of the proof are developed from [43]. In [43], within a reasonably long interval, any orbit of T will come back to a neighborhood of the starting point in an almost periodic manner. However, in our case, the conjugated cocycle might be the perturbation of a parabolic matrix, and consequently those points with the second variable φ not close to 0 will leave the neighborhood of the starting point without returning in the future, which is the main difference compared to [43]. The key observation here is that in this situation the orbit of the second variable will approach 0. Inspired by this, we approximate the orbit of (θ, φ) by the orbit of $(\theta, 0)$ by a periodic sequence with period q_{k_j}. Then we decompose the periodic sequence into short average of Dirichlet characters[2], reducing the problem to control the average of multiplicative function on a typical interval, which is the content of the following lemma.

Lemma 7.1 ([43])**.** *For all $L, Q, M \in \mathbb{N}$ and any periodic function $D : \mathbb{N} \to \mathbb{C}$ of period Q with $|D| \leq 1$,*

$$|\mathbb{E}_{L\leq n<L+MQ}\mu(n)D(n)|^2 \leq Q\mathbb{E}_{d|Q,\chi \bmod^* \frac{Q}{d}} \left|\mathbb{E}_{\frac{L}{d}\leq r<\frac{L}{d}+M\frac{Q}{d}}\mu(r)\chi(r)\right|^2,$$

where the first average on the right hand side is taken over all pairs (d, χ) such that $d|Q$ and χ is a Dirichlet character of conductor $\frac{Q}{d}$.

So as to control $\left|\mathbb{E}_{\frac{L}{d}\leq r<\frac{L}{d}+M\frac{Q}{d}}\mu(r)\chi(r)\right|$, we will use a recent result in [37] that allows to bound averages of a non-pretentious multiplicative function in random short intervals. We first recall the definition of pretentiousness.

For multiplicative functions $\nu, \nu' : \mathbb{N} \to \mathbb{C}$ with $|\nu|, |\nu'| \leq 1$, define $\mathbb{D}(\nu, \nu', X) = \left(\sum_{p\leq X} \frac{1-\Re(\nu(p)\overline{\nu'(p)})}{p}\right)^{\frac{1}{2}}$, and the function below measures how closely ν pretends to be n^{it}:

$$M(\nu, X) = \inf_{|t|\leq X} \mathbb{D}(\nu, n^{it}, X)^2.$$

By the same discussions in [43], we know that

$$\lim_{X\to\infty} M(\mu\chi, X) = \infty.$$

We will apply the following proposition to $\nu(n) = \mu(n)\chi(n)$.

Proposition 7.2 ([37])**.** *Let ν be a multiplicative function with $|\nu| \leq 1$ and $X \geq l \geq 10$. Then*

$$\mathbb{E}_{X\leq L<2X}\left|\mathbb{E}_{L\leq n<L+l}\ \nu(n)\right|^2 \lesssim e^{-M(\nu,X)}M(\nu, X) + (\log X)^{-\frac{1}{50}} + \left(\frac{\log\log l}{\log l}\right)^2.$$

Proof of Proposition 3.6. Since trigonometric polynomials are dense in $C^0(\mathbb{T}^1 \times \mathbb{RP}^1)$, it suffices to prove the orthogonality for every $f(\theta, \varphi) = e^{2\pi i(\iota_1\theta+2\iota_2\varphi)}, \iota_1, \iota_2 \in \mathbb{Z}$. Fix a function $f(\theta, \varphi) = e^{2\pi i(\iota_1\theta+2\iota_2\varphi)}$ and $0 < \eta \ll 1$. Suppose k_j is sufficiently large and $q_{k_j+1} \geq e^{\beta q_{k_j}/2}$. Let $\tilde{C}$ be the constant in the proposition. Then for j large

[2]We say a function $\chi : \mathbb{Z} \to \mathbb{C}$ is a Dirichlet character modulus k if it has the following properties: (1) $\chi(n) = \chi(n + k)$ for all n; (2) If gcd(n,k)>1 then $\chi(n) = 0$, if gcd(n,k)=1 then $\chi(n) \neq 0$; (3) $\chi(mn) = \chi(m)\chi(n)$ for all integers m and n.

enough, we have $6\tilde{C}^2e^{-(\frac{\beta}{4}-2\tilde{C}\eta_j-\tau_j)q_{k_j}}<1$, $\frac{16\tilde{C}\eta_j}{\tau_j}<\xi$, and $4\tilde{C}q_{k_j}<e^{\frac{1}{7}\xi q_{k_j}^{1/2}}$. For any $(\theta,\varphi)\in\mathbb{T}^1\times\mathbb{RP}^1$, denote

$$x_0=(\theta,\varphi),\quad \tilde{x}_0=T_{(0,W_j)^{-1}}x_0=:(\theta,\tilde{\varphi}),\quad T=T_{(\alpha,A)},\quad \tilde{T}=T_{(\alpha,A_j(I+G_j))},\quad \tilde{x}_n=\tilde{T}^n\tilde{x}_0.$$

Let $N_0<\frac{N}{2}$ be an integer. Then,

$$\begin{aligned}
&\mathbb{E}_{n<N}\mu(n)f(T^nx_0)=\mathbb{E}_{n=N_0}^{N-1}\mu(n)f(T^nx_0)+O(\frac{N_0}{N})\\
=\;&\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{n=0}^{M_jq_{k_j}-1}\mu(L+n)f(T^{L+n}x_0)+O(\frac{M_jq_{k_j}}{N})+O(\frac{N_0}{N})\\
=\;&\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{n=0}^{q_{k_j}-1}\mathbb{E}_{m=0}^{M_j-1}\mu(L+n+mq_{k_j})f(T^{L+n+mq_{k_j}}x_0)+O(\frac{M_jq_{k_j}}{N})+O(\frac{N_0}{N}).
\end{aligned}$$

Now we approximate $f(T^{L+n+mq_{k_j}}x_0)$ by some periodic sequence. We first approximate it by $T_{(0,W_j)}(T_{(0,A_j)^{mq_{k_j}}}\tilde{x}_{L+n})$. Then use the nature of parabolic matrix A_j to count the number of m where $0\le m\le M_j-1$ that the second variable of $T_{(0,A_j)^{mq_{k_j}}}\tilde{x}_{L+n}$ is away from 0, and show that it is quite small comparing to M_j.

Lemma 7.3. *For any $L,m,n\in\mathbb{N}$ with $m<M_j\le e^{\frac{1}{7}-\frac{\xi}{2}\tau_jq_{k_j}}$, we have*

$$\begin{aligned}
&d\left(T^{L+n+mq_{k_j}}x_0,\ T_{(0,W_j)}(T_{(0,A_j)^{mq_{k_j}}}\tilde{x}_{L+n})\right)\\
&\lesssim M_j^3q_{k_j}^3(1+M_jq_{k_j}|c_j|)^4\|W_j\|_{C^1}^4\|G_j\|_{C^0}+\tfrac{M_j\|W_j\|_{C^1}^6}{q_{k_j+1}},
\end{aligned}$$

where $A_j=\begin{pmatrix}1&c_j\\0&1\end{pmatrix}$ with j large enough.

Proof. By the fact that $(0,W_j)^{-1}\circ(\alpha,A)\circ(0,W_j)=(\alpha,A_j(I+G_j))$, we have $T^{L+n+mq_{k_j}}x_0=T_{(0,W_j)}(T_{(\alpha,A_j(I+G_j))^{mq_{k_j}}}(\tilde{x}_{L+n}))$. Then,

$$\begin{aligned}
&d\left(T^{L+n+mq_{k_j}}x_0,\ T_{(0,W_j)}(T_{(0,A_j)^{mq_{k_j}}}\tilde{x}_{L+n})\right)\\
\le\;&d\left(T_{(0,W_j)}(T_{(\alpha,A_j(I+G_j))^{mq_{k_j}}}\tilde{x}_{L+n}),\ T_{(0,W_j(\cdot-mq_{k_j}\alpha))}(T_{(\alpha,A_j)^{mq_{k_j}}}\tilde{x}_{L+n})\right)\\
&+d\left(T_{(0,W_j(\cdot-mq_{k_j}\alpha))}(T_{(\alpha,A_j)^{mq_{k_j}}}\tilde{x}_{L+n}),\ T_{(0,W_j)}(T_{(0,A_j)^{mq_{k_j}}}\tilde{x}_{L+n})\right).
\end{aligned}$$

Since A_j is a constant matrix, then $(\alpha,A_j)^{mq_{k_j}}=(mq_{k_j}\alpha,A_j^{mq_{k_j}})$, $(0,A_j)^{mq_{k_j}}=(0,A_j^{mq_{k_j}})$, and thus $(0,W_j(\cdot-mq_{k_j}\alpha))\circ(\alpha,A_j)^{mq_{k_j}}=(mq_{k_j}\alpha,W_j(\cdot)A_j^{mq_{k_j}})$, $(0,W_j)\circ(0,A_j)^{mq_{k_j}}=(0,W_j(\cdot)A_j^{mq_{k_j}})$. Therefore,

$$d\left(T_{(0,W_j(\cdot-mq_{k_j}\alpha))}(T_{(\alpha,A_j)^{mq_{k_j}}}\tilde{x}_{L+n}),\ T_{(0,W_j)}(T_{(0,A_j)^{mq_{k_j}}}\tilde{x}_{L+n})\right)\le\|mq_{k_j}\alpha\|_{\mathbb{T}}\le\frac{m}{q_{k_j+1}}.$$

Denoting $\Delta W_{j,m}(\cdot) = W_j(\cdot)^{-1} W_j(\cdot - mq_{k_j}\alpha)$, then for $0 \leq m < M_j$ with j sufficiently large,

$$\begin{aligned} & \|\Delta W_{j,m} - I\|_{C^0} \\ \leq\ & \|W_j\|_{C^0} \cdot \|W_j(\cdot - mq_{k_j}\alpha) - W_j(\cdot)\|_{C^0} \\ \leq\ & \|W_j\|_{C^1} \cdot \|W_j\|_{C^1} \cdot \|mq_{k_j}\alpha\|_{\mathbb{T}} \leq \frac{m\|W_j\|_{C^1}^2}{q_{k_j+1}} \\ \leq\ & \tilde{C}^2 e^{-(\frac{\beta}{2} - 2\tilde{C}\eta_j - \tau_j)q_{k_j}} < \frac{1}{6}. \end{aligned}$$

By Lemma 6.2,

$$\begin{aligned} & d\left(T_{(0,W_j)}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n}),\ T_{(0,W_j(\cdot)\Delta W_{j,m}(\cdot))}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n})\right) \\ \leq\ & 2C^* \|W_j\|_{C^1}^4 \|\Delta W_{j,m} - I\|_{C^0} \lesssim \frac{m\|W_j\|_{C^1}^6}{q_{k_j+1}}. \end{aligned}$$

Moreover, by Lemma 6.1 and 6.4, we have

$$\begin{aligned} & d\left(T_{(0,W_j)}(T_{(\alpha,A_j(I+G_j))}{}^{mq_{k_j}} \tilde{x}_{L+n}),\ T_{(0,W_j)}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n})\right) \\ \leq\ & C^* \|W_j\|_{C^1}^4 d\left(T_{(\alpha,A_j(I+G_j))}{}^{mq_{k_j}} \tilde{x}_{L+n},\ T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n}\right) \\ \leq\ & 16 C^{*2} (mq_{k_j})^3 \|W_j\|_{C^1}^4 \|A_j^{mq_{k_j}}\|^4 \|G_j\|_{C^0} \\ \lesssim\ & m^3 q_{k_j}^3 (1 + mq_{k_j}|c_j|)^4 \|W_j\|_{C^1}^4 \|G_j\|_{C^0}, \end{aligned}$$

since $4m\|G_j\|_{C^0}^{\frac{1}{3}} \leq 4e^{\frac{1}{7} - \frac{\xi}{2}\tau_j q_{k_j}} \tilde{C} e^{-\tau_j q_{k_j}} < 1$ for j large enough.

In conclusion, we get

$$\begin{aligned} & d\left(T^{L+n+mq_{k_j}} x_0,\ T_{(0,W_j)}(T_{(0,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n})\right) \\ \leq\ & d\left(T_{(0,W_j)}(T_{(\alpha,A_j(I+G_j))}{}^{mq_{k_j}} \tilde{x}_{L+n}),\ T_{(0,W_j)}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n})\right) \\ & + d\left(T_{(0,W_j)}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n}),\ T_{(0,W_j(\cdot)\Delta W_{j,m}(\cdot))}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n})\right) \\ & + d\left(T_{(0,W_j(\cdot - mq_{k_j}\alpha))}(T_{(\alpha,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n}),\ T_{(0,W_j)}(T_{(0,A_j)}{}^{mq_{k_j}} \tilde{x}_{L+n})\right) \\ \lesssim\ & M_j^3 q_{k_j}^3 (1 + M_j q_{k_j}|c_j|)^4 \|W_j\|_{C^1}^4 \|G_j\|_{C^0} + \frac{M_j \|W_j\|_{C^1}^6}{q_{k_j+1}}. \end{aligned}$$

□

Moreover, for any $0 < \tilde{\eta} \ll 1, (\bar{\theta}, \bar{\varphi}) \in \mathbb{T}^1 \times \mathbb{RP}^1$, we denote

$$\bar{\varphi}^{(m)} := \pi_2 \circ T_{(0,A_j)}{}^{mq_{k_j}} (\bar{\theta}, \bar{\varphi}) = \pi_2 \circ T_{(0,A_j^{q_{k_j}})^m} (\bar{\theta}, \bar{\varphi})$$

and $I_0(\bar{\theta}, \bar{\varphi}, \tilde{\eta}) := \{m \in \mathbb{N} : |\bar{\varphi}^{(m)}| > \tilde{\eta}\}$. Let $v_* = (a_*, 1)^T \in \mathbb{R}^2$ be any vector and its angle is φ_*. Then the length for a_* such that $|\tan \varphi_*| > \tilde{\eta}$ is no more than $\frac{2}{\tilde{\eta}}$. Moreover, we have $A_j^{q_{k_j}} v_* = (a_* + q_{k_j} c_j, 1)^T$, implying that the effect of $A_j^{q_{k_j}}$ on v_* is to add $q_{k_j} c_j$

on its first variable. Since $I_0(\bar\theta,\bar\varphi,\tilde\eta)\subseteq\{m\in\mathbb{N}:|\tan(\bar\varphi^{(m)})|>\tilde\eta\}$, then we obtain that

$$\#I_0(\bar\theta,\bar\varphi,\tilde\eta)\le\frac{2}{\tilde\eta q_{k_j}|c_j|}+1.$$

Therefore, we have

$$\begin{aligned}
&\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{n=0}^{q_{k_j}-1}\mathbb{E}_{m=0}^{M_j-1}\mu(L+n+mq_{k_j})\left(f\circ T_{(0,W_j)}({T_{(0,A_j)}}^{mq_{k_j}}\tilde x_{L+n})-f\circ T_{(0,W_j)}(\theta_{L+n},0)\right)\\
=\;&\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{n=0}^{q_{k_j}-1}\frac{1}{M_j}\left(\sum_{\substack{0\le m<M_j\\ m\in I_0(\theta_{L+n},\tilde\varphi_{L+n},\tilde\eta)}}+\sum_{\substack{0\le m<M_j\\ m\notin I_0(\theta_{L+n},\tilde\varphi_{L+n},\tilde\eta)}}\right)\\
&\mu(L+n+mq_{k_j})\left(f\circ T_{(0,W_j)}({T_{(0,A_j)}}^{mq_{k_j}}\tilde x_{L+n})-f\circ T_{(0,W_j)}(\theta_{L+n},0)\right).
\end{aligned}$$

Since $\#I_0(\bar\theta,\bar\varphi,\tilde\eta)\le\frac{2}{\tilde\eta q_{k_j}|c_j|}+1$ for any $0<\tilde\eta\ll1,(\bar\theta,\bar\varphi)\in\mathbb{T}^1\times\mathbb{RP}^1$, then

$$\begin{aligned}
&\frac{1}{M_j}\sum_{\substack{0\le m<M_j\\ m\in I_0(\theta_{L+n},\tilde\varphi_{L+n},\tilde\eta)}}\mu(L+n+mq_{k_j})\left(f\circ T_{(0,W_j)}({T_{(0,A_j)}}^{mq_{k_j}}\tilde x_{L+n})-f\circ T_{(0,W_j)}(\theta_{L+n},0)\right)\\
\le\;&\frac{2}{M_j}\left(\frac{2}{\tilde\eta q_{k_j}|c_j|}+1\right)\lesssim\frac{1}{M_j\tilde\eta q_{k_j}|c_j|}.
\end{aligned}$$

Moreover, for those m such that $|\tilde\varphi_{L+n}^{(m)}|\le\tilde\eta$, we have

$$\begin{aligned}
&|f\circ T_{(0,W_j)}({T_{(0,A_j)}}^{mq_{k_j}}\tilde x_{L+n})-f\circ T_{(0,W_j)}(\theta_{L+n})|\\
\le\;&\|f\|_{C^1}d\left(T_{(0,W_j)}({T_{(0,A_j)}}^{mq_{k_j}}\tilde x_{L+n}),T_{(0,W_j)}(\theta_{L+n},0)\right)\\
\overset{Lemma\ 6.1}{\le}\;&C_*\|f\|_{C^1}\|W_j\|_{C^1}^4d\left((\theta_{L+n},\tilde\varphi_{L+n}^{(m)}),(\theta_{L+n},0)\right)\\
\lesssim\;&\|W_j\|_{C^1}^4\tilde\eta,
\end{aligned}$$

and thus

$$\begin{aligned}
&\frac{1}{M_j}\sum_{\substack{0\le m<M_j\\ m\notin I_0(\theta_{L+n},\tilde\varphi_{L+n},\tilde\eta)}}\mu(L+n+mq_{k_j})\left(f\circ T_{(0,W_j)}({T_{(0,A_j)}}^{mq_{k_j}}\tilde x_{L+n})-f\circ T_{(0,W_j)}(\theta_{L+n},0)\right)\\
\lesssim\;&\|W_j\|_{C^1}^4\tilde\eta.
\end{aligned}$$

In conclusion, we get

$$\begin{aligned}
&\mathbb{E}_{n<N}\mu(n)f(T^nx_0)\\
=\;&\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{n=0}^{q_{k_j}-1}\mathbb{E}_{m=0}^{M_j-1}\mu(L+n+mq_{k_j})f\circ T_{(0,W_j)}(\theta_{L+n},0)\\
+\;&O\left(\|W_j\|_{C^1}^4\tilde\eta+\frac{1}{M_j\tilde\eta q_{k_j}|c_j|}+M_j^7q_{k_j}^7|c_j|^4\|W_j\|_{C^1}^4\|G_j\|_{C^0}+\frac{M_j\|W_j\|_{C^1}^6}{q_{k_j+1}}+\frac{M_jq_{k_j}}{N}+\frac{N_0}{N}\right).
\end{aligned}$$

Let $D_l:=f\circ T_{(0,W_j)}(\theta_l,0)$, and for each $L\in\mathbb{N}$ construct a function $D_L:\mathbb{N}\to\mathbb{C}$ by $D_L(n)=D_l$ where l is the unique integer in $[L,L+q_{k_j})$ such that $l=n(\mathrm{mod}\,q_{k_j})$. Then D_L is periodic with period q_{k_j} and $|D_L|=1$.

Now we let $\tilde{\eta} = e^{-4\tilde{C}q_{k_j}\eta_j}\eta \leq \eta$, $M_j = [e^{(\frac{1}{7}-\frac{\xi}{2})\tau_j q_{k_j}}]$ and $J = \lceil \log_2 \frac{1}{\eta} \rceil$. Choose $N_0 = \lfloor 2^{-J}N \rfloor$, and then $\frac{N_0}{N} \leq \eta$. We have for j sufficiently large,

$$
\begin{aligned}
&|\mathbb{E}_{n<N}\mu(n)f(T^n x_0)| \\
= \quad & |\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{n=L}^{L+M_j q_{k_j}-1}\mu(n)D_L(n)| + O(\|W_j\|_{C^1}^4\tilde{\eta} + \frac{N_0}{N}) \\
& +O(\frac{1}{M_j\tilde{\eta}q_{k_j}|c_j|} + M_j^7 q_{k_j}^7 |c_j|^4 \|W_j\|_{C^1}^4 \|G_j\|_{C^0} + \frac{M_j\|W_j\|_{C^1}^6}{q_{k_j+1}}) + O(\frac{M_j q_{k_j}}{N}) \\
\overset{Cauchy-Schwarz}{\lesssim} \quad & \left(\mathbb{E}_{L=N_0}^{N-1}\left|\mathbb{E}_{n=L}^{L+M_j q_{k_j}-1}\mu(n)D_L(n)\right|^2\right)^{\frac{1}{2}} + \eta + \frac{e^{-q_{k_j}(\frac{\xi}{2}\tau_j - 4\tilde{C}\eta_j)}}{\eta q_{k_j}} \\
& +q_{k_j}^7 e^{-q_{k_j}(\frac{7}{2}\xi\tau_j - 4\tilde{C}\eta_j)} + e^{-q_{k_j}(\frac{\beta}{2}-(\frac{1}{7}-\frac{\xi}{2})\tau_j - 6\tilde{C}\eta_j)} + \frac{q_{k_j}e^{(\frac{1}{7}-\frac{\xi}{2})\tau_j q_{k_j}}}{N} \\
\overset{Lemma\ 7.1}{\lesssim} \quad & \left(q_{k_j}\cdot\mathbb{E}_{L=N_0}^{N-1}\mathbb{E}_{d|q_{k_j},\chi \bmod^* \frac{q_{k_j}}{d}}\left|\mathbb{E}_{n=\frac{L}{d}}^{\frac{L}{d}+\frac{M_j q_{k_j}}{d}-1}\mu(n)\chi(n)\right|^2\right)^{\frac{1}{2}} \\
& +\eta + \frac{e^{-\frac{\xi}{4}\tau_j q_{k_j}}}{\eta} + e^{-2\xi\tau_j q_{k_j}} + e^{-\frac{\beta}{4}q_{k_j}} + \frac{e^{(\frac{1}{7}-\frac{5}{14}\xi)\tau_j q_{k_j}}}{N}.
\end{aligned}
$$

Cut $[N_0, N)$ into J dyadic interval: $[2^{-\iota}N, 2^{-\iota+1}N)$ for $\iota = 1, \ldots, J$. Let $\tilde{\rho}_\chi(X) := e^{-M(\mu\chi,X)}M(\mu\chi, X) + (\log X)^{-\frac{1}{50}}$, $\tilde{\rho}_{q_{k_j}}(X) := \max_{\chi \bmod q_{k_j}} \tilde{\rho}_\chi(X)$, and $\rho_{q_{k_j}}(X) := \sup_{X'\geq X}\tilde{\rho}_{q_{k_j}}(X)$. Then $\rho_{q_{k_j}}$ is a positive function independent of χ that decreases to 0 as $X \to \infty$. By Proposition 7.2, for each pair (d, χ) that $d|q_{k_j}$ and χ is a Dirichlet character of conductor $\frac{q_{k_j}}{d}$, and every $\iota \in [1, J]$,

$$
\begin{aligned}
& \mathbb{E}_{2^{-\iota}N\leq L<2^{-\iota+1}N}\left|\mathbb{E}_{\frac{L}{d}\leq r<\frac{L}{d}+\frac{M_j q_{k_j}}{d}}\mu(r)\chi(r)\right|^2 \\
\lesssim \quad & \rho_{q_{k_j}}\left(\frac{2^{-\iota}N}{d}\right) + \left(\frac{\log\log\frac{M_j q_{k_j}}{d}}{\log\frac{M_j q_{k_j}}{d}}\right)^2 \\
\lesssim \quad & \rho_{q_{k_j}}\left(\frac{\eta N}{2q_{k_j}}\right) + \left(\frac{\log\log M_j}{\log M_j}\right)^2 \\
\lesssim \quad & \rho_{q_{k_j}}\left(\frac{\eta N}{2q_{k_j}}\right) + \frac{\log^2(\tau_j q_{k_j})}{\tau_j^2 q_{k_j}^2}.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
& |\mathbb{E}_{n<N}\mu(n)f(T^n x_0)| \\
\lesssim \quad & \left(q_{k_j}\cdot\rho_{q_{k_j}}(\frac{\eta N}{2q_{k_j}}) + \tau_j^{-2}q_{k_j}^{-1}\log^2(\tau_j q_{k_j})\right)^{\frac{1}{2}} + \eta + \frac{e^{-\frac{\xi}{4}\tau_j q_{k_j}}}{\eta} + \frac{e^{(\frac{1}{7}-\frac{5}{14}\xi)\tau_j q_{k_j}}}{N}.
\end{aligned}
$$

Once η is fixed, then for sufficiently large q_{k_j} with $q_{k_j+1} \geq e^{\beta q_{k_j}/2}$,

$$\tau_j^{-2}q_{k_j}^{-1}\log^2(\tau_j q_{k_j}) < q_{k_j}^{-2\epsilon}\log^2 q_{k_j} \lesssim \eta^2, \quad e^{-\frac{\xi}{4}\tau_j q_{k_j}} \lesssim \eta^2.$$

Fix such a q_{k_j}, and then

$$|\mathbb{E}_{n<N}\mu(n)f(T^n x_0)| \lesssim (q_{k_j}\cdot\rho_{q_{k_j}}(\frac{\eta N}{2q_{k_j}})+\eta^2)^{1/2}+\eta+\frac{e^{(\frac{1}{7}-\frac{5}{14}\xi)\tau_j q_{k_j}}}{N}.$$

Since η, q_{k_j}, M_j are now all fixed and $\rho_{q_{k_j}}$ is a function that decays to 0, then for sufficiently large N, we have $q_{k_j}\cdot\rho_{q_{k_j}}(\frac{\eta N}{2q_{k_j}})\lesssim\eta^2$ and $\frac{e^{(\frac{1}{7}-\frac{5}{14}\xi)\tau_j q_{k_j}}}{N}\lesssim\eta$, which implies the result.

□

Acknowledgments

The authors would like to thank Tianyuan Mathematical Center in Southwest (No. 11826102). W. Huang was supported by NSFC grant (12090012, 12031019, 11731003). J. Wang was supported by NSFC grant (11971233), the Outstanding Youth Foundation of Jiangsu Province (No. BK20200074), and Qing Lan Project of Jiangsu province. Z. Wang was supported by NSF grant (DMS-1753042). Q. Zhou was partially supported by National Key R&D Program of China (2020YFA0713300), NSFC grant (12071232), The Science Fund for Distinguished Young Scholars of Tianjin (No. 19JCJQJC61300) and Nankai Zhide Foundation.

References

[1] A. Avila, The absolutely continuous spectrum of the almost Mathieu operator, arXiv preprint (2008), arXiv:0810.2965 [math.DS].

[2] A. Avila, Almost reducibility and absolute continuity I, arXiv preprint (2010), arXiv:1006.0704 [math.DS].

[3] A. Avila, KAM, Lyapunov exponents and the spectral dichotomy for one-frequency Schrödinger operators, preprint.

[4] A. Avila, Global theory of one-frequency Schrödinger operators, Acta Math., **215**, 1-54 (2015).

[5] A. Avila, B. Fayad, and R. Krikorian, A KAM Scheme for $SL(2,\mathbb{R})$ cocycles with Liouvillean frequencies, Geom. Funct. Anal., **21**, 1001-1019 (2011).

[6] A. Avila, S. Jitomirskaya, Almost localization and almost reducibility, J. Eur. Math. Soc., **12**, 93-131 (2010).

[7] A. Avila, S. Jitomirskaya, Hölder continuity of absolutely continuous spectral measures for one-frequency Schrödinger operators, Comm. Math. Phys., **301**, 563-581 (2011).

[8] A. Avila, S. Jitomirskaya and C.A. Marx, Spectral theory of extended Harper's model and a question by Erdös and Szekeres. Invent. math. 210, 283-339 (2017).

[9] A. Avila, K. Khanin and M. Leguil, Invariant graphs and spectral type of Schrödinger operators, Pure and Applied Functional Analysis., **5**, 1257-1277 (2020).

[10] A. Avila, Y. Last, M. Shamis and Q. Zhou, On the abominable properties of the Almost Mathieu operator with well approximated frequencies. Preprint.

[11] A.Avila, J.You, and Z. Zhou, Sharp Phase transitions for the almost Mathieu operator, Duke Math. J. **166**, 2697-2718 (2017)

[12] A. Avila, J. You, and Q. Zhou, Dry ten Martini problem in the non-critical case, preprint.

[13] K. Bjerklöv, T. Jäger, Rotation numbers for quasiperiodically forced circle maps: Mode-locking vs strict monotonicity, J. Amer. Math. Soc. **22**, no. 2, 353-362 (2009).

[14] J. Bourgain and S. Jitomirskaya, Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential, J. Statist. Phys. **108**, 1203-1218 (2002).

[15] H. Davenport, On some infinite series involving arithmetical functions II, Quarterly Journal of Mathematics, **1**, 313-320 (1937).

[16] A. de Faveri, Möbius disjointness for $C^{1+\varepsilon}$ skew products, Int. Math. Res. Not. IMRN, doi:10.1093/imrn/rnaa185 (2020).

[17] L. H. Eliasson, Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation, Comm. Math. Phys. **146**(3), 447-482 (1992).
[18] L. H. Eliasson, Almost reducibility of linear quasi-periodic systems. Proceedings of Symposia inn Pure Mathematics. Volume 69, 2001
[19] S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk, Sarnak's conjecture: what's new, Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics, 163-235. Springer, Cham (2018).
[20] L. Ge and J. You, Arithmetic version of Anderson localization via reducibility, Geom. Funct. Anal., 30 (2020), 1370-1401.
[21] M. R. Herman, Construction d'un difféomorphisme minimal d'entropie topologique non nulle, Ergod. Th. & Dynam. Sys. **1**(1), 65-76 (1981).
[22] M. R. Herman, Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractere local d'un théoreme d'Arnold et de Moser sur le tore de dimension 2, Commen. Math. Helv. **58**(1), 453-502 (1983).
[23] X. Hou and J. You, Almost reducibility and non-perturbative reducibility of quasi-periodic linear systems, Invent. Math. **190**, 209-260 (2012).
[24] W. Huang and Y. Yi, Almost periodically forced circle flows, J. Funct. Anal., **257**(3), 832-902 (2009).
[25] W. Huang, Z. Wang, and X. Ye, Measure complexity and Möbius disjointness, Adv. Math., **347**, 827-858 (2019).
[26] W. Huang, Z. Wang, and G. Zhang, Möbius disjointness for topological models of ergodic systems with discrete spectrum, J. Mod. Dyn., **14**, 277-290 (2019).
[27] H. Iwaniec and E. Kowalski, Analytic number theory, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Soc. (2004).
[28] T. Jäger, The creation of strange non-chaotic attractors in non-smooth saddle-node bifurcations, Mem. Amer. Math. Soc., 201, no. 945, vi+106 pp. (2009).
[29] R. Johnson, and J. Moser, The rotation number for almost periodic potentials, Comm. Math. Phys. **84**, 403-438 (1982).
[30] A. Kanigowski, M. Lemańczyk and M. Radziwiłł, Rigidity in dynamics and Möbius disjointness, arXiv preprint, arXiv:1905.13256 (2019).
[31] R. Krikorian, J. Wang, J. You and Q. Zhou, Linearization of quasiperiodically forced circle flow beyond Brjuno condition. Comm. Math. Phys. **358**(1), 81-100 (2018).
[32] J. Kułaga-Przymus and M. Lemańczyk, The Möbius function and continuous extensions of rotations, Monatsh. Math., **178**, 553-582 (2015).
[33] Y. Last, Spectral theory of Sturm-Liouville operators on infinite intervals: a review of recent developments. *Sturm-Liouville Theory.* Birkhäuser Basel, 99-120 (2005).
[34] M. Leguil, J. You, Z. Zhao, and Q. Zhou, Asymptotics of spectral gaps of quasi-periodic Schrödinger operators, arXiv:1712.04700 [math.DS].
[35] J. Liu and P. Sarnak, The Möbius function and distal flows, Duke Math. J., **164**(7), 1353-1399 (2015).
[36] C. Marx and S.Jitomirskaya, Dynamics and spectral theory of quasi-periodic Schrödinger-type operators. Ergod. Th. & Dynam. Sys. **37**, 2353-2393 (2017).
[37] K. Matomäki, M. Radziwiłł, and T. Tao, An averaged form of Chowla's conjecture, Algebra Number Theory **9**(9), 2167-2196 (2015).
[38] J. Puig, A nonperturbative Eliasson's reducibility theorem, Nonlinearity **19**(2), 355-376 (2006).
[39] P. Sarnak, Three lectures on the Möbius function, randomness and dynamics, Institute for Advanced Study, New Yersey (2011).
[40] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Quantised Hall conductance in a two dimensional periodic potential, Phys. Rev. Lett., **49**, 405-408 (1982).
[41] J. Wang and T. Jäger, Abundance of mode-locking for quasiperiodically forced circle maps, Comm. Math. Phys. **353**(1), 1-36 (2017).
[42] J. Wang, Q. Zhou, and T. Jäger, Genericity of mode-locking for quasiperiodically forced circle maps, Adv. Math. **348**, 353-377 (2019).
[43] Z. Wang, Möbius disjointness for analytic skew products, Invent. Math. **209**, 175-196 (2017).

[44] J.-C. Yoccoz, Some questions and remarks about $SL(2,\mathbb{R})$ cocycles, Modern Dynamical Systems and Applications, 447-458. Cambridge University Press, Cambridge, 2004.
[45] J. You and Q. Zhou, Embedding of analytic quasi-periodic cocycles into analytic quasi-periodic linear systems and its applications, Comm. Math. Phys. **323**, 975-1005(2013).
[46] Z. Zhang, On topological genericity of the mode-locking phenomenon, Math. Ann. **376**(1-2), 707-728 (2020).

(W. Huang) CAS Wu Wen-Tsun Key Laboratory of Mathematics, School of Mathematical Sciences, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China
Email address: wenh@mail.ustc.edu.cn

(J. Wang) Department of Mathematics, School of Science, Nanjing University of Science and Technology, Nanjing, 210094, P.R. China
Email address: jing.wang@njust.edu.cn

(Z. Wang) Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA
Email address: zhirenw@psu.edu

(Q. Zhou) Chern Institute of Mathematics and LPMC, Nankai University, Tianjin 300071, China
Email address: qizhou@nankai.edu.cn