UNIFORM POSITIVITY AND CONTINUITY OF LYAPUNOV EXPONENTS FOR A CLASS OF C^2 QUASIPERIODIC SCHRÖDINGER COCYCLES

YIQIAN WANG AND ZHENGHE ZHANG

Abstract. We show that for a class of C^2 quasiperiodic potentials and for any fixed *Diophantine* frequency, the Lyapunov exponent of the corresponding Schrödinger cocycles, as a function of energies, are uniformly positive and weakly Hölder continuous. As a corollary, we obtain that the corresponding integrated density of states is weakly Hölder continuous as well. Our approach is of purely dynamical systems, which depends on a detailed analysis of asymptotic stable and unstable directions. We also apply it to more general $\mathrm{SL}(2;\mathbb{R})$ cocycles, which in turn can be applied to get uniform positivity and continuity of Lyapuonv exponents around unique nondegenerate extremal points of any smooth potential, and to a certain class of C^2 Szegő cocycles.

Keywords: Lyapunov exponents; Quasiperiodic potentials; Schrödinger operators

Contents

Y.W. was supported by the National Natural Science Foundation of China grant No. 11271183.
Z.Z. was partially supported by NSF grant DMS-1316534.

1. Introduction

Consider the family of *Schrödinger operators* $H_{\alpha,\lambda v,x}$ on $\ell^2(\mathbb{Z}) \ni u = (u_n)_{n\in\mathbb{Z}}$:

$$(H_{\alpha,\lambda v,x}u)_n = u_{n+1} + u_{n-1} + \lambda v(x+n\alpha)u_n. \tag{1}$$

Here $v \in C^r(\mathbb{R}/\mathbb{Z},\mathbb{R}), r \in \mathbb{N}\cup\{\infty,\omega\}$ is the *potential*, $\lambda \in \mathbb{R}$ *coupling constant*, $x \in \mathbb{R}/\mathbb{Z}$ *phase*, and $\alpha \in \mathbb{R}/\mathbb{Z}$ *frequency*. For simplicity, we may sometimes leave $\alpha, \lambda v$ in $H_{\alpha,\lambda v,x}$ implicit. Let $\Sigma(H_x)$ be the spectrum of the operator. Then it is well-known that

$$\Sigma(H_x) \subset [-2+|\lambda|\inf v, 2+|\lambda|\sup v]. \tag{2}$$

Moreover, for irrational α, due to a theorem of Johnson [Jo], $\Sigma(H_x)$ is phase-independent. This follows from minimality of the irrational rotation, see also [Z2] for a more recent proof. Let $\Sigma_{\alpha,\lambda v}$ denote the common spectrum in this case.

Consider the eigenvalue equation $H_x u = Eu$. Then there is an associated cocycle map which is denoted as $A^{(E-\lambda v)} \in C^r(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$, and is given by

$$A^{(E-\lambda v)}(x) = \begin{pmatrix} E-\lambda v(x) & -1 \\ 1 & 0 \end{pmatrix}. \tag{3}$$

Then $(\alpha, A^{(E-\lambda v)})$ defines a family of dynamical systems on $(\mathbb{R}/\mathbb{Z})\times\mathbb{R}^2$, which is given by $(x,w) \mapsto (x+\alpha, A^{(E-\lambda v)}(x)w)$ and is called the *Schrödinger cocycle*. The nth iteration of dynamics is denoted by $(\alpha, A^{(E-\lambda v)})^n = (n\alpha, A_n^{(E-\lambda v)})$. Thus,

$$A_n^{(E-\lambda v)}(x) = \begin{cases} A^{(E-\lambda v)}(x+(n-1)\alpha)\cdots A^{(E-\lambda v)}(x), & n \geq 1; \\ I_2, & n = 0; \\ [A_{-n}^{(E-\lambda v)}(x+n\alpha)]^{-1}, & n \leq -1. \end{cases}$$

The relation between operator and cocycle is the following. $u \in \mathbb{C}^{\mathbb{Z}}$ is a solution of the equation $H_{\lambda,x}u = Eu$ if and only if

$$A_n^{(E-\lambda v)}(x)\begin{pmatrix} u_0 \\ u_{-1} \end{pmatrix} = \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix},\ n \in \mathbb{Z}.$$

This says that $A_n^{(E-\lambda v)}$ generates the n-step transfer matrices for the operator (1).

The Lyapunov Exponent (LE for short), $L(E,\lambda)$, of this cocycle is given by

$$L(E,\lambda) = \lim_{n\to\infty}\frac{1}{n}\int_{\mathbb{R}/\mathbb{Z}} \ln\|A_n^{(E-\lambda v)}(x)\|dx = \inf_n \frac{1}{n}\int_{\mathbb{R}/\mathbb{Z}} \ln\|A_n^{(E-\lambda v)}(x)\|dx \geq 0.$$

The limit exists and is equal to the infimum since $\{\int_{\mathbb{R}/\mathbb{Z}} \ln \|A_n^{(E-\lambda v)}(x)\|dx\}_{n\geq 1}$ is a subadditive sequence. Then by Kingman's Subadditive Ergodic Theorem, we also have for irrational α,

$$L(E,\lambda) = \lim_{n\to\infty} \frac{1}{n} \ln \|A_n^{(E-\lambda v)}(x)\| \text{ for } a.e.\ x \in \mathbb{R}/\mathbb{Z}.$$

The integrated density of states (IDS for short), $N(E)$, is given by

$$N(E) = \lim_{n\to\infty} \frac{1}{n} \mathrm{card}\,\{(-\infty, E) \cap \Sigma(H_{n,x})\} \text{ for } a.e.\ x \in \mathbb{R}/\mathbb{Z}.$$

Here $H_{n,x}$ denotes the restriction of the operator $H_{\lambda,x}$ to $[1,n]$ with Dirichlet boundary condition $u_0 = u_{n+1} = 0$, $\Sigma(H_{n,x})$ the set of eigenvalues of $H_{n,x}$, and card the cardinality of a set. It is well known that the convergence is independent of Lebesgue almost every $x \in \mathbb{R}/\mathbb{Z}$. Moreover, the Lyapunov exponent L and the integrated density of states N are related via the following famous Thouless' Formula

$$L(E) = \int \log |E - E'| dN(E'), \tag{4}$$

which basically says that L is the Hilbert transform of N and vice versa. It is well-known that Hilbert transform preserves Hölder or some weak Hölder continuity (e.g. the continuity results we obtained in Theorem 2 in Section 1.1), see [GoSc] for some detailed description. In particular, Hölder and weak Hölder continuity pass from L to N and vice versa.

1.1. **Statement of main results: Theorem 1 and 2.** In this paper, from now on, we assume $v \in C^2(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ satisfy the following conditions. Assume $\frac{dv}{dx} = 0$ at exactly two points, one is minimal and the other maximal, which are denoted by z_1 and z_2. Assume that these two extremals are non-degenerate. In other words, $\frac{d^2 v}{dx^2}(z_j) \neq 0$ for $j = 1, 2$.

Fix two positive constants τ, γ. We say α is satisfying a *Diophantine* condition $DC_{\tau,\gamma}$ if

$$|\alpha - \frac{p}{q}| \geq \frac{\gamma}{|q|^\tau} \text{ for all } p, q \in \mathbb{Z} \text{ with } q \neq 0.$$

It is a standard result that for any $\tau > 2$,

$$DC_\tau := \bigcup_{\gamma > 0} DC_{\tau,\gamma}$$

is of full Lebesgue measure. We fix a $\tau > 2$ and a $\alpha \in DC_\tau$. Then, we would like to show the following results.

Theorem 1. *Let α and v be as above. Consider the Schrödinger cocycle with potential v and coupling constant λ. Let $L(E,\lambda)$ be the associated Lyapunov exponents. Then for all $\varepsilon > 0$, there exist a $\lambda_0 = \lambda_0(\alpha, v, \varepsilon) > 0$ such that*

$$L(E,\lambda) > (1-\varepsilon)\log\lambda \tag{5}$$

for all $(E,\lambda) \in \mathbb{R} \times [\lambda_0, \infty)$.

Theorem 2. *Let α and v be in Theorem 1. Consider the Schrödinger cocycle with potential v and coupling constant λ. Then there exist a $\lambda_1 = \lambda_1(\alpha, v) > 0$ such that for any fixed $\lambda > \lambda_1$, if we let $L(E)$ be the Lyapunov exponents and $N(E)$ integrated density of states (IDS), then for all $E, E' \in [\lambda \inf v - 2, \lambda \sup v + 2]$, it holds that*

$$|L(E) - L(E')| + |N(E) - N(E')| < Ce^{-c(\log |E-E'|^{-1})^\sigma}, \tag{6}$$

where $c, C > 0$ *depends on* α, v, λ, *and* $0 < \sigma < 1$ *on* α.

By the discussion following (2), $\mathbb{R} \setminus [\lambda \inf v - 2, \lambda \sup v + 2]$ is a subset of the resolvent set, in which $N(E)$ clearly stays constant. Due to a theorem of Johnson [Jo], for irrational frequency, $(\alpha, A^{(E-\lambda v)})$ is uniform hyperbolic ($\mathcal{UH}$ for short) if and only if E is in the resolvent set. See again [Z2] for a more recent proof. Then it is standard result that $L(E)$ is smooth in the $\mathcal{UH}$ region, see e.g. [A1, Section 2.1]. Thus, in particular, for these α, v, λ as in Theorem 2, $L(E)$ and $N(E)$ are weak Hölder continous functions of $E \in \mathbb{R}$.

1.2. **Remarks on positivity of Lyapunov exponents.** Positivity of LE for Schrödinger cocycle is closely related to the spectral properties of the corresponding Schrödinger operators. For instance, by Ishii-Pastur-Kotani [I, P, Ko1], for general bounded ergodic potential, positivity of LE for almost every energy is equivalent to the absence of absolutely continuous spectrum for almost every phase.

Moreover, positivity of LE for all energies is closely related to the Anderson Localization phenomenon. In fact, for the type of potentials considered in Theorem 1, Anderson Localization has been established by Sinai and Fröhlich-Spencer-Wittwer [Sin, FrSpWi]. Note in [FrSpWi], the authors also assumed that the potentials are even functions. These authors developed some inductive multi-scale procedures to get exponentially decaying eigenstates. One could extract a similar result as Theorem 1, that is, $L(E, \lambda) > \frac{1}{2} \log \lambda$ for all $E \in \mathbb{R}$, from the proofs in [Sin, FrSpWi]. Very recently, Bjerklöv also obtain among other things a similar result, $L(E, \lambda) > \frac{2}{3} \log \lambda$ for all $E \in \mathbb{R}$, via his approach, see [Bj1].

Clearly, the estimate (5) obtained in this paper is stronger. Combined with some additional arguments, it actually leads to a version of Large deviation theorem (LDT for short) that is crucial for the proof of Theorem 2, which is the first result of this kind. See Section 1.3 for the further remarks.

On the other hand, positivity of LE for Schrödinger cocycles, or more generally, $\mathrm{SL}(2, \mathbb{R})$ cocycles, is one of the central topics in dynamical systems. Thus, it has been extensively studied by both dynamicists and mathematical physicists. For different base dynamics, both the mechanisms and phenomena are very different. Let us list some of the related results.

For the i.i.d. potentials, Furstenberg [Fu] showed that, among other things, LE is uniformly positive for all energies. For ergodic potentials Kotani [Ko2] showed that LE is positive for almost every energy if the potential is non-deterministic. Moreover, Kotani [Ko2] showed that ergodic potential taking finitely many values is non-deterministic if it is aperiodic, hence the corresponding LE is positive for a.e. E. Based on this result of Kotani, together with some new interesting ingredients, Avila-Damanik [AD] showed that for generic continuous potentials defined on compact metric spaces, if the ergodic measure of the base dynamics is non-atomic, then LE is positive for almost every energy.

For doubling map on the unit circle or Anosov diffeomorphism on two dimensional torus, see Chulaevsky-Spencer and Bourgain-Schlag [ChuSp, BoSc]. For skew shifts, see Bourgain-Goldstein-Schlag, Bourgain, and Krüger [BoGoSc, Bou2, Bou3, Kr1, Kr2]. For limit periodic potentials, of which the base dynamics is minimal translations on Cantor group, see Avila [A4]. Let us remark also that Avila [A1] showed that positivity of LE is a dense phenomenon on any suitable base dynamics and in any usual regularity classes.

The most intensively studied cases are quasiperiodic potentials. For real analytic potentials, the first breakthrough is due to Herman [H]. By subharmonicity, among other things, the author showed that LE is uniformly positive for trigonometric polynomials. These techniques have been further developed by Sorets-Spencer [SoSp] for arbitrary one-frequency nonconstant real analytic potentials and for large disorders. Same results for *Diophantine* multi-frequency were established by Bourgain-Schlag [BoGo] and Goldstein-Schlag [GoSc]. Bourgain [Bou3] obtained the same results for any rational independent multi-frequency. Based on new results in [A2], Zhang [Z1] gave a different proof of the [SoSp] results. He also applied it to a certain class of analytic Szegő cocycle and obtained the uniform positivity of the associated LE.

Most results mentioned in the above paragraph do not require the *Diophantine* type of conditions for frequency since one has subharmonicity. For a class of Gevrey potentials and strong *Diophantine* frequencies, see Klein [Kl]. Eliasson [E] also gets some related results for a certain class of Gevrey potentials and for some strong *Diophantine* frequencies.

For smooth potentials, it seems that a complicated induction and some *Diophantine* type of conditions are necessary to take care of the small divisor type of problems. Other than works in [FrSpWi, Sin], some recent works may be found in [Bj2, Cha] for more general smooth potentials. In [Bj2], the author used techniques that are close in spirit to Benedicks-Carleson [BeCa] type of techniques for Henón map, and a positive measure of frequencies and energies are excluded. In [Cha], the author used multi-scale analysis, and uniform positivity of LE for some C^3 potentials is obtained by excluding a positive measure of frequencies and by varying the potentials in some typical way.

The method used in this paper is of purely dynamical systems, which is a further development of those techniques from Young [Y], and close also in spirit to Benedicks-Carleson [BeCa]. The techniques in [Y] have been applied to Schrödinger cocycles by Zhang [Z1], and some results concerning positivity of LE for general smooth potentials and for fixed *Brjuno* frequencies have been obtained. Roughly speaking, these techniques based on some detailed analysis of asymptotic stable and unstable directions. The key idea is to classify the ways that they intersect with each other. Then, one need to develop some induction schemes to show that these ways are all the possibilities of intersection between them.

In those cases considered in [Y, Z1], one again needs to exclude a positive measure of energies to get the nonresonance condition. Then it is showed that under the nonresonance condition, the n-step stable and n-step unstable directions always intersect in a transversal way, which makes the induction easier. And in this case, a C^1 type of estimates of the asymptotic stable and unstable direction is sufficient. The nonresonance condition also makes sure that for the survived parameters, the dynamical systems are *nonuniformly hyperbolic* ($\mathcal{NUH}$ for short). This is due to the fact that the intersection between asymptotic stable and unstable directions persist in larger and larger time scale, which eventually implies the intersection of stable and unstable directions, hence, $\mathcal{NUH}$. Back to the model in Theorem 1, while the statement of Theorem 1 does not necessarily distinguish energies between the spectrum and the resolvent set, we actually have the following Corollary of [Z1, Theorem B′].

Corollary 1. *Let α and v be as in Theorem 1. Then for each $\lambda > \lambda_0$, there exists a $\Omega_{\alpha,\lambda v} \subset \Sigma_{\alpha,\lambda v}$ such that*

$$\lim_{\lambda\to\infty} \frac{\mathrm{Leb}(\Omega_{\alpha,\lambda v})}{\lambda(\sup v - \inf v)} = 1,$$

and for each $E \in \Omega_{\alpha,\lambda v}$, there exists some $x \in \mathbb{R}/\mathbb{Z}$ such that the eigenvalue equation $H_{\alpha,\lambda v,x}u = Eu$ admits some exponentially decaying eigenvectors.

Remark 1. *The last statement of Corollary 1 implies that $|u_n| < Ce^{-L|n|}$ for all $n \in \mathbb{Z}$ for some constant $C, L > 0$, which is kind of Anderson Localization phenomenon. Also, in Corollary 1, we can actually relax the* Diophantine *condition to the* Brjuno *condition, see* [Z1, Theorem B′].

In this paper, we will not exclude any parameter. Thus, the main difficulty of the cases considered in this paper is the occurrence of 'resonance'. This leads to bifurcation of the way n-step stable direction intersecting with n-step unstable direction: our analysis shows that 'resonance' leads to some tangential way of intersection or even separation of n-step stable and unstable directions, which leads to *uniformly hyperbolic* ($\mathcal{UH}$) systems, see Figure 3. In fact, to start with, one encounters with tangential intersections of first step stable and unstable directions. Thus one needs some nondegenerate conditions to get started. And a new induction scheme that includes both 'nonresonance' and 'resonance' cases needs to be introduced. Moreover, to deal with the tangential type of intersection, a C^2 type of estimate of the asymptotic stable and unstable directions is required.

1.3. **Remarks on regularity of Lyapunov exponents.** Much work has been devoted to the regularity properties of Lyapunov exponents (LE) and integrated density of states (IDS) as well. By the discussion following (4), we focus on the regularity of LE here.

On the regularity of LE for real analytic quasi-periodic potentials, a series of positive results have been obtained in the 2000s. It starts with the work of Goldstein-Schlag [GoSc] where they obtained some sharp version of large deviation theorems (LDT) for real analytic potentials with strong *Diophantine* frequency, developed a powerful tool, the Avalanche Principle, and proved Hölder or weak Hölder continuity of $L(E)$ in the regime of positive LE. Notice that LDT for real analytic potentials with *Diophantine* frequency was first established in [BoGo] in order to get Anderson Localization. This also illustrates the power and importance of LDT.

Avalanche Principle involves only long finite products of matrices, see Section 5.2. Thus, the key to apply the method in [GoSc] to other cases is to establish LDT. For other type of base dynamics, Bourgain-Goldstein-Schlag [BoGoSc] obtained the results for skewshift base dynamics, Bourgain-Schlag [BoSc] for doubling map and Anosov diffeomorphism. For lower regularity case, [Kl] got the results for a class of Gevrey potentials. These results concern continuity with respect to energies. For wider class of cocycle maps, Jitomirskaya-Koslover-Schulteis [JiKoSch] get the continuity of LE for a class of analytic quasiperiodic $\mathrm{M}(2,\mathbb{C})$ cocycles which is applicable to general quasi-periodic Jacobi matrices or orthogonal polynomials on the unit circle in various parameters. Jitomirskaya-Marx [JiMar1] later extended it to all (including singular) $\mathrm{M}(2,\mathbb{C})$ cocycles. Hölder continuity for $\mathrm{GL}(d,\mathbb{C})$ cocycles, $d \geq 2$, was recently obtained in Schlag [Sc] and Duarte-Klein [DuKl]. All the results stated above, except strongly mixing cases, require *Diophantine* condition.

An arithmetic version of large deviations and an arithmetic inductive scheme were developed in [BoJi] allowing to obtain joint continuity of LE for $\mathrm{SL}(2,\mathbb{C})$ cocycles, in frequency and cocycle map, at any irrational frequencies. This result has been crucial in later proofs of the Ten Martini problem [AvJi], Avila's global theory of one-frequency cocycles [A2, A3], and other important developments. It was extended to multi-frequency case by Bourgain [Bou3] and to general $\mathrm{M}(2,\mathbb{C})$ case in [JiMar2]. More recently, a completely different proof, not using LDT or Avalanche principle, and extending to the general $\mathrm{M}(d,\mathbb{C})$, $d \geq 2$, case was developed in Avila-Jitomirskaya-Sadel [AvJiSa]. All these results however rely heavily on analyticity of the cocycle map.

Thus, Theorem 2 in this paper is striking in the sense that it provides the first positive result on the continuity of LE and weak Hölder continuity of IDS on E (note log-Hölder continuity of IDS on E holds for general ergodic bounded potentials, see [CrSim]) for C^r, $r \leq \infty$, quasi-periodic potentials. More concretely, surprisingly, it turns out that some version of LDT follows naturally from our induction scheme, see Section 5 for details. Thus, combined with the Avalanche Principle, Theorem 2 follows essentially from the same argument as in [GoSc].

We remark that Theorem 2 has an analog as in [JiKoSch], that is, we can prove continuity of LE with respect to the C^2 *cos*-type of potentials.

For other related results, Avila-Krikorian [AK] recently studied so-called monotonic cocycles which are a class of smooth or analytic cocycles non-homotopic to constant. They proved that the LE is smooth or even analytic, respectively. In comparison, the regularity of LE cannot be better (as far as the modulus of continuity is concerned) than 1/2-Hölder continuous for cocycles homotopic to constant which automatically includes the category of Schrödinger cocycles. However, Avila [A2] recently showed that if one stratifies the energies or some real analytic family of real analytic potentials in some natural way, then the LE is in fact real analytic.

There are many negative results on the positivity and continuity of LE for non-analytic cases. It is well known that in C^0-topology, discontinuity of LE holds true at every non-uniformly hyperbolic cocycle, see [Fur, Kn, T]. Moreover, motivated by *Mañé* [Ma1, Ma2], Bochi [Boc1, Boc2] proved that with an ergodic base system, any non-uniformly hyperbolic $\mathrm{SL}(2,\mathbb{R})$-cocycle can be approximated by cocycles with zero LE in the C^0 topology.

Based also on the the method of Young[Y], Wang-You [WaYo1] constructed examples to show that LE can be discontinuous even in the space of C^∞ Schrödinger cocycles. Recently, Wang-You [WaYo2] has improved the result in [WaYo1] by showing that in C^r topology, $1 \leq r \leq \infty$, there exists Schrödinger cocycles with positive LE that can be approximated by ones with zero LE. The example in [WaYo2] also showed that the nondegenerate condition for the potential in Theorem 1 and 2 is necessary for positivity and continuity of LE. Jitomirskaya-Marx [JiMar2] constructed examples showing that LE of $\mathrm{M}(2,\mathbb{C})$ cocycles is discontinuous in C^∞ topology.

Finally, let us remark that the continuity of LE for Schrödinger cocycles is also expected to play important roles in studying Cantor spectrum, typical localization length, phase transition, etc, for quasi-periodic Schrödinger operators.

1.4. **Generalization and further comments.** Though Theorem 1 is our primary interest, our method is not restricted to Schrödinger cocycles. What we proved is actually a more general version concerning smooth quasiperiodic $\mathrm{SL}(2,\mathbb{R})$ cocycles, see Corollary 5 of Appendix Section B. In particular, we obtain the following corollary of Corollary 5 and [Z1, Theorem B′]. We say $v \in C^2(\mathbb{R}/\mathbb{Z},\mathbb{R})$ has a unique maximal point if the set $\{x : v(x) = \max_{y\in\mathbb{R}/\mathbb{Z}} v(y)\}$ consists of a single point (for simplicity, we state only the maximal point case. The minimal point case can be stated similarly). Then we have the following corollary.

Corollary 2. *Let α be as in Theorem 1. Assume $v \in C^2(\mathbb{R}/\mathbb{Z},\mathbb{R})$ has a unique nondegenerate maximal point which is denoted by x_0. Then there exists a $r > 0$ such that for each $\varepsilon > 0$, there exists a $\lambda_0 = \lambda_0(\alpha, v, \varepsilon, r)$ such that for all $(E,\lambda) \in \lambda[v(x_0) - r, v(x_0) + r] \times (\lambda_0, \infty)$,*

$$L(E,\lambda) > (1-\varepsilon)\log\lambda.$$

Moreover, for any fixed $\lambda > \lambda_0$ and for all $E, E' \in \lambda[v(x_0) - r, v(x_0) + r]$, it holds that

$$|L(E) - L(E')| + |N(E) - N(E')| < Ce^{-c(\log|E-E'|^{-1})^{\sigma}},$$

where $c, C > 0$ depends on $v, \alpha, \varepsilon, r, \lambda$, and $0 < \sigma < 1$ on α. Finally, we have

$$\lim_{\lambda\to\infty} \frac{1}{\lambda r}\mathrm{Leb}\,\{\Sigma_{\alpha,\lambda v} \cap \lambda[v(x_0) - r, v(x_0)]\} = 1.$$

In other words, the LE is positive and continuous for all energies around the unique non-degenerate extremals of potentials for large disorders. This corollary says that, in some sense, the positivity of LE is a local property with respect to the initial 'critical interval' of the potential. Corollary 5 can also be applied to a certain class of quasiperiodic Szegő cocycles, see Corollary 6 of Section B. For details and other applications, see Section B.

To sum up, we believe that our method may have the following further development. Firstly, although the computation will be much more complicated, it is possible that our techniques can be used to analyze more general smooth potentials. For instance, instead of C^2 estimate of Lemma 4, we may need C^r for $r < \infty$. Moreover, we may need to deal with the new types of resonance, e.g. resonance between the type I and type II functions of Definition 2.

Secondly, since our method is based on a detailed analysis of asymptotic stable and unstable directions, it has the advantage in showing the occurrence of $\mathcal{UH}$. In other words, it has the advantage in showing Cantor spectrum for the type of potentials in Theorem 1, or even for more possible potentials. In fact, based on the induction scheme developed in this paper, we have already showed the Cantor spectrum for the same α and v as in Theorem 1, see [WaZ].

Thirdly, it is possible to relax *Diophantine* condition to *Brjuno* or even weak *Liouville* conditions in Theorem 1 and 2. Moreover, it is also possible to improve the index σ in (6) to 1 which is nothing other than the Hölder continuity. We do not pursue these goals here in order to keep this paper to a reasonable length.

Finally, the idea of analyzing the asymptotic stable and unstable directions is probably not restricted to one-frequency quasiperiodic case. These techniques are also considered to be promising in [A2, AK].

1.5. **Structure of the Paper.** The structure of the remaining part of this papers is as follows. In Section 2, we state a series of technical lemmas. We first reduce the Schrödinger cocycles to its polar decomposition form so that we can get started with our induction. Then we state the series of Lemmas that will be used to control the derivatives of asymptotic stable and unstable directions and the norms the iteration of cocycles, and concatenation of sequence of matrix-maps. Then, we classify the types of functions that will be used to describe all possible ways the n-step stable direction intersecting with n-step unstable direction. Finally, we state and prove a easy corollary which actually builds the bridge of concatenation of sequence of matrix-maps and our classification of the intersection between asymptotic stable and unstable directions . The proof of our induction and Theorem 1 and 2 are just some repeated applications of these lemmas.

In Section 3, we will get started with our induction. We will start with step 1 and move one step forward to step 2. So we get to know all possible cases that will occur in our induction. In Section 4, we state and prove our induction. In Section 5, we prove our main results. We first show a key lemma, which easily implies Theorem 1 and a version of LDT. Finally, we prove Theorem 2. In appendix Section A, we prove Lemma 1–6 that are given in Section 2. In Section B, we state a more general version of Theorem 1 and 2, and give some applications.

Acknowledgments. Z.Z. would like to thank his advisors Artur Avila and Amie Wilkinson for suggesting the project of uniformly positive Lyapunov exponents for smooth potentials, for helpful discussions, and for their encouragement and continuous support. He also would like to thank Vadim Kaloshin for suggesting this project, for some helpful discussion, and for showing a note joint with Anton Gorodetski that gives some helpful hints. We are grateful to Michael Goldstein for some helpful discussion and for showing us a manuscript of him which gives us some positive hints. It is our pleasure to thank Jiangong You for some helpful discussions. We also owe our thanks to Svetlana Jitomirskaya for detailed comments and suggestions. Finally, we also would like to thank the referee for many detailed comments which improve greatly the presentation of this paper.

2. Preliminary: statement of Some Technical Lemmas

From now on, if not stated otherwise, let $C,\ c$ be some universal positive constants depending only on v and α, where C is large and c small. Let $\frac{p_s}{q_s}$ be the sth continued fraction approximants of frequency α. Then it is a standard result that $\alpha \in DC_\tau$ if and only if there is some $c > 0$ such that $q_{s+1} < cq_s^{\tau-1}$ for all $s \geq 1$. We will sometimes use this equivalent condition. For two positive real number $a, b > 0$, by $a \gg b$ or $b \ll a$, we mean that a is sufficiently larger than b.

For $\theta \in \mathbb{R}/(2\pi\mathbb{Z})$, let

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}(2, \mathbb{R}).$$

Define the map

$$s : \mathrm{SL}(2, \mathbb{R}) \to \mathbb{RP}^1 = \mathbb{R}/(\pi\mathbb{Z})$$

so that $s(A)$ is the most contraction direction of $A \in \mathrm{SL}(2, \mathbb{R})$. Let $\hat{s}(A) \in s(A)$ be an unit vector. Thus, $\|A \cdot \hat{s}(A)\| = \|A\|^{-1}$. Abusing the notation a little, let

$$u : \mathrm{SL}(2, \mathbb{R}) \to \mathbb{RP}^1 = \mathbb{R}/(\pi\mathbb{Z})$$

be that $u(A) = s(A^{-1})$. Then for $A \in \mathrm{SL}(2, \mathbb{R})$, it is clear that

$$A = R_u \cdot \begin{pmatrix} \|A\| & 0 \\ 0 & \|A\|^{-1} \end{pmatrix} \cdot R_{\frac{\pi}{2}-s}, \tag{7}$$

where $s, u \in [0, 2\pi)$ are some suitable choices of angles correspond to the directions $s(A), u(A) \in \mathbb{R}/(\pi\mathbb{Z})$.

2.1. **Polar decomposition of Schrödinger cocycles: Lemma 1.** Instead of proving Theorem 1 directly, we will use the following equivalent form of the cocycle map (3).

Lemma 1. *Let $\mathcal{I} \subset \mathbb{R}$ be some compact interval. For $x \in \mathbb{R}/\mathbb{Z}$ and $t \in \mathcal{I}$, define the following cocycles map*

$$A(x) = \Lambda(x) \cdot R_{\phi(x,t)} := \begin{pmatrix} \lambda(x) & 0 \\ 0 & \lambda^{-1}(x) \end{pmatrix} \cdot \begin{pmatrix} \frac{t-v(x)}{\sqrt{(t-v(x))^2+1}} & \frac{-1}{\sqrt{(t-v(x))^2+1}} \\ \frac{1}{\sqrt{(t-v(x))^2+1}} & \frac{t-v(x)}{\sqrt{(t-v(x))^2+1}} \end{pmatrix}, \tag{8}$$

where $\cot \phi(x,t) = t - v(x)$. *Assume*

$$\lambda(x) > \lambda, \ \left| \frac{d^m \lambda(x)}{dx^m} \right| < C\lambda, \ m = 1, 2. \tag{9}$$

Then to prove Theorem 1 and 2, it suffices to consider the cocycle map (8).

The proof of Lemma 1 will be given in Section A.1.

2.2. **Simplification of most contracted directions: Lemma 2.** We first reduce the estimate of most contracted directions to the estimate of some concrete functions that we are able to deal with.

Recall that an *elementary function* is a single variable function built from a finite number of exponentials, logarithms, constants, and nth roots, trigonometric functions and their inverses through composition and combinations using the four elementary operations $(+, -, \times, \div)$.

We call $Op(x)$ is a *one-time operation* if It is one of the following

$$x^c, \ e^x, \ \log(x), \ \sin(x), \ \sin^{-1}(x),$$

where c a constant. We assume the elementary functions we considered are always in their simplest forms. In other words, these forms have as few one-time operations as possible. Then we say an elementary function is *splittable* if It is a sum of at least two but finitely many nontrivial elementary functions with the fewest one-time operations. Otherwise, we say It is *non-splittable*. For instance, $\frac{\sin x}{1+x^2}$ is non-splittable while $x(\tan x + e^x)$ is splittable.

We call n-variable function $g(x_1, \ldots, x_n)$ *elementary* if It is elementary in each variable. We say g is *non-splittable* if It is non-splittable in each variable. For instance, $x_1 + x_2$ is splittable while $x_1 x_2$ is non-splittable.

Let g and h be n-variable elementary functions. Fix a $r \geq 1$, consider $e_i \in C^r(I, \mathbb{R}), 1 \leq i \leq n$. Then we define $G(x) = g[e_1(x), \ldots, e_n(x)]$ and $H(x) = h[e_1(x), \ldots, e_n(x)]$. We say $G(x)$ dominates $H(x)$ if

$$\|H(x)\|_\infty \ll \|G(x)\|_\infty.$$

Assume $g = g_1 + g_2$ with g_1 and g_2 elementary functions. Similarly, define $G_i(x) = g_i(e_1, \ldots, e_n)$ $i = 1, 2$. Thus $G(x) = G_1(x) + G_2(x)$. As function of x, $\frac{d^m G}{dx^m}$, $0 \leq m \leq r$, are no longer elementary. But It is still make sense to consider them as elementary functions of e_i, $1 \leq i \leq n$. Then we have the following definition.

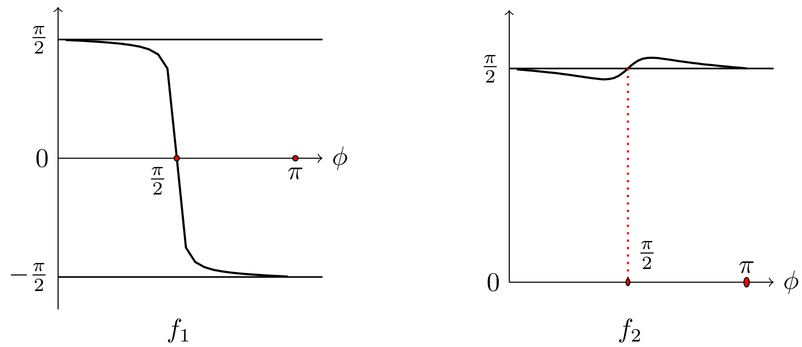

FIGURE 1. Graph of simplified most contracted directions

Definition 1. *We say G_1 is the C^0 dominant term of G if G_1 dominates G_2. Moreover, for each $1 \le l \le r$, we say G_1 is the C^l dominant term of G if in addition the following holds:*

for each $1 \le m \le l$, $\frac{d^m G_2}{dx^m}$ is dominated by some non-splittable term of $\frac{d^m G_1}{dx^m}$.

For each $0 \le l \le r$, we denote the relation "G_1 is the C^l dominant term of G" by

$$G \precsim_l G_1.$$

Remark 2. *The definition implies if $G \precsim_l G_1$ with $1 \le m \le l$, then it suffices to estimate each non-splittable term of $\frac{d^m G_1}{dx^m}$ to get the upper-bound of $\left|\frac{d^m G}{dx^m}\right|$.*

Then we define the following two functions which essentially play the role of simplified most contracted directions:

$$f_1(x, \phi) = \tan^{-1}(x^2 \cot\phi) = \frac{\pi}{2} - \cot^{-1}(x^2 \cot\phi),$$

$$f_2(x, y, \phi) = \tan^{-1}\left[x^2\left(\cot\phi + y^{-4}\tan\phi\right)\right] = \frac{\pi}{2} - \cot^{-1}\left[x^2\left(\cot\phi + y^{-4}\tan\phi\right)\right].$$

For some fixed $x > 0, y > 0$, we clearly have for all $\phi \in \mathbb{R}$,

$$|f_2(x, y, \phi)| \ge x^2 y^{-2}.$$

For some fixed $x \gg y > 0$, consider f_1 and f_2 as function of ϕ, then we have Figure (1) as graphs of f_1 and f_2.

Now we consider functions $e_1(x), e_2(x), \theta(x) \in C^2(I, \mathbb{R})$ with $e_1(x) > 0$ and $e_2(x) > 0$. For simplicity, x will be left implicit for these functions. Then we have the following lemma.

Lemma 2. *Let $e_1, e_2, \theta \in C^2(I, \mathbb{R})$ be as above. Consider the function*

$$s(x) = s[E(x)],\ u(x) = u[E(x)] = s[E(x)^{-1}] : I \to \mathbb{RP}^1,$$

where $I \subset \mathbb{RP}^1$ is a connected interval and

$$E(x) := \begin{pmatrix} e_2 & 0 \\ 0 & e_2^{-1} \end{pmatrix} R_\theta \begin{pmatrix} e_1 & 0 \\ 0 & e_1^{-1} \end{pmatrix}.$$

Define the following functions

$$F_1(x) = f_1(e_1, \theta),\ \widetilde{F}_1(x) = \frac{\pi}{2} - f_1(e_2, \theta),$$

$$F_2(x) = f_2(e_1, e_2, \theta),\ \widetilde{F}_2(x) = \frac{\pi}{2} - f_2(e_2, e_1, \theta).$$
$$e_3(x) = \|E(x)\|.$$

Set $e_0 = \min_{x\in I}\{e_1(x), e_2(x)\}$*. Then we have the following results.*

- *If* $e_0 \gg 1$ *and*

$$\inf_{x\in I} |\theta(x) - \frac{\pi}{2}| \gg e_0^{-1}, \tag{10}$$

then it holds that

$$s \precsim_2 F_1,\ u \precsim_2 \widetilde{F}_1,\ e_3 \precsim_2 e_1 e_2 |\cos\theta|. \tag{11}$$

- *If* $e_2(x) \gg e_1(x) \gg 1$ *for each* $x \in I$*, then it holds that*

$$s \precsim_2 F_1,\ u \precsim_2 \widetilde{F}_2,\ e_3 \precsim_2 \sqrt{(e_1 e_2 \cos\theta)^2 + (e_1^{-1} e_2 \sin\theta)^2}. \tag{12}$$

- *if* $e_1(x) \gg e_2(x) \gg 1$ *for each* $x \in I$*, then it holds that*

$$s \precsim_2 F_2,\ u \precsim_2 \widetilde{F}_1,\ e_3 \precsim_2 \sqrt{(e_1 e_2 \cos\theta)^2 + (e_1 e_2^{-1} \sin\theta)^2}. \tag{13}$$

2.3. **Estimates of derivatives of most contracted directions: Lemma 3–6.** The following series of Lemmas will be quite involved in our induction scheme. Basically, under suitable conditions, they deals with the product of sequence of $\mathrm{SL}(2,\mathbb{R})$ matrices maps that are defined on small intervals of $\mathbb{R}/\mathbb{Z}$. To get exponential growth of the norm of the products for larger and larger time scale, on one hand we need to control the geometrical properties of the forward and backward most contraction directions. On the other hand, we also need to control the derivatives of the norms. And we need to deal with both *resonance* and *nonresonance* cases. Consider the product $E_2 \cdot E_1$, we say we are in *nonresonance* case if

$$|s(E_2) - u(E_1)|^{-1} \ll \min\{\|E_1\|,\ \|E_2\|\}.$$

Otherwise, we say we are in *resonance* case. Proofs of Lemmas 3–6 can be found in Section A.3–A.4.

Lemma 3. *Let* $E(x)$*,* $e_0 = \min\{e_1, e_2\} \gg 1$ *be as in the Lemma 2 and* $e_3(x) = \|E(x)\|$*. Assume* $0 < \eta \ll 1$*. Suppose that for all* $x \in I$*,* $j, m = 1, 2$*, we have*

$$\left|\frac{d^m e_j}{dx^m}(x)\right| < C e_j^{1+m\eta};\ \left|\frac{d^m \theta}{dx^m}\right|,\ |\theta - \frac{\pi}{2}|^{-1} < C e_0^{\eta}.$$

Then we have

$$\left\|s - \frac{\pi}{2}\right\|_{C^2} < C e_1^{-(2-5\eta)},\ \|u\|_{C^2} < C e_2^{-(2-5\eta)}; \tag{14}$$

$$\left|\frac{d^m e_3}{dx^m}(x)\right| < C e_3^{1+m\eta} \text{ for all } x \in I \text{ and } m = 1, 2. \tag{15}$$

Then we move Lemma 3 forward to the product of n matrices in nonresonance case for arbitrary $n \in \mathbb{Z}_+$, which is as follows.

Let $I \subset \mathbb{T}$ be a interval. Consider a sequence of maps

$$E(\ell) \in C^2(I, \mathrm{SL}(2,\mathbb{R})),\ 0 \le \ell \le n-1.$$

Let $s(\ell) = s[E(\ell)]$, $u(\ell) = u[E(\ell)]$, $\lambda(\ell) = \|E(\ell)\|$, and $\Lambda(\ell) = \begin{pmatrix} \lambda(\ell) & 0 \\ 0 & \lambda(\ell)^{-1} \end{pmatrix}$. By (7), it holds that

$$E(\ell) = R_{u(\ell)} \Lambda(\ell) R_{\frac{\pi}{2} - s(\ell)}.$$

Set for each $0 \le \ell \le n-1$,

$$E_k(\ell) = \begin{cases} E(k-1+\ell)\cdots E(\ell), & 1 \le k \le n-\ell; \\ Id, & k = 0; \\ [E_{-k}(\ell+k)]^{-1}, & -\ell \le k \le -1 \end{cases}$$

For $k \ge 1$, let $s_k(\ell) = s(E_k(\ell))$, $u_k(\ell) = s(E_{-k}(\ell))$, $\lambda_k(\ell) = \|E_k(\ell)\|$ and $\Lambda_k(\ell) = \begin{pmatrix} \lambda_k(\ell) & 0 \\ 0 & \lambda_k(\ell)^{-1} \end{pmatrix}$. Again from (7), it holds that

$$E_k(\ell) = R_{u_k(\ell+k)} \Lambda_k(\ell) R_{\frac{\pi}{2} - s_k(\ell)}.$$

Lemma 4. *Let $E(\ell)$ and $E_k(\ell)$ be as above. Let*

$$0 < \eta \ll 1 \ll \lambda_0 := \min_{0 \le \ell \le n-1} \{\lambda(\ell)\}.$$

Assume that for every $x \in I$, $m = 1, 2$ and $0 \le \ell \le n-1$,

$$\left|\frac{d^m \lambda(\ell)}{dx^m}\right| < C[\lambda(\ell)]^{1+m\eta}; \ \left|\frac{d^m s(\ell)}{dx^m}\right|, \ \left|\frac{d^m u(\ell)}{dx^m}\right|, \ |s(\ell) - u(\ell-1)|^{-1} < C\lambda_0^{\eta}. \tag{16}$$

Then we have that

$$\|u(n-1) - u_n(n)\|_{C^2} < C\lambda_0^{-(2-5\eta)}, \ \|s(0) - s_n(0)\|_{C^2} < C\lambda_0^{-(2-5\eta)}; \tag{17}$$

$$\left|\frac{d^m \lambda_n(0)}{dx^m}(x)\right| < C\lambda_n(0)^{1+m\eta}, \ m = 1, 2; \tag{18}$$

$$\|\lambda_n(0)\| > \left(\prod_{\ell=0}^{n-1} \lambda(\ell)\right)^{1-\eta}. \tag{19}$$

Remark 3. *By* [Z1, Lemma 11]*, merely the assumptions $|s(\ell) - u(\ell-1)|^{-1} < C\lambda_0^{\eta}$, $0 \le \ell < n$, imply that for each $1 \le \ell \le n-1$,*

$$\|\lambda_k(\ell)\| \ge \left(\prod_{l=0}^{k-1} \lambda(l+\ell)\right)^{1-\eta} \ge \lambda_0^{k(1-\eta)}, \ 1 \le k \le n-\ell. \tag{20}$$

On the other hand, By the proof of Lemma 3 and 4 in Section A.2, it is not difficult to see that in order to get a C^0 version of (17), one only needs to assume that the norm of the sequence of matrices are large and $|s^{(\ell)} - u^{(\ell-1)}|^{-1}$ is not large with respect to norms. If in addition, one needs C^1 version of (17), then one just needs to add the corresponding C^1 control of the norm maps, s and u. In particular, the C^1 version of Lemma 4 is essentially the same with [Y, Lemma 3].

For the *resonance* case, consider $E_2 \cdot E_1$, $s(E_2) - u(E_1)$ may pass through 0. With the help of Lemma 2, we will show that some good estimate still holds true if $\|E_2\| \gg \|E_1\|$ or $\|E_1\| \gg \|E_2\|$. We first estimate the derivatives of the norm functions, and give the upper-bound of the most contraction direction.

Lemma 5. *Let $E(x) = E_2(x)E_1(x)$. Define $e_3(x) = \|E(x)\|$ and $e_0 = \min\{e_1, e_2\}$. Assume $0 < \eta \ll 1 \ll e_0$ and $0 < \beta \ll 1$. Suppose $e_1 \le e_2^{\beta}$ or $e_2 \le e_1^{\beta}$, and for $\theta(x) = s[E_2(x)] - u[E_1(x)]$ and each $x \in I$, $j, m = 1, 2$, it holds that*

$$\left|\frac{d^m e_j}{dx^m}(x)\right| < Ce_j^{1+m\eta}; \ \left|\frac{d^m \theta}{dx^m}\right| < Ce_0^{\eta}.$$

Then we have for $m = 1, 2$,

$$\left|\frac{d^m s[E(x)]}{dx^m}\right| < Ce_1^{4+2\eta},\ \left|\frac{d^m u[E(x)]}{dx^m}\right| < Ce_3^{-\frac{3}{2}}\ \text{if } e_1 \le e_2^{\beta}; \tag{21}$$

$$\left|\frac{d^m u[E(x)]}{dx^m}\right| < Ce_2^{4+2\eta},\ \left|\frac{d^m s[E(x)]}{dx^m}\right| < Ce_3^{-\frac{3}{2}}\ \text{if } e_2 \le e_1^{\beta}; \tag{22}$$

$$\left|\frac{d^m e_3}{dx^m}(x)\right| < Ce_3^{1+m\eta+2m\eta\beta}. \tag{23}$$

However, in the resonance case, we also need a C^2 lower bound near C^1 degenerate points. Here again, from the proof of Lemma 2 and 6, we are allowed to reduce the estimates of most contracted directions to estimates of simpler functions in (12) and (13).

We first define the following three types of functions, which basically classify all the possible ways that the n-step stable directions intersecting with unstable directions. In particular, the type III functions are going to describe the resonance case, from which we also have a bifurcation procedure.

Let $B(x, r) \subset \mathbb{T}$ be the ball centered around $x \in \mathbb{T}$ with radius r. For a connected interval $J \subset \mathbb{T}$ and constant $0 < a \le 1$, let aJ be the subinterval of J with the same center and whose length is $a|J|$. Let $I \subset \mathbb{T}$ be a connected interval. Without loss of generality, let $I = B(0, r)$ and l_0 satisfy $l_0 \gg r^{-1} \gg 1$. Consider for some small $\beta > 0$, a function $l : I \to \mathbb{R}$ such that

$$l(x) > l_0 \text{ and } \frac{dl^m}{dx^m}(x) < l(x)^{1+\beta},\ \forall x \in I,\ m = 1, 2. \tag{24}$$

Then we define the following types of functions, see Figure 2 for their graphs.

Definition 2. *Let I and l be as above. Let $f \in C^2(I, \mathbb{RP}^1)$. Then*

- f **is of type** I *if we have the following.*
  - $\|f\|_{C^2} < C$ *and* $f(x) = 0$ *has only one solution, say* x_0*, which is contained in* $\frac{I}{3}$*;*
  - $\frac{df}{dx} = 0$ *has at most one solution on* I *while* $|\frac{df}{dx}| > r^2$ *for all* $x \in B(x_0, \frac{r}{2})$*;*
  - *let* $J \subset I$ *be the subinterval such that* $\frac{df}{dx}(J) \cdot \frac{df}{dx}(x_0) \le 0$*, then* $|f(x)| > cr^3$ *for all* $x \in J$*.*

  Let I_+ *denotes the case* $\frac{df}{dx}(x_0) > 0$ *and* I_- *for* $\frac{df}{dx}(x_0) < 0$*.*
- f **is of type** II *if we have the following.*
  - $\|f\|_{C^2} < C$ *and* $f(x) = 0$ *has at most two solutions which are contained in* $\frac{I}{2}$*;*
  - $\frac{df}{dx}(x) = 0$ *has one solution which is contained in* $\frac{I}{2}$*;*
  - $f(x) = 0$ *has one solution if and only if it is the* x *such that* $\frac{df}{dx}(x) = 0$*;*
  - *finally,* $\left|\frac{d^2 f}{dx^2}\right| > c$ *whenever* $|\frac{df}{dx}| < r^2$*.*
- f **is of type** III *if for* $l : I \to \mathbb{R}$ *as in (24)*

$$f = \tan^{-1}\left(l^2[\tan f_1(x)]\right) - \frac{\pi}{2} + f_2, \tag{25}$$

where either f_1 *is of type* I_+ *and* f_2 *of type* I_-*, or* f_1 *is of type* I_- *and* f_2 *of type* I_+*.*

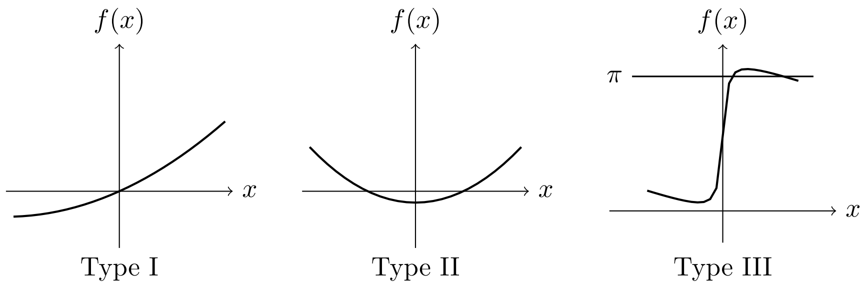

Figure 2. Graph of Type I, II and III functions

A simple case of type I function is when $\left|\frac{df}{dx}(x)\right| > r^2$ for all $x \in I$. The reason we define type III function as in (25) is due to (12) and (13).

The following lemma for f of type III actually plays the key role for the lower-bound estimate of the geometric properties of most contraction directions in resonance case. Without loss of generality, let f be as in (25) with f_1 be type I_+ and f_2 type I_- throughout this section. We may further assume that $f_1(0) = 0$ and $f_2(d) = 0$ with $0 \le d \le \frac{2}{3}r$. Let

$$X = \{x \in I : \mathbb{RP}^1 \ni |f(x)| = \inf_{y \in I} |f(y)|\}.$$

Then it is easy to see that X contains at most two points, say $X = \{x_1, x_2\}$ with $x_1 \le x_2$. Then we have the following lemma.

Lemma 6. *Let f be of type* III. *Let $r^2 \le \eta_j \le r^{-2}$, $0 \le j \le 4$. Then*

$$|x_1| < Cl^{-\frac{3}{4}},\ |x_2 - d| < Cl^{-\frac{3}{4}}. \tag{26}$$

In particular, if $f(x_1) = f(x_2) = 0$, then

$$0 < x_1 \le x_2 < d. \tag{27}$$

If $f(x_1) = f(x_2) \neq 0$, then

$$x_1 = x_2. \tag{28}$$

Then we consider the following two different cases:

$d < \frac{r}{3}$: then there exist two distinct points $x_3,\ x_4 \in B(x_1, \eta_0 l^{-1})$ such that

$$\frac{df}{dx}(x_j) = 0 \text{ for } j = 3, 4.$$

Here we set x_4 to be the point such that $x_1 \le x_4 \le x_2$. Then x_3 is a local minimum with

$$f(x_3) > \eta_1 l^{-1} - \pi. \tag{29}$$

See Figure 3 for positions of x_j, $j = 1, 2, 3, 4$. Moreover, it holds that

$$\left|\frac{d^2 f}{dx^2}(x)\right| > c \text{ whenever } \left|\frac{df}{dx}(x)\right| \le r^2 \text{ for } x \in B(X, \frac{r}{6}) \tag{30}$$

and $|f(x)| > cr^3$ for all $x \notin B(X, \frac{r}{6})$.

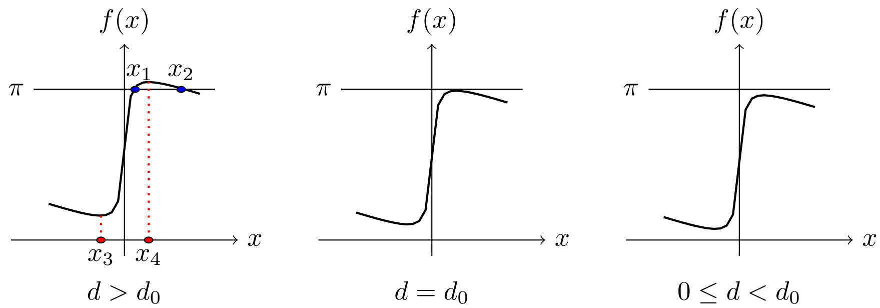

FIGURE 3. Bifurcation between $\mathcal{UH}$ and $\mathcal{NUH}$

$d \geq \frac{r}{3}$: then $\frac{df}{dx} = 0$ may have one or two solutions, among which the one between x_1 and x_2 always exists. In other words, there might exist another local minimum x_3 or not while x_4 always exists. In any case, it always holds in this case that

$$|f(x)| > cr^3, \ x \notin B(x_1, Cl^{-\frac{1}{4}}) \cup B(x_2, \frac{r}{4}); \ \|f - f_2\|_{C^2} < Cl^{-\frac{3}{2}} \ \text{on } B(x_2, \frac{r}{4}). \tag{31}$$

Finally, we have the following bifurcation as d varies. There is a $d_0 = \eta_2 l^{-1}$ such that:

- *if $d > d_0$, then $f(x) = 0$ has two solutions;*
- *if $d = d_0$, then $f(x) = 0$ has exactly one tangential solution. In other words, $x_1 = x_2 = x_4$ and $f(x_4) = 0$;*
- *if $0 \leq d < d_0$, then $f(x) \neq 0$ for all $x \in I$. Moreover, we have*

$$\min_{x \in I} |f(x)| = -\eta_3 l^{-1} + \eta_4 d.$$

See Figure 3 for the bifurcation procedure.

Remark 4. *Let $I' = B(x_0, r')$ with $r' < \frac{r}{2}$, then the restriction of type* I *f on I' is still of type* I *for I'. Let $I'' = B(0, r'')$ with $r'' \leq r$, then the restriction of type* II *f on I'' is still of type* II *for I'' if all solutions to $f(x) = 0$ and $\frac{df}{dx} = 0$ are contained in $\frac{I''}{2}$. Let $I''' \subset I$ be any connected interval containing $B(x_1, Cl^{-\frac{1}{4}})$. We also call the restriction of the type* III *f on I''' is again of type* III *for I'''. Note that if d is sufficiently close to d_0 and r''' is sufficiently small, then the restriction of type* III *function f to $B(x_1, r''')$ may become type* II*. However, for simplicity, we still call this restriction is of type* III.

Corollary 3. *Let $f : I \to \mathbb{RP}^1$ be of type* I, II *or* III. *Define*

$$X = \{x \in I : |f(x)| = \min_{y \in I} |f(y)|\} = \begin{cases} \{x_0\}, & \text{if } f \text{ is of type I} \\ \{x_1, x_2\}, & \text{if } f \text{ is of type II or III.} \end{cases}$$

In case f is of type III, *we further assume $d := |x_1 - x_2| < \frac{r}{3}$. Then for any $0 < r' < r$, we have that*

$$|f(x)| > cr'^3, \ \text{for all } x \notin B(X, r'). \tag{32}$$

For the case that f is of type III, *we have the same estimate (32) for $Cl^{-\frac{1}{4}} < r' < r$ if $d \geq \frac{r}{3}$.*

Proof. Let us consider the case that f is of type I first. If $r > \frac{r}{2}$, then (32) is obtained by definition. If $r' \leq \frac{r}{2}$, then for all $x \notin B(X, r')$, it holds that

$$|f(x)| > r^2|x - x_0| > r^2 r' > cr'^3.$$

If f is of type II, then clearly for some $d',\ d'' \geq 0$ satisfying $d' + d'' = |x - x_j|$, we have that for all $x \notin B(X, r')$, it holds that

$$|f(x)| > cd'^2 + r^2|d''| > c|x - x_j|^3 > cr'^3.$$

If f is of type III and $d < \frac{r}{3}$. Then (32) follows from Lemma 6 directly for $r' > \frac{r}{6}$. If $|x_1 - x_3| \leq r' \leq \frac{r}{6}$, we partition $B(X, r')$ as

$$B(X, r') = [x_1 - r', x_3] \cup [x_3, x_1] \cup J.$$

Then for the part J, the corresponding growth of f follows from (30) of Lemma 6 and the same argument for type II functions. For the part $[x_3, x_1]$, the issue is that $f(x)$ may increase too fast from near $-\pi$ to near 0. However, by (29) of Lemma 6, we have $|x_3 - x_1| \leq \eta_0 l^{-1}$ and

$$|f(x_3)| > \eta_1 l^{-1} - \pi > c|x_3 - x_1|^3 - \pi,$$

which is also a local minimal. Hence, we have the corresponding growth of $f(x)$. For the part $[x_1 - r', x_3]$, again by (30) of Lemma 6, we have corresponding growth as those for type II function .

If $0 < r' < |x_1 - x_3|$, we partition $B(X, r')$ as $B(X, r') = [x_1 - r', x_1] \cup J$. Then it can be treated similarly as the case $|x_1 - x_3| \leq r' \leq \frac{r}{6}$.

If $d \geq \frac{r}{3}$ and $Cl^{-\frac{1}{4}} < r' < r$, then it follows from Lemma 6 and the same argument as the one for type I functions. □

3. Getting Started

Consider the sequence $\{\lambda_n\}_{n \geq N}^{\infty}$ by $\log \lambda_n = \log \lambda_{n-1} - C\frac{\log q_n}{q_{n-1}} \log \lambda_{n-1}$ with $\lambda_N = \lambda$. It is easy to see that for all ε, there exists a N such that λ_n decreases to some λ_∞ with $\lambda_\infty > \lambda^{1-\varepsilon}$. For two finite sets $C_j \subset \mathbb{R}/\mathbb{Z}$, $j = 1, 2$, we define $|C_1 - C_2| = \min_{c_1 \in C_1, c_2 \in C_2} |c_1 - c_2|$.

For $n \geq 1$, let $s_n(x) = s[A_n(x)]$ and $u_n(x) = s[A_{-n}(x)]$. Note they may depend on the parameter t. These two functions will play the role of n-step stable and unstable directions. We call them n-step stable and unstable directions, respectively, since it is not very difficult to see that they converge to the stable and unstable directions in case one has a positive Lyapunov exponent, see, for example, the proof of [Z2, Theorem 1]. Obviously, we have that $u_1(x) = 0$ and

$$s_1^t(x) = \frac{\pi}{2} - \phi(x, t) = \frac{\pi}{2} - \cot^{-1}[t - v(x)] = \tan^{-1}[t - v(x)].$$

Let us define the following function, g_1^t, which is the difference between the first step stable and unstable direction:

$$g_1^t(x) := s_1(x) - u_1(x) = \tan^{-1}[t - v(x)]. \tag{33}$$

It is not difficult to see that we only need to consider

$$t \in \mathcal{I} := [\inf v - \frac{2}{\lambda_0}, \sup v + \frac{2}{\lambda_0}] \text{ for all } \lambda > \lambda_0,$$

see, for example, Lemma 11 of [Z1]. From now on, let us restrict t to this interval, and the dependence of g_1^t on $t \in \mathcal{I}$ will be left implicit.

By (33), it is a straightforward computation to see that for all $t \in \mathcal{I}$,

$$\|g_1\|_{C^2} \le C, \text{ and } c \le \left| \frac{d^m g_1}{dx^m} / \frac{d^m(t-v)}{dx^m} \right| \le C \tag{34}$$

for $m = 0, 1$ and all $x \in \mathbb{R}/\mathbb{Z}$. Thus, for all $t \in \mathcal{I}$, $\frac{dg_1}{dx} = 0$ have the same solutions as $\frac{dv}{dx} = 0$, which are z_1 and z_2. Moreover, it is a straightforward calculation to see that $|\frac{d^2 g_1}{dx^2}(z_j)| > c$ for all $t \in \mathcal{I}$. Clearly, there exists a $r > 0$ such that on $B(z, r)$, we have for all $t \in \mathcal{I}$,

$$g_1(x) = g_1(z_j) + \frac{d^2 g_1}{dx^2}(z_j)(x - z_j)^2 + o(x^2) \tag{35}$$

We also assume that, for the above r and for all $t \in \mathcal{I}$,

$$\left| \frac{dg_1}{dx}(x) \right| > cr, \text{ for all } x \notin B(z_j, r). \tag{36}$$

By choosing N sufficiently large, we may assume that $q_N^{-2\tau} \ll r$. Let

$$C_1 = \{y : |g_1(y)| = \min_{x \in \mathbb{R}/\mathbb{Z}} |g_1(x)|\} \text{ and } I_1 = B(C_1, \frac{1}{2q_N^{2\tau}}).$$

Clearly, C_1 contains at most two points. So we may let

$$C_1 = \{c_{1,1}, c_{1,2}\} \text{ and } I_{1,j} = B(c_{1,j}, \frac{1}{2q_N^{2\tau}}),\ j = 1, 2.$$

Note it is possible that $c_{1,1} = c_{1,2}$.

3.1. **Step 1.** For each $t \in \mathcal{I}$, we have one of the following cases.

$(1)_{\mathrm{I}}$ **(Type** I**)** I_1 consists of two disjoint connected intervals. In other words, $I_{1,1} \cap I_{1,2} = \varnothing$. Then, it is easy to see that $g_1(c_{1,j}) = 0$ and $|c_{1,1} - c_{1,2}| \ge \frac{1}{q_N^{2\tau}}$. Then by (35) and (36), it is straightforward that g_1 is of type I on $I_{1,1}$ and $I_{1,2}$. Furthermore, if g_1 is of type I_+ on $I_{1,1}$, then it is of type I_- on $I_{1,2}$, vice versa. Let $(1)_{\mathrm{I}}$ denotes this case.

$(1)_{\mathrm{II}}$ **(Type** II**)** I_1 consists of one connected interval. Hence $0 \le |c_{1,1} - c_{1,2}| < \frac{1}{q_N^{2\tau}}$. Clearly, g_1 is of type II on I_1 in this case. Let $(1)_{\mathrm{II}}$ denote this case.

Thus, by Corollary 3, we have that for each t and each $x \notin I_1$, $|g_1(x)| > cq_N^{-6\tau}$. Let $\eta'_N = \frac{C \log q_N}{\log \lambda_N} \ll \frac{C \log q_{N+1}}{q_N} \ll 1$. Fix a connected interval $I \subset \mathbb{R}/\mathbb{Z}$ and $\ell < \lambda_N^{\frac{1}{2}}$. Assume that $x + j\alpha \notin I_1$ for all $x \in I$ and for all $1 \le j \le \ell - 1$. Then by Lemma 1 and Lemma 4, we have that for $x \in I$, it holds that

$$\|s_\ell(x) - s_1(x)\|_{C^2},\ \|u_\ell(x + \ell\alpha) - u_1(x + \ell\alpha)\|_{C^2} \le C\lambda_N^{-\frac{3}{2}}, \tag{37}$$

$$\|A_\ell(x)\| \ge \lambda_N^{(1-\eta'_N)\ell},\ \left| \frac{d^m \|A_\ell(x)\|}{dx^m} \right| < \|A_\ell(x)\|^{1+m\eta'_N},\ m = 1, 2. \tag{38}$$

3.2. **From step 1 to step 2.** Define $q_N - 1 < r_1^{\pm} : I_1 \to \mathbb{Z}^+$ to be the smallest positive number j such that $j > q_N - 1$ and $T^{\pm j}x \in I_1$ for $x \in I_1$, respectively. Thus, there exist three possible cases:

- I_1 is in case $(1)_{\mathrm{II}}$. Thus $r_1^{\pm}(x)$ is the actual first return time by the *Diophantine* condition and we call this the non-resonance case;

- I_1 is in case $(1)_{\mathrm{I}}$ and $r_1^{\pm}(x)$ is the actual first return time for all $x \in I_1$. We also call this the non-resonance case and denote it by $(1)_{\mathrm{I},NR}$;
- I_1 is in case $(1)_{\mathrm{I}}$ and $r_1^{\pm}(x)$ is the second return time for some $x \in I_1$. We call this the resonance case and denote it by $(1)_{\mathrm{I},R}$.

Let $r_1^{\pm} = \min_{x\in I_1} r_1^{\pm}(x)$ and $r_1 = \min\{r_1^+, r_1^-\}$.

3.2.1. **Nonresonance case**. By the *Diophantine* condition, it is easy to see that in cases $(1)_{\mathrm{II}}$, it holds that $r_1 \geq q_N^2$. First note that in this case I_1 is a connected interval of length at most $\frac{2}{q_N^{2\tau}}$. Thus, assume $x+n\alpha \in I_1$ for some n and $x \in I_1$, then by the *Diophantine* condition $\gamma n^{-\tau+1} < \|n\alpha\|_{\mathbb{R}/\mathbb{Z}} < 2q_N^{-2\tau}$, where $\|\cdot\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance to the nearest integers. Hence $n > cq_N^{\frac{2\tau}{\tau-1}} > q_N^2$. In case $(1)_{\mathrm{I},NR}$, by definition, it holds that $r_1 \geq q_N$.

Applying (37) and (38) to the case $I = I_1$ and $\ell = r_1^{\pm}$, we then get

$$\|s_{r_1^+}(x) - s_1(x)\|_{C^2},\ \|u_{r_1^-}(x) - u_1(x)\|_{C^2} \leq C\lambda_N^{-\frac{3}{2}} \tag{39}$$

$$\|A_{\pm r_1^{\pm}}(x)\| \geq \lambda_N^{(1-\eta_N')r_1^{\pm}},\quad \left|\frac{d^m\|A_{\pm r_1^{\pm}}(x)\|}{dx^m}\right| < \|A_{\pm r_1^{\pm}}(x)\|^{1+m\eta_N'},\ m = 1,2. \tag{40}$$

Now we consider the function $s_{r_1^+}$, $u_{r_1^-} : I_1 \to \mathbb{RP}^1$ and define $g_2 = s_{r_1^+} - u_{r_1^-} : I_1 \to \mathbb{RP}^1$. Thus we get that as a functions on I_1,

$$\|g_2 - g_1\|_{C^2} < C\lambda_N^{-\frac{3}{2}}. \tag{41}$$

Combined with our assumption, it is clear that in cases $(1)_{\mathrm{I},NR}$ and $(1)_{\mathrm{II}}$, we must be in the following cases.

$(2)_{\mathrm{I}}$ **(Type** I**)** I_1 consists of two disjoint connected intervals. In other words, $I_{1,1} \cap I_{1,2} = \varnothing$. Furthermore g_2 is of type I on $I_{1,1}$ and $I_{1,2}$. In addition, if g_2 is of type I_+ on $I_{1,1}$, then it is of type I_- on $I_{1,2}$, vice versa. Let $(2)_{\mathrm{I}}$ denote this case.

$(2)_{\mathrm{II}}$ **(Type** II**)** I_1 consists of one connected interval. g_2 is of type II. Let $(2)_{\mathrm{II}}$ denote this case.

In both cases, if we let $C_2 = \{y : |g_2(y)| = \min_{x\in\mathbb{R}/\mathbb{Z}} |g_2(x)|\} = \{c_{2,1}, c_{2,2}\}$, then we clearly have $|c_{1,j} - c_{2,j}| < C\lambda_N^{-1}$.

3.2.2. **Resonance case**. Now we consider the case $(1)_{\mathrm{I},R}$, let $0 < k < q_N$ be the actual first return time for some x. Note in this case, we must have that $(I_{1,1} \pm k\alpha) \cap I_{1,2} \neq \varnothing$. Without loss of generality, assume $(I_{1,1} + k\alpha) \cap I_{1,2} \neq \varnothing$. Also, by the *Diophantine* condition and by the same argument as in nonresonance case, it is not difficult to see that once we have resonance, $(I_{1,1} + j\alpha) \cap I_{1,2} = \varnothing$ for all j such that $|j| < q_N^2$ and $|j| \neq k$ for all $x \in I_1$. Hence, in this case, no matter whether $r_1^{\pm}(x)$ is the first return or the second return time for $x \in \mathrm{I}_1$, we have $r_1^{\pm}(x) > q_N^2$ for all $x \in I_1$.

Now, for $x \in I_{1,1}$, let us consider

$$A_{r_1^+}(x) = A_{r_1^+ - k}(x + k\alpha)A_k(x) \text{ and } A_{-r_1^-}(x).$$

Clearly, for $A_{r_1^+-k}(x+k\alpha)$, $A_k(x)$ and $A_{-r_1^-}(x)$, (37) and (38) can be applied. Then we get the following facts. First we have

$$\|u_1(x)-u_{r_1^-}(x)\|_{C^2},\ \|s_k(x)-s_1(x)\|_{C^2},\ \|u_k(x+k\alpha)-u_1(x+k\alpha)\|_{C^2} \text{ and}$$

$$\|s_{r_1^+-k}(x+k\alpha)-s_1(x+k\alpha)\|_{C^2}<C\lambda_N^{-\frac{3}{2}};$$

Secondly, let $\nu=r_1^+-k$, $-r_1^-$ or k. Then

$$\left|\frac{d^m\|A_\nu(x)\|}{dx^m}\right|<\|A_\nu(x)\|^{1+m\eta'_N} \text{ for } m=1,2.$$

Finally, again for $\nu=r_1^+-k$, $-r_1^-$ or k,

$$\|A_\nu(x)\|>\lambda_{N+1}^{|\nu|}.$$

Now, let us focus on

$$A_{r_1^+}(x)=A_{r_1^+-k}(x+k\alpha)A_k(x).$$

Let $l_k=\|A_k(x)\|$ and $l'=\|A_{r_1^+-k}(x+k\alpha)\|$. Then by definition, we have

$$A_{r_1^+}(x)=R_{u_{r_1^+-k}(x+r_1^+\alpha)}\begin{pmatrix} l' & 0\\ 0 & l'^{-1}\end{pmatrix}R_{\frac{\pi}{2}-s_{r_1^+-k}(x+k\alpha)+u_k(x+k\alpha)}\begin{pmatrix} l_k & 0\\ 0 & l_k^{-1}\end{pmatrix}R_{\frac{\pi}{2}-s_k(x)}.$$

Clearly, we have

$$\lambda_N^{(1-\eta'_N)k}\le l_k<(C\lambda_N)^k<(C\lambda_N)^{q_N}\ll\lambda_N^{(1-\eta'_N)(q_N^2-k)}\le\lambda_N^{(1-\eta'_N)(r_1^+-k)}<l',$$

which implies that $\|A_{r_1^+}(x)\|>\lambda_{N+1}^{r_1^+}$. Moreover, by Lemma 5, it is easy to see that for sufficiently large λ, we have for all $x\in I_1$ and $m=1,2$,

$$\left|\frac{d^m\|A_{r_1^+}(x)\|}{dx^m}\right|<C\|A_{r_1^+}(x)\|^{1+m\eta'_N+Cm\frac{\eta'_N}{q_N}}, \tag{42}$$

where $\log\lambda_{N+1}>(1-\frac{C}{q_N})\log\lambda_N$.

Now, let us consider the function g_2. It is enough to consider $g_2: I_{1,1}\to\mathbb{RP}^1$. Clearly, $s_{r_1^+}(x)=s[B(x)]: I_1\to\mathbb{RP}^1$ for the following B.

$$B(x)=\begin{pmatrix} l' & 0\\ 0 & l'^{-1}\end{pmatrix}R_{\frac{\pi}{2}-s_{r_1^+-k}(x+k\alpha)+u_k(x+k\alpha)}\begin{pmatrix} l_k & 0\\ 0 & l_k^{-1}\end{pmatrix}R_{\frac{\pi}{2}-s_k(x)}.$$

Let $g'_{1,1}=s_k-u_{r_1^-}: I_{1,1}\to\mathbb{RP}^1$ and $g'_{1,2}=s_{r_1^+-k}-u_k: I_{1,2}\to\mathbb{RP}^1$. Thus we have

$$\left\|g'_{1,j}-g_1\right\|_{C^2,I_{1,j}}<C\lambda_N^{-\frac{3}{2}},\ j=1,2. \tag{43}$$

Thus $g'_{1,j}$ is of the same type as g_1 on $I_{1,j}$. Let $\bar{c}_{1,j}\in I_{1,j}$ be the zero of $g'_{1,j}$. Note as in the nonresonance case, we have

$$|\bar{c}_{1,j}-c_{1,j}|<C\lambda_N^{-\frac{3}{2}} \text{ for } j=1,2. \tag{44}$$

By the first estimate of (12) of Lemma 2, to do the C^2 estimate of g_2, it suffices to take

$$g_2(x)=\tan^{-1}\left(l_k^2\tan[g'_{1,2}(x+k\alpha)]\right)-\frac{\pi}{2}+g'_{1,1}(x),\ x\in I_{1,1}. \tag{45}$$

Similarly, for $x\in I_{1,2}$, by considering

$$A_{r_1^+}(x) \text{ and } A_{-r_1^-}(x)=A_{-r_1^-+k}(x-k\alpha)A_{-k}(x),$$

we may take

$$g_2(x) = \tan^{-1}\left(l_k^2 \tan[g'_{1,1}(x-k\alpha)]\right) - \frac{\pi}{2} + g'_{1,2}(x),\ x \in I_{1,2}. \tag{46}$$

Thus, by Lemma 6, g_2 are of type III on $I_{1,1} \cup (I_{1,2} - k\alpha)$ and $I_{1,2} \cup (I_{1,1} + k\alpha)$. Let $d := \bar{c}_{1,1} + k\alpha - \bar{c}_{1,2}$ and assume without loss of generality $d \geq 0$. Then, depending on the size of d, we get the following for step 2.

(**Nonresonance**) There is a d_1 close to $\frac{q_N^{-2\tau}}{2}$ such that for $d > d_1$, we have

$$\|g_2 - g_1\|_{C^2} < \|g_2 - g'_{1,j}\|_{C^2} + \|g'_{1,j} - g_1\|_{C^2} < Cl_k^{-\frac{3}{2}} + C\lambda_N^{-\frac{3}{2}} < C\lambda_N^{-\frac{3}{2}}.$$

by (31) and (43). Thus, this basically goes to the case $(2)_{\mathrm{I}}$. In other words, g_2 is of the same type as g_1 on $I_{1,j}$, $j = 1, 2$.

(**Weak resonance**) If $d \leq d_1$, then the drastic change part of graph of $\tan^{-1}(l_k^2 \tan[g'_{1,j_1}(x \pm k\alpha)])$ gradually enters I_{1,j_2}, where $j_1 \neq j_2 \in \{1, 2\}$. Hence g_2 is of type III by Remark 4. By Lemma 6, there is a $d_0 = \eta_1 l_k^{-1} \leq d_1$ with $q_N^{-4\tau} < \eta_1 < q_N^{-4\tau}$ such that for $d > d_0$, $g_2(x) = 0$ has two solutions. We say this comes from the case $(1)_{\mathrm{I},WR}$.

(**Strong resonance**) If $0 \leq d \leq d_0$, then $g_2(x) = 0$ has only one or even no solution. Here are the essential differences between the resonance case and the nonresonance case. The reason is that we have tangential intersections or even separation of r_1-step stable and unstable directions. The separation may also lead to $\mathcal{UH}$ (see Remark 7). We say this comes from the case $(1)_{\mathrm{I},SR}$.

We say the last two cases are in case $(2)_{\mathrm{III}}$ and come from case $(1)_{\mathrm{I},R}$. Now let us focus on the case $(2)_{\mathrm{III}}$ and consider without loss of generality g_2 on $I_{1,1}$, which is given by (45). By the (27) and (28) of Lemma 6, we may let $c_{2,1} \in I_{1,1}$ be the minimal point of g_2 that is closer to $\bar{c}_{1,1}$ than other ones. Thus, by (26), we have $|\bar{c}_{1,1} - c_{2,1}| < Cl_k^{-\frac{3}{4}} < C\lambda_N^{-\frac{3}{4}}$. This together with (44), clearly implies that

$$|c_{1,1} - c_{2,1}| < C\lambda_N^{-\frac{3}{4}}.$$

Similarly, by considering (46), we find a minimal point, $c_{2,2} \in I_{1,2}$, of g_2 such that

$$|c_{1,2} - c_{2,2}| < C\lambda_N^{-\frac{3}{4}}.$$

We say $c_{2,j}$ comes essentially from $c_{1,j}$ for $j = 1, 2$. It is possible that g_2 has one or two minimum points on $I_{1,j}$. Let $C'_2 = \{c'_{2,1}, c'_{2,2}\}$ with $c'_{2,1} \in I_{1,2}$ and $c'_{2,2} \in I_{1,1}$ be the possible extra minimum points of g_2. By (45), (46), (27) and (28), we have that $c'_{2,2}$ is closer to $\bar{c}_{1,2} - k\alpha$ than $c_{2,1}$, and $c'_{2,1}$ closer to $\bar{c}_{1,1} + k\alpha$ than $c_{2,2}$. On the other hand, $c'_{2,j}$ is essentially on the k-orbit of $c_{2,j}$, $j = 1, 2$, which is illustrated by the following lemma.

Lemma 7. *Assume we have either* $|g_2(c_{2,1})| < C\lambda_{N+1}^{-\frac{1}{10}r_1}$ *or* $|g_2(c_{2,2})| < C\lambda_{N+1}^{-\frac{1}{10}r_1}$, *then*

$$|c_{2,1} + k\alpha - c'_{2,1}|,\ |c_{2,2} - k\alpha - c'_{2,2}| < C\lambda_{N+1}^{-\frac{1}{30}r_1}.$$

To prove Lemma 7, we need the following lemma and corollary.

Lemma 8. *Let* $E = E_2E_1 \in \mathrm{SL}(2, \mathbb{R})$ *such that* $\|E_2\| \gg \|E_1\| \gg 1$. *Then we have*

$$|E_1^{-1} \cdot s(E_2) - s(E)| < C\|E\|^{-2},\ |s(E_2) - E_1 \cdot s(E)| < C\|E_2\|^{-2}. \tag{47}$$

Proof. Let $\hat{s} \in s$ be an unit vector. By polar decomposition, it suffices to consider the case

$$E_2 = \begin{pmatrix} e_2 & 0 \\ 0 & e_2^{-1} \end{pmatrix} R_\theta, \ E_1 = \begin{pmatrix} e_1 & 0 \\ 0 & e_1^{-1} \end{pmatrix}.$$

Then $s(E_2) = \frac{\pi}{2} - \theta$ and $\tan[E_1^{-1} \cdot s(E_2)] = e_1^2 \cot\theta$. Assume $\theta \neq 0$, otherwise it is trivial. Let $w \in E_1^{-1} \cdot s(E_2)$ be a unit vector. Then we have

$$\begin{aligned}
\|Ew\| &= \frac{1}{\sqrt{1 + e_1^4 \cot^2\theta}} \left\| E \begin{pmatrix} 1 \\ e_1^2 \cot\theta \end{pmatrix} \right\| = \frac{e_1 e_2^{-1} |\sin\theta|^{-1}}{\sqrt{1 + e_1^4 \cot^2\theta}} \\
&= \frac{1}{\sqrt{e_1^{-2} e_2^2 \sin^2\theta + e_1^2 e_2^2 \cot^2\theta}} = C\|E\|^{-1},
\end{aligned}$$

which clearly implies the first inequality of (47).

For the proof of the second inequality, let $w' \in E_1 \cdot s(E)$ be a unit vector. Then we have

$$\|E_2 w'\| = \frac{1}{\|E_1 \hat{s}(E)\|} \|E_2 E_1 \hat{s}(E)\| = \frac{1}{\|E\| \cdot \|E_1 \hat{s}(E)\|}.$$

Since $\|E\| \gg \|E_1\|$, by the first inequality of (47), we could replace $\hat{s}(E)$ by a unit vector, w'', in $E_1^{-1} \cdot s(E_2)$ in the above estimate. Thus we get

$$\begin{aligned}
\|E_2 w'\| &= \frac{C}{\|E\| \cdot \|E_1 w''\|} = \frac{C\sqrt{1 + e_1^4 \cot^2\theta}}{\|E\|} \left\| E_1 \begin{pmatrix} 1 \\ e_1^2 \cot\theta \end{pmatrix} \right\|^{-1} \\
&= \frac{C\sqrt{e_1^{-2} \sin^2\theta + e_1^2 \cos^2\theta}}{\sqrt{e_1^{-2} e_2^2 \sin^2\theta + e_1^2 e_2^2 \cot^2\theta}} = Ce_2^{-1} \\
&= C\|E_2\|^{-1},
\end{aligned}$$

which implies the second inequality of (47). □

The following corollary is an immediate consequence of Lemma 8.

Corollary 4. *Let E_1 and E_1 be as in Lemma 8. Let $E = E_1 \cdot E_2$, then we have*

$$|E_1 \cdot u(E_2) - u(E)| < C\|E\|^{-2}, \ |u(E_2) - E_1^{-1} \cdot u(E)| < C\|E_2\|^{-2}. \tag{48}$$

Proof. Note that $u(E) = s(E^{-1})$, $u(E_2) = s(E_2^{-1})$ and $E^{-1} = E_2^{-1} \cdot E_1^{-1}$. This reduces the proof to the case in Lemma 8. □

Remark 5. *It is probably interesting to point out that Lemma 8 and Corollary 4 are one more precise version of the first estimate of (12) and the second estimate of (13) in case of C^0 estimate.*

Now we are ready to prove Lemma 7.

Proof. (**Proof of Lemma 7**). For $x \in I_{1,1}$, let us consider the following

$$A_{r_1^+}(x + k\alpha) \cdot A_k(x) \cdot A_{r_1^-}(x - r_1^- \alpha).$$

By Lemma 8 and Corollary 4, we obtain

$$|s_{r_1^+}(x) - A_k(x)^{-1} \cdot s_{r_1^+}(x + k\alpha)|, \ |u_{r_1^-}(x) - A_k(x)^{-1} \cdot u_{r_1^-}(x + k\alpha)| < C\lambda_{N+1}^{-2r_1}.$$

Hence, we have

$$\left| g_2(x) - [A_k(x)^{-1} \cdot s_{r_1^+}(x + k\alpha) - A_k(x)^{-1} \cdot u_{r_1^-}(x + k\alpha)] \right| < C\lambda_{N+1}^{-2r_1}. \tag{49}$$

For $A \in \mathrm{SL}(2,\mathbb{R})$, consider the induced map $A : \mathbb{RP}^1 \to \mathbb{RP}^1 = \mathbb{R}/(\pi\mathbb{Z})$. For $\theta \in \mathbb{RP}^1$, let $\hat{\theta} \in \theta$ be a unit vector. Then a direct computation shows that $\frac{dA}{d\theta}(\theta) = \|A\hat{\theta}\|^{-2}$. Together with (49) and fact that $l_k = \|A_k(x)\| < \lambda^k \ll \lambda_{N+1}^{r_1}$, this clearly implies that

$$|g_2(x) - M \cdot g_2(x + k\alpha)| < C\lambda_{N+1}^{-2r_1},\ l_k^{-2} \leq M \leq l_k^2. \tag{50}$$

Note here g_2 is a function on $I_{1,1}$. Similarly, we get that, as a function on $I_{1,2}$, g_2 satisfies

$$\left| g_2(x) - [A_k(x) \cdot s_{r_1^+}(x - k\alpha) - A_k(x) \cdot u_{r_1^-}(x - k\alpha)] \right| < C\lambda_{N+1}^{-2r_1}.$$

Hence,

$$|g_2(x) - M \cdot g_2(x - k\alpha)| < C\lambda_{N+1}^{-2r_1},\ l_k^{-2} \leq M \leq l_k^2. \tag{51}$$

Without loss of generality, assume that $|g_2(c_{2,1})| < \lambda_{N+1}^{-\frac{1}{10}r_1}$. Then, by (51), we have

$$|g_2(c_{2,1} + k\alpha)| < Cl_k^2\lambda_{N+1}^{-\frac{1}{10}r_1} < C\lambda_{N+1}^{-\frac{1}{15}r_1}.$$

Clearly, $0 \leq |g_2(c'_{2,1})| = |g_2(c_{2,2})| \leq |g_2(c_{2,1} + k\alpha)| < C\lambda_{N+1}^{-\frac{1}{15}r_1}$. Now $c'_{2,1}$ is always the minimal point of g_2 on $I_{1,2}$ that is closer to $\bar{c}_{1,1} + k\alpha$, hence $c_{2,1} + k\alpha$, than other ones. Thus, by Lemma 6, we get that

$$|c_{2,1} + k\alpha - c'_{2,1}| < C\lambda_{N+1}^{-\frac{1}{30}r_1}.$$

Note that if $g_2(c_{2,1}) \neq 0$, then $c_{2,1} = c'_{2,2}$. If $g_2(c_{2,1}) = 0$, then $g_2(c'_{2,2}) = 0$. In any case, we get a similar relation between $c'_{2,2}$ and $c_{2,2}$, concluding the proof. $\square$

Finally, note that in case $(2)_{\mathrm{III}}$, we also have the following estimate by (21) and (22) of Lemma 5

$$\|g_2\|_{C^2} < Cl_k^5 < C\lambda^{5q_N}.$$

3.3. **The starting lemma.**

To conclude, at step 2, we have the following lemma.

Lemma 9 (The starting Lemma). *Let $g_1 = s_1 - u_1 = \tan^{-1}(t - v) : \mathbb{R}/\mathbb{Z} \to \mathbb{RP}^1$. Define*

$$C_1 = \{c_{1,1}, c_{1,2}\} = \{y : |g_1(y)| = \min_{x \in \mathbb{R}/\mathbb{Z}} |g_1(x)|\},\ I_{1,j} = \{x : |x - c_{1,j}| \leq \frac{1}{2q_N^{2\tau}}\}$$

and $I_1 = I_{1,1} \cup I_{1,2}$. Let $q_N - 1 < r_1^{\pm}(x) : I_1 \to \mathbb{Z}^+$ be the first return time after time $q_N - 1$, where r_1^+ is the forward return and r_1^- backward. Let $r_1^{\pm} = \min_{x \in \mathbb{R}/\mathbb{Z}} r_1^{\pm}(x)$, $r_1 = \min\{r_1^+, r_1^-\}$ and $r_0 = 1$. Let $g_2 = s_{r_1^+} - u_{r_1^-} : I_1 \to \mathbb{RP}^1$ and define $\eta_N = \frac{\log q_{N+1}}{q_N}$ and $\log \lambda_{N+1} > (1 - C\eta_N) \log \lambda_N$. Then there exists a set

$$C_2 = \{c_{2,1}, c_{2,2}\}, \quad c_{2,j} \in \{y : |g_2(y)| = \min_{x \in I_{1,j}} |g_2(x)|\} \textit{ such that}$$

$$|c_{1,j} - c_{2,j}| < C\lambda_N^{-\frac{3}{4}},\ j = 1, 2; \tag{52}$$

Moreover, for all $x \in I_1$ and $m = 1, 2$, it holds that

$$\|A_{\pm r_1^{\pm}}(x)\| > \lambda_{N+1}^{r_1^{\pm}},\ \frac{d^m \|A_{\pm r_1^{\pm}}(x)\|}{dx^m} < \|A_{\pm r_1^{\pm}}(x)\|^{1+m\eta_N}; \tag{53}$$

if I_1 consists of one connected interval, then g_2 is of type II *on I_1; if $I_{1,1} \cap I_{1,2} = \varnothing$, then in nonresonance case, g_2 is either of type* I*; in resonance case g_2 is of type* III *on each $I_{1,j}$. In other words, we have three different cases: $(2)_{\mathrm{I}}$, $(2)_{\mathrm{II}}$ and $(2)_{\mathrm{III}}$.*

In case $(2)_{\mathrm{I}}$ *and* $(2)_{\mathrm{II}}$, *there is no other minimal point of* g_2 *than those in* C_2, *and we have that*

$$\|g_2 - g_1\|_{C^2} \le C\lambda_N^{-\frac{3}{2}} \text{ and } \|g_2\|_{C^2} < C. \tag{54}$$

Moreover, in case $(2)_{\mathrm{I}}$, *if* g_2 *is of type* I_+ *on* $I_{1,1}$, *then it is of type* I_- *on* $I_{1,2}$, *vice versa.*

In case $(2)_{\mathrm{III}}$, *there are two more minimal points* $c'_{2,1}, c'_{2,2}$ *with* $c'_{2,1} \in I_{1,2}$, $c'_{2,2} \in I_{1,1}$ *such that* $g_2(c'_{2,1}) = g_2(c_{2,2})$ *and* $g_2(c'_{2,2}) = g_2(c_{2,1})$. *However,* $c_{2,j}$ *is always the one that comes essentially from* $c_{1,j}$ *for* $j = 1, 2$ *while it is possible that* $c'_{2,1} = c_{2,2}$ *and* $c'_{2,2} = c_{2,1}$.

Moreover, if $|g_2(c_{2,j})| < C\lambda_{N+1}^{-\frac{1}{10}r_1}$ *for* $j = 1$ *or* $j = 2$, *then we have*

$$|c_{2,1} + k\alpha - c'_{2,1}|,\ |c_{2,2} - k\alpha - c'_{2,2}| < C\lambda_{N+1}^{-\frac{1}{30}r_1} \tag{55}$$

with $1 \le k < q_N$. *In this case, we also have*

$$r_1 \ge q_N^2 \text{ and } \|g_2\|_{C^2} \le C\lambda^{5q_N}. \tag{56}$$

Remark 6. *In Section 3, instead of* $r_1^{\pm}(x)$, *sometimes, we used* $r_1^{\pm}$, r_1 *and even* $r_1^{\pm} \pm k$, $1 \le k \le q_N - 1$. *The difference between the usage of* $r_1^{\pm}(x)$, $r_1^{\pm}$, r_1 *and* $r_1 \pm k$ *are negligible. In fact, by the second estimate of (21) and (22) in Lemma 5, it is easy to see that these differences produces errors of order at most* $\lambda_{N+1}^{-\frac{3}{2}r_1}$, *which is clearly not important in all the necessary estimates. Similarly, in the following discussions, we will not distinguish the difference between* $r_i^{\pm}(x)$, $r_i^{\pm}$, r_i *and* $r_i \pm k$, *where* $i \ge 1$ *and* $1 \le k \le q_{N+i-1}$.

4. The Induction

Now we are ready to do the induction.

4.1. Statement of the induction theorem.

We formulate our induction as the following theorem.

Theorem 3 (Induction Theorem)**. Step (i+1).** *Let* $g_i = s_{r_{i-1}^+} - u_{r_{i-1}^-} : I_{i-1} \to \mathbb{RP}^1$. *Assume we have*

$$C_i = \{c_{i,1}, c_{i,2}\},\ c_{i,j} \in \{y : |g_i(y)| = \min_{x \in I_{i-1,j}} |g_i(x)|\},$$

$$I_{i,j} = \{x : |x - c_{i,j}| \le \frac{1}{2^i q_{N+i-1}^{2\tau}}\} \text{ and}$$

$$I_i = I_{i,1} \cup I_{i,2}.$$

Let $q_{N+i-1} < r_i^{\pm}(x) : I_i \to \mathbb{Z}^+$ *be the first return time after time* q_{N+i-1}. *Let* $r_i = \min\{r_i^+, r_i^-\}$ *with* $r_i^{\pm} = \min_{x \in I_i} r_i^{\pm}(x)$. *Let* $g_{i+1} = s_{r_i^+} - u_{r_i^-} : I_i \to \mathbb{RP}^1$ *and assume we have*

$$C_{i+1} = \{c_{i+1,1}, c_{i+1,2}\}, \quad c_{i+1,j} \in \{y : |g_{i+1}(y)| = \min_{x \in I_{i,j}} |g_{i+1}(x)|\} \text{ such that}$$

$$|c_{i,j} - c_{i+1,j}| < C\lambda_{N+i}^{-\frac{3}{4}r_{i-1}},\ j = 1, 2, \tag{57}$$

where $\log\lambda_{N+i} > (1-C\eta_{N+i-1})\log\lambda_{N+i-1}$ *with* $\eta_{N+i-1} = \frac{\log q_{N+i}}{q_{N+i-1}}$. *Assume for all* $x \in I_i$, $m = 1, 2$, *it holds that*

$$\|A_{\pm r_i^{\pm}}(x)\| > \lambda_{N+i}^{r_i^{\pm}}, \quad \frac{d^m\|A_{\pm r_i^{\pm}}(x)\|}{dx^m} < \|A_{\pm r_i^{\pm}}(x)\|^{1+mC\sum_{j=N}^{N+i-1}\eta_j}, \tag{58}$$

and $g_{i+1} : I_i \to \mathbb{RP}^1$ *is of type* I, II *or* III, *which are denoted as cases* $(i+1)_{\mathrm{I}}$, $(i+1)_{\mathrm{II}}$ *and* $(i+1)_{\mathrm{III}}$.

Assume that in case $(i+1)_{\mathrm{I}}$ *and* $(i+1)_{\mathrm{II}}$, *there is no other minimal point of* g_{i+1} *than those in* C_{i+1}, *and we have*

$$\|g_{i+1} - g_i\|_{C^2} \le C\lambda_{N+i-1}^{-\frac{3}{2}r_{i-1}} \text{ and } \|g_{i+1}\|_{C^2} \le C. \tag{59}$$

Moreover, in case $(i+1)_{\mathrm{I}}$, *assume that if* g_{i+1} *is of type* I_+ *on* $I_{i,1}$, *then it is of type* I_- *on* $I_{i,2}$, *vice versa.*

Assume that in case $(i+1)_{\mathrm{III}}$, *there are two more minimal points* $c'_{i+1,1}, c'_{i+1,2}$ *with* $c'_{i+1,1} \in I_{1,2}$, $c'_{i+1,2} \in I_{1,1}$ *such that* $g_{i+1}(c'_{i+1,1}) = g_{i+1}(c_{i+1,2})$ *and* $g_{i+1}(c'_{i+1,2}) = g_{i+1}(c_{i+1,1})$. *Assume* $c_{i+1,j}$ *is always the one comes essentially from* $c_{i,j}$ *while it is possible that* $c'_{i+1,1} = c_{i+1,2}$ *and* $c'_{i+1,2} = c_{i+1,1}$.

If $|g_{i+1}(c_{i+1,j})| < C\lambda_{N+i}^{-\frac{1}{10}r_i}$ *for* $j = 1$ *or* $j = 2$, *assume we have*

$$|c_{i+1,1} + k\alpha - c'_{i+1,1}|, \ |c_{i+1,2} - k\alpha - c'_{i+1,2}| < C\lambda_{N+i}^{-\frac{1}{30}r_i} \tag{60}$$

with $1 \le k < q_{N+i-1}$. *In this case, we also assume that*

$$r_i \ge q_{N+i-1}^2 \text{ and } \|g_{i+1}\|_{C^2} \le C\lambda^{5q_{N+i-1}}. \tag{61}$$

Then we have the following.

Step (i+2). *Let* $I_{i+1,j} = \{x : |x - C_{i+1,j}| \le \frac{1}{2^{i+1}q_{N+i}^{2\tau}}\}$ *and* $I_{i+1} = I_{i+1,1} \cup I_{i+2,j}$. *Let* $q_{N+i} < r_{i+1}^{\pm}(x) : I_{i+1} \to \mathbb{Z}^+$ *be the first return time after time* q_{N+i}. *Let* $r_{i+1}^{\pm} = \min_{x\in I_{i+1}} r_{i+1}^{\pm}(x)$ *and* $r_i = \min\{r_i^+, r_i^-\}$. *Let* $g_{i+2} = s_{r_{i+1}^+} - u_{r_{i+1}^-} : I_{i+1} \to \mathbb{RP}^1$. *Define* $\eta_{N+i} = \frac{\log q_{N+i+1}}{q_{N+i}}$ *and* $\log\lambda_{N+i+1} > (1-C\eta_{N+i})\log\lambda_{N+i}$. *Then there exists a set*

$$C_{i+2} = \{c_{i+2,1}, \ c_{i+2,2}\}, \quad c_{i+2,j} \in \{y : |g_{i+2}(y)| = \min_{x\in I_{i+1,j}} |g_{i+2}(x)|\}$$

such that

$$|c_{i+1,j} - c_{i+2,j}| < C\lambda_{N+i}^{-\frac{3}{4}r_i}, \ j = 1, 2; \tag{62}$$

and for all $x \in I_{i+1}$ *and* $m = 1, 2$, *it holds that*

$$\|A_{\pm r_{i+1}^{\pm}}(x)\| > \lambda_{N+i+1}^{r_{i+1}^{\pm}}, \quad \frac{d^m\|A_{\pm r_{i+1}^{\pm}}(x)\|}{dx^m} < \|A_{\pm r_{i+1}^{\pm}}(x)\|^{1+mC\sum_{j=N}^{N+i}\eta_j} \tag{63}$$

and $g_{i+2} : I_{i+1} \to \mathbb{RP}^1$ *is of types* I, II *or* III, *which are denoted as case* $(i+2)_{\mathrm{I}}$, $(i+2)_{\mathrm{II}}$ *and* $(i+2)_{\mathrm{III}}$. *In addition, in case* $(i+2)_{\mathrm{I}}$ *and* $(i+2)_{\mathrm{II}}$, *there is no other minimal point of* g_{i+2} *than those in* C_{i+2}, *and we have*

$$\|g_{i+2} - g_{i+1}\|_{C^2} \le C\lambda_{N+i}^{-\frac{3}{2}r_i} \text{ and } \|g_{i+2}\|_{C^2} \le C. \tag{64}$$

Moreover, in case $(i+2)_{\mathrm{I}}$, *if* g_{i+2} *is of type* I_+ *on* $I_{i+1,1}$, *then it is of type* I_- *on* $I_{i+1,2}$, *vice versa.*

In case $(i+2)_{\mathrm{III}}$, there are two more minimal points $c'_{i+2,1}, c'_{i+2,2}$ such that $g_{i+2}(c'_{i+2,1}) = g_{i+2}(c_{i+2,2})$ and $g_{i+2}(c'_{i+2,2}) = g_{i+2}(c_{i+2,1})$. However, $c_{i+2,j}$ is always the one that comes essentially from $c_{i+1,j}$ while it is possible that $c'_{i+2,1} = c_{i+2,2}$ and $c'_{i+2,2} = c_{i+2,1}$. Moreover, it holds that

- *if $|g_{i+1}(c_{i+1,j})| > C\lambda_{N+i}^{-\frac{1}{10}r_i}$, $j = 1, 2$, then so are $|g_{i+2}(c_{i+2,j})|$, $j = 1, 2$;*
- *if $|g_{i+2}(c_{i+2,j})| < C\lambda_{N+i+1}^{-\frac{1}{10}r_{i+1}}$ for $j = 1$ or $j = 2$, then*

$$|c_{i+2,1} + k\alpha - c'_{i+2,1}|,\ |c_{i+2,2} - k\alpha - c'_{i+2,2}| < C\lambda_{N+i+1}^{-\frac{1}{30}r_{i+1}} \tag{65}$$

  with $1 \le k < q_{N+i}$.

In this case, we also have

$$r_{i+1} \ge q_{N+i}^2 \text{ and } \|g_{i+2}\|_{C^2} \le C\lambda^{5q_{N+i}}. \tag{66}$$

4.2. **Proof of the induction.** To study the $g_{i+2} : I_{i+1} \to \mathbb{RP}^1$, we need to consider $g_{i+1} : I_{i+1} \to \mathbb{RP}^1$. Let us start with the case $(i+1)_{\mathrm{I}}$. Clearly, $g_{i+1} : I_{i+1} \to \mathbb{RP}^1$ inherited all the conditions of $g_{i+1} : I_i \to \mathbb{RP}^1$ that are assumed in the induction step $(i+1)_{\mathrm{I}}$. As in Section 2.1, we divide it into cases $(i+1)_{\mathrm{I},NR}$ and $(i+1)_{\mathrm{I},R}$. We first consider the nonresonance case $(i+1)_{\mathrm{I},NR}$.

Case $(\mathbf{i+1})_{\mathrm{I},\mathbf{NR}}$. In this case, $(I_{i+1,1} + l\alpha) \cap I_{i+1,2} = \varnothing$ for all $|l| \le q_{N+i} - 1$. Hence $r_{i+1}^{\pm}(x)$ is the actually first return time for each $x \in I_{i+1}$. If $r_{i+1}^{+} = r_i^{+}$, then $g_{i+2} = g_{i+1}$ and we have nothing to say. Otherwise, let $0 \le j_m \le r_{i+1}^{+}$, $0 \le m \le p$ be the times such that

$$j_{m+1} - j_m \ge q_{N+i-1},\ 0 \le m \le p-1 \text{ and } x + j_m\alpha \in I_i \setminus I_{i+1},\ 1 \le m \le p-1, \tag{67}$$

where $j_0 = 0$ and $j_p = r_{i+1}^{+}$. Now consider the sequence

$$A_{j_1}(x), A_{j_2}(x + j_1\alpha), \ldots, A_{j_p - j_{p-1}}(x + j_{p-1}\alpha). \tag{68}$$

Then, we have the following.

First, for $1 \le m \le p$, $\|A_{j_m - j_{m-1}}(x + j_{m-1}\alpha)\| \ge \lambda_{N+i}^{j_m - j_{m-1}} \ge \lambda_{N+i}^{q_{N+i-1}}$ and $\left|\frac{d^m \|A_{r_i}^{+}(x)\|}{dx^m}\right| < \|A_{r_i}^{+}(x)\|^{1+mC\sum_{j=N}^{N+i-2}\eta_j}$ by induction assumption (58).

Secondly, $|g_{i+1}(x + j_m\alpha)| \ge 2^{-2(i+1)}q_{N+i}^{-4\tau} > \lambda_{N+i}^{-\eta_{N+i-1}q_{N+i-1}}$ by induction assumption on the g_{i+1}, (67) and Corollary 3.

Finally, from (59), we have $\|g_{i+1}\|_{C^2} < C$.

Thus, the sequence (68) satisfies all the conditions in Lemma 4. Hence, we get

- $\|A_{r_{i+1}^{+}}(x)\| \ge \lambda_{N+i+1}^{r_{i+1}^{+}}$ with $\log\lambda_{N+i+1} = (1 - C\eta_{N+i-1})\log\lambda_{N+i}$;
- $\left|\frac{d^m \|A_{r_{i+1}}^{+}(x)\|}{dx^m}\right| < \|A_{r_{i+1}}^{+}(x)\|^{1+mC\sum_{j=N}^{N+i-1}\eta_j}$;
- $\|s_{r_{i+1}^{+}} - s_{r_i^{+}}\|_{C^2} \le C\lambda_{N+i}^{-\frac{3}{2}r_i}$ on I_{i+1}.

Similarly, we get the estimate for backward sequences. Thus, we also get that

$$\|g_{i+2} - g_{i+1}\|_{C^2} \le C\lambda_{N+i}^{-\frac{3}{2}r_i} \text{ on } I_{i+1},$$

which together with $\|g_{i+1}\|_{C^2} < C$ clearly implies (64). This also clearly implies the existence of two minimal points of $|g_{i+2}(x)|$, denoted by $c_{i+2,1}$ and $c_{i+2,2}$ such that

$$|c_{i+2,j} - c_{i+1,j}| < C\lambda_{N+i}^{-r_i} < C\lambda_{N+i}^{-\frac{3}{4}r_i}.$$

By (64) and the induction assumption for $g_{i+1} : I_{i+1} \to \mathbb{RP}^1$, $g_{i+2} : I_{i+1} \to \mathbb{RP}^1$ is clearly in case $(i+2)_{\mathrm{I}}$. Moreover, if g_{i+2} is of type I_+ on $I_{i+1,1}$, then it is of type I_- on $I_{i+1,2}$, vice versa.

Case $(\mathbf{i+1})_{\mathrm{I},\mathbf{R}}$. In this case, there exists a $r_i \le k \le q_{N+i} - 1$ such that

$$x + k\alpha \in I_{i+1,2} \text{ for some } x \in I_{i+1,1}.$$

By the *Diophantine* condition, $r_{i+1} \ge q_{N+i}^2$. Again, let $1 \le j_m \le r_{i+1}$, $0 \le m \le p$ be the times such that

$$j_{m+1} - j_m \ge q_{N+i-1},\ 0 \le m \le p-1 \text{ and } x + j_m\alpha \in I_i \setminus I_{i+1},\ 1 \le m \le p-1,$$

where we set $j_0 = 0$ and $j_p = r_{i+1}$. Let $l \in \{1, \dots, p-1\}$ be that $j_l < k < j_{l+1}$. Consider for $x \in I_{i+1,1}$,

$$A_{r_{i+1}^+ - k}(x + k\alpha) \cdot A_k(x) \text{ and } A_{-r_{i+1}^-}(x), \text{ where } A_k(x) = A_{k-j_l}(x + j_l\alpha) \cdots A_{j_1}(x)$$
$$\text{and } A_{r_{i+1}^+ - k}(x + k\alpha) = A_{j_p - j_{p-1}}(x + j_{p-1}) \cdots A_{j_{l+1}-k}(x + k\alpha).$$

Clearly, for $A_k(x)$, $A_{r_{i+1}^+ - k}(x + k\alpha)$ and $A_{-r_{i+1}^-}(x)$, the same argument of $(i+1)_{\mathrm{I},NR} \to (i+2)_{\mathrm{I}}$ is applicable. Thus, by Lemma 4, we get the following. Let $\nu = r_{i+1}^+ - k$, $-r_{i+1}^-$ or k. Then we have

First, $\|A_\nu(x)\| \ge \lambda_{N+i+1}^{|\nu|}$ with $\log \lambda_{N+i+1} = (1 - C\eta_{N+i-1}) \log \lambda_{N+i}$.

Secondly, $\left|\frac{d^m \|A_\nu(x)\|}{dx^m}\right| < \|A_\nu(x)\|^{1 + mC\sum_{j=N}^{N+i} \eta_j}$.

Finally, $\|s_k(x) - s_{r_i^+}(x)\|_{C^2}$, $\|u_k(x + k\alpha) - u_{r_i^-}(x + k\alpha)\|_{C^2}$ and

$$\|s_{r_{i+1}^+ - k}(x + k\alpha) - s_{r_i^+}(x + k\alpha)\|_{C^2} < C\lambda_{N+i}^{-\frac{3}{2}r_i}.$$

By the same reason, for $x \in I_{i+1,2}$, we could get a similar result when we consider

$$A_{r_{i+1}^+}(x) \text{ and } A_{-r_{i+1}^-}(x) = A_{-r_{i+1}^- + k}(x - k\alpha) \cdot A_{-k}(x).$$

Then everything follows from the same argument as of the case $(1)_{\mathrm{I},R}$ in Section 2.1. More concretely, let $g'_{i+1,1} = s_k - u_{r_{i+1}^-} : I_{i+1,1} \to \mathbb{RP}^1$, $g'_{i+1,2} = s_{r_{i+1}^+ - k} - u_k : I_{i+1,2} \to \mathbb{RP}^1$ and

$$\bar{c}_{i+1,1} \in I_{i+1,1}, \bar{c}_{i+1,2} \in I_{i+1,2} \text{ where } g'_{i+1,1}(\bar{c}_{i+1,1}) = 0,\ g'_{i+1,2}(\bar{c}_{i+1,2}) = 0.$$

Clearly, we have for $j = 1, 2$,

$$\left\|g'_{i+1,j} - g_{i+1}\right\|_{C^2, I_{i+1,j}} < C\lambda_{N+i}^{-\frac{3}{2}r_i} \text{ and } |\bar{c}_{i+1,j} - c_{i+1,j}| < C\lambda_{N+i}^{-r_i}. \tag{69}$$

This implies that $g'_{i+1,j}$ are of type I on $I_{i+1,j}$, one of which is of type I_+ and the other I_-. Let $d = \bar{c}_{i+1,1} + k\alpha - \bar{c}_{i+1,2}$. To estimate the geometric properties of g_{i+2}, it is sufficient to take

$$g_{i+2}(x) = \begin{cases} \tan^{-1}(l_k^2 \tan[g'_{i+1,2}(x + k\alpha)]) - \frac{\pi}{2} + g'_{i+1,1}(x), & x \in I_{i+1,1}, \\ \tan^{-1}(l_k^2 [\tan g'_{i+1,1}(x - k\alpha)]) - \frac{\pi}{2} + g'_{i+1,2}(x), & x \in I_{i+1,2}. \end{cases}$$

Then depending on the size of d, we have the following.

- $(i+1)_{\mathrm{I},R} \to (i+2)_{\mathrm{I}}$ if $d_0 \le d < \frac{2}{2^{i+1}q_{N+i}^{2\tau}}$ for some d_0 close to $\frac{1}{2^{i+1}q_{N+i}^{2\tau}}$.
- $(i+1)_{\mathrm{I},R} \to (i+2)_{\mathrm{III}}$ if $0 \le d < d_0$.

By (26) of Lemma 6, there is a minimal point, $c_{i+2,j}$, of $g_{i+2,j}$ that is always closer to $\bar{c}_{i+1,j}$ than other ones such that

$$|c_{i+2,j} - \bar{c}_{i+1,j}| < Cl_k^{-\frac{3}{4}} < C\lambda_{N+i}^{-\frac{3}{4}r_i} \tag{70}$$

for $j = 1, 2$. (69) and (70) clearly implies that

$$|c_{i+2,j} - c_{i+1,j}| < C\lambda_{N+i}^{-\frac{3}{4}r_i},\ j = 1, 2.$$

Finally, in this case, by (21), (22) of Lemma 5 and (59), (61) of Theorem 3, we get

$$\|g_{i+2}\|_{C^2} < Cl_k^{-4.5} < C\lambda^{5q_{N+i}}.$$

Case $(\mathbf{i+1})_{\mathrm{II}}$. In this case, let $d = c_{i+1,1} - c_{i+1,2}$. By definition, $|d| < \frac{2}{2^i q_{N+i-1}^{2\tau}}$. Then we have the following.

If $|d| \ge \frac{1}{2^{i+1}q_{N+i}^{2\tau}}$, $I_{i+1,1} \cap I_{i+1,2} = \varnothing$. Then, by induction assumption, it is not difficult to see that $g_{i+1} : I_{i+1} \to \mathbb{RP}^1$ are essentially as in the case $(i+1)_{\mathrm{I}}$. Thus we can do it as in the previous step and get $(i+2)_{\mathrm{I}}$ or $(i+2)_{\mathrm{III}}$.

If $d < \frac{1}{2^{i+1}q_{N+i}^{2\tau}}$, then I_{i+1} consists of one interval. Thus $r_{i+1}^{\pm} > q_{N+i}^2$ is the actually the first return. Moreover there exist two minimal points $c_{i+2,1}$ and $c_{i+2,2}$ of g_{i+2} such that

$$\|g_{i+2} - g_{i+1}\|_{C^2} < C\lambda_{N+i}^{-\frac{3}{2}r_i} \text{ and } |c_{i+2,j} - c_{i+1,j}| < C\lambda_{N+i}^{-\frac{3}{4}r_i}.$$

The estimate of $\|A_{\pm r_{i+1}^{\pm}}(x)\|$ together with its derivatives follows from the same argument as in case $(i+1)_{\mathrm{I}}$. So we will be in the case $(i+2)_{\mathrm{II}}$.

Case $(\mathbf{i+1})_{\mathrm{III}}$. In this case, other than $c_{i+1,j} \in I_{i,j}$, $j = 1, 2$, we have minimal points, $c'_{i+1,1} \in I_{i,2}$ and $c'_{i+1,2} \in I_{i,1}$, of g_{i+1} such that if $|g_{i+1}(c_{i+1,j})| < C\lambda_{N+i}^{-\frac{1}{10}r_i}$, then

$$|c'_{i+1,2} + k\alpha - c_{i+1,2}|,\ |c_{i+1,1} + k\alpha - c'_{i+1,1}| < \lambda_{N+i}^{-\frac{1}{30}r_i},\ 0 < k \le q_{N+i-1} - 1.$$

Let $d = c_{i+1,1} - c'_{i+1,2}$. Then there exists a d_0 close to $\frac{2}{2^{i+1}q_{N+i}^{2\tau}}$ such that the following holds true.

If $d > d_0$, then it is easy to see that $g_{i+1} : I_{i+1} \to \mathbb{RP}^1$ are essentially in the case $(i+1)_{\mathrm{I}}$. So we get the corresponding case in step $(i+2)$. Note in this case it is necessary that $g_{i+1}(c_{i+1,j}) = 0$.

If $0 < d \le d_0$, then we know that $(I_{i+1,1} + k\alpha) \cap I_{i+1,2} \ne \varnothing$ for $k < q_{N+i-1}$. Let us first assume d is not too small so that $|g_{i+1}(c_{i+1,j})| < C\lambda_{N+i}^{-\frac{1}{10}r_i}$. The other case will be dealt with later. In any case, by the *Diophantine* condition, we have that $r_{i+1}^{\pm} > q_{N+i}^2$. Hence, for any $x \in I_{i+1}$, whenever $x + l\alpha \in I_i$ we must have $x + l\alpha \notin I_{i+1}$ for any $1 \le |l| \le r_{i+1} - k$. Thus, by Lemma 4, it is not difficult to see that, as functions on I_{i+1}, $\|g_{i+2} - g_{i+1}\|_{C^2} < \lambda_{N+i}^{-\frac{3}{2}r_i}$. This implies the existence of two minimal points $c_{i+2,1}$ and $c_{i+2,2}$ of g_{i+2} such that

$$|c_{i+1,j} - c_{i+2,j}| < C\lambda_{N+i}^{-\frac{3}{4}r_i};$$

and together with (61), we have

$$\|g_{i+2}\|_{C^2} < C\lambda^{5q_{N+i-1}} < C\lambda^{5q_{N+i}}.$$

The estimate of $\|A_{\pm r_{i+1}^{\pm}}(x)\|$ together with its derivatives comes from the same argument as previous cases. Thus, we are in case $(i+2)_{\mathrm{III}}$.

In addition, from $(i+1)_{\mathrm{III}}$ to $(i+2)_{\mathrm{III}}$ that, we need to show that if $|g_{i+1}(c_{i+1,j})| > C\lambda_{N+i}^{-\frac{1}{10}r_i}$, $j = 1,2$, then so are $|g_{i+2}(c_{i+2,j})|$, $j = 1,2$. This is a consequence of Lemma 4. Indeed, fix arbitrary $x \in I_{i+1}$, let $0 \le j_m \le r_{i+1}^{+}$, $0 \le m \le p$ be as in case $(i+1)_{\mathrm{I},NR}$. Consider the following sequence

$$A_{j_1}(x), \ldots, A_{j_{m+1}-j_m}(x + j_m\alpha), \ldots, A_{j_p - j_{p-1}}(x + j_{p-1}\alpha)$$

and we clearly have

- $\|A_{j_{m+1}-j_m}(x + j_m\alpha)\| \ge \lambda_{N+i}^{r_i}$ and
- $|g_{i+1}(x + j_m\alpha)| = C|s_{j_{m+1}-j_m}(x + j_m\alpha) - u_{j_m - j_{m-1}}(x + j_m\alpha)| > C\lambda_{N+i}^{-\frac{1}{10}r_i}$.

Then, by Remark 3 following the proof of Lemma 4 in Section A.2 in case of C^0 estimate, we get as functions on I_{i+1},

$$\|s_{r_{i+1}^{+}} - s_{r_i^{+}}\| < C\lambda_{N+i}^{-\frac{3}{2}r_i}.$$

Similarly, we get $\|u_{r_{i+1}^{-}} - u_{r_i^{-}}\| < C\lambda_{N+i}^{-\frac{3}{2}r_i}$. Thus, we get $\|g_{i+2} - g_{i+1}\|_{I_{i+1}} < C\lambda_{N+i}^{-\frac{3}{2}r_i}$. This clearly implies that

$$\|g_{i+2}(x)\| > C\lambda_{N+i}^{-\frac{1}{10}r_i} \text{ for all } x \in I_{i+1}.$$

Finally, whenever we are in case $(i+2)_{\mathrm{III}}$, let $c'_{i+2,1}$ and $c'_{i+2,2}$ be the two extra minimal points of g_{i+2}. We need to show that if $|g_{i+2}(c_{i+2,j})| < C\lambda_{N+i+1}^{-\frac{1}{10}r_{i+1}}$ for $j = 1$ or $j = 2$, then

$$|c_{i+2,1} + k\alpha - c'_{i+2,1}|,\ |c_{i+2,2} - k\alpha - c'_{i+2,2}| < C\lambda_{N+i+1}^{-\frac{1}{30}r_{i+1}}$$

for some $1 \le k < q_{N+i}$. As the proof of Lemma 7, this is just another an application of Lemma 8 and Corollary 4 to the following

$$A_{r_{i+1}^{+}}(x + k\alpha) \cdot A_k(x) \cdot A_{r_{i+1}^{-}}(x - r_{i+1}^{-}\alpha).$$

Then everything follows the same proof of Lemma 7. This concludes the proof of our induction.

Remark 7. *This above inequality essentially shows that* $|g_{n+1}(x)| > C\lambda_{N+i}^{-\frac{1}{10}r_i}$ *for all* $n \ge i$ *and for all* $x \in I_n$. *Since it is clear that the* $\|g_{n+1} - g_n\|_{I_n} < C\lambda_{N+n-1}^{-\frac{3}{2}r_{n-1}}$, $n \ge i+1$, *we have*

$$C\sum_{n=i+1}^{\infty} \lambda_{N+n-1}^{-\frac{3}{2}r_{n-1}} \le C\lambda_{N+i}^{-\frac{3}{2}r_i}.$$

Thus, we will not need to worry about these parameters in the future since Lemma 4 will be applicable for all $n \ge i$. *In fact, the reason we have this property is that, by Lemma 11 of* [Z1], *it is not difficult to see that* $(\alpha, A) \in \mathcal{UH}$ *for these parameters.*

Hence there must exist a $N_0 \in \mathbb{Z}^+$ and $\gamma_0 > 0$ such that $|s_n(x) - u_n(x)| > \gamma_0$ for all $n \geq N_0$ and for all $x \in \mathbb{R}/\mathbb{Z}$. In particular, whenever we have $|g_{i+1}(x)| > C\lambda_{N+i}^{-\frac{1}{10}r_i}$ for all $x \in I_i$ for some i, then the induction can basically stop at this step i. Thus, in our induction, we essentially only need to deal with the case $|g_{i+1}(x)| < C\lambda_{N+i}^{-\frac{1}{10}r_i}$ for some $x \in I_i$.

Remark 8. *As we see in the proof of induction, the case $(i+2)_{\mathrm{III}}$ may come from $(i+1)_{\mathrm{III}}$. In fact, it is possible that the strong resonance occurs at step $i_0 \ll i$ so that $(i+1)_{\mathrm{III}}$ comes from $(i_0)_{\mathrm{III}}$. Then, as pointed out in Remark 4, $g_{i+1} : I_{i,j} \to \mathbb{RP}^1$, $j = 1, 2$, may become type* II *and we may have $c_{i+1,j_1} = c'_{i+1,j_2}$, $j_1 \neq j_2 \in \{1, 2\}$ and $\lambda_N^{-k} > \frac{1}{2^{i+1}q_{N+i}^{2\tau}}$. Thus, at step $(i+2)$, one may worry that if there is resonance between two type* II *functions, which may lead to some new bifurcation that is not included in our induction. However, the last part of Theorem 3 and Remark 7 imply that either we have $|g_n(c_{n,j})| > C\lambda_{N+i}^{-\frac{1}{10}r_i}$ for all $n \geq i+1$, hence, we do not need to worry about this case in the future. Or we have*

$$|c'_{i+1,1} - k\alpha + c_{i+1,1}|,\ |c'_{i+1,2} + k\alpha - c_{i+1,2}| < C\lambda_{N+i}^{-\frac{1}{30}r_i} \ll |I_{i+1,j}|,$$

which implies that there will be no new type of bifurcation within time q_{N+i}.

5. Positivity and Continuity: Proof of Theorem 1 and 2

In this section we will first deduce a key lemma from Theorem 3. Then we will see that Theorem 1 and a version of large deviation theorem (LDT) are direct consequences of the key lemma. Finally, combine with the so-called Avalanche Principle, the LDT will imply the weak Hölder continuity of Lyapunov exponents and integrated density of states (IDS) with respect to the energy.

5.1. **A key lemma.** Consider $E \in [\lambda \inf v - C, \lambda \sup v + C]$. Thus for any $\varepsilon > 0$, we have for large λ

$$\log \|A^{(E-\lambda v)}(x)\| < (1 + \frac{\varepsilon}{2}) \log \lambda \tag{71}$$

for all E in question and all $x \in \mathbb{R}/\mathbb{Z}$.

Let $\frac{p_s}{q_s}$ be the sth continued fraction approximants of α, which is from Theorem 1 or 2. Note we have by *Diophantine* condition

$$q_{s+1} < c q_s^{\tau-1},\ s \in \mathbb{Z}_+. \tag{72}$$

We need the following proposition which is a standard result for *Diophantine* translation on torus. For a simple proof, see [ADZ, Lemma 6].

Proposition 1. *For arbitrary fixed* Diophantine *frequency, there is a polynomial $P = P(\alpha) \in \mathbb{R}[X]$ such that, for any interval $I \subset \mathbb{R}/\mathbb{Z}$ and for each $x \in \mathbb{R}/\mathbb{Z}$,*

$$x + \ell\alpha \in I \text{ for some } 0 < \ell < P(|I|^{-1}).$$

Consider $I_n = I_{n,1} \cup I_{n,2}$ with $I_{n,j} = B(c_{n,j}, 2^{-n} q_{N+n-1}^{-2\tau})$, $n \geq 1$, $j = 1, 2$, which is from the induction Theorem 3. Recall by Theorem 3, there are associated times $r_n^{\pm}(x)$, $r_n = \min_{x \in I_n} r_n^{\pm}(x) \geq q_{N+n-1}$, such that the following holds.

- $\|A_\nu(x)\| \geq \lambda_{N+n}^{|\nu|}$ for each $x \in I_n$, $\nu = r_n^{\pm}(x)$ or r_n.
- $|s_{r_n^+(x)}(x) - s_{r_n}(x)|$, $|u_{r_n^-(x)}(x) - u_{r_n}(x)| < \lambda_{N+n}^{-\frac{3}{2}r_n}$ for each $x \in I_n$, see, e.g. Remark 6.

- $g_{n+1} : I_n \to \mathbb{RP}^1$ is of type I, II or III, where $g_{n+1}(x) = s_{r_n}(x) - u_{r_n}(x)$.

Recall $C_{n+1} = \{c_{n+1,1}, c_{n+1,2}\}$ is a set of minimal points of g_{n+1}, which is also from Theorem 3. Without loss of generality, let $R_n = P(|I_n|^{-1}) = q_{N+n-1}^C$ for some constant C depending only on α. By definition, $R_n \geq \max_{x\in I_n}\{r_n^{\pm}(x)\} \geq r_n$. Note that by (72)

$$R_n^{2\tau} = q_{N+n-1}^{2\tau C} > q_{N+n}^{2C} = R_{n+1}^2 > R_n^2. \tag{73}$$

Let $\mathcal{I}_n = B(C_{n+1}, e^{-\delta q_{N+n-1}}) = B(c_{n+1,1}, e^{-\delta q_{N+n-1}}) \bigcup B(c_{n+1,2}, e^{-\delta q_{N+n-1}})$,

$$\mathcal{D}_n = \bigcup_{\ell=1}^{R_n^{2\tau}} (\mathcal{I}_n - \ell\alpha) \text{ and } \mathcal{B}_n = \bigcup_{\ell=1}^{R_n} (\mathcal{D}_n - \ell\alpha). \tag{74}$$

Clearly, if we set $\sigma = \frac{1}{2\tau C}$, then for sufficiently large n and for each $i \in [R_n^2, R_n^{2\tau}]$

$$\mathrm{Leb}(\mathcal{B}_n) \leq 4q_{N+n-1}^{(2\tau+1)C} e^{-\delta q_{N+n-1}} \leq e^{-\frac{1}{2}\delta q_{N+n-1}} \leq e^{-\frac{1}{2}\delta i^{\sigma}}. \tag{75}$$

Then we have the following key lemma.

Lemma 10. *Let be v, α as in Theorem 1. Then for all $\varepsilon > 0$, there exists $\lambda_2 = \lambda_2(v, \alpha, \varepsilon)$, $\delta = \delta(\varepsilon) > 0$, and $n_0 = n_0(\alpha, \varepsilon) \in \mathbb{Z}^+$, such that for each $\lambda > \lambda_2$, each $i \in [R_n^2, R_n^{2\tau}]$ with $n \geq n_0$, and each $x \notin \mathcal{B}_n$, it holds that*

$$\|A_i(x)\| \geq \lambda^{(1-\frac{\varepsilon}{2})i}. \tag{76}$$

Proof. Consider a $x \notin \mathcal{B}_n$. Then by the choice of R_n, we have that $x_\ell = x + \ell\alpha \in I_n$ for some $0 < \ell < R_n$. Then write $A_i(x)$ as

$$A_i(x) = A_{i-\ell}(x_\ell) \cdot A_\ell(x).$$

Now let us focus on $A_{i-\ell}(x_\ell)$ with $x_\ell \in I_n$. Let $j_0 = 0$ and $j_m \leq i - \ell$, $1 \leq m \leq p$, be the all the possible times such that

$$j_m - j_{m-1} = r_n^+(x_\ell + j_{m-1}\alpha) \geq q_{N+n-1}.$$

Note $i - \ell - j_p \leq R_n$.

By our choice of x and the definition of $r_n^+ : I_n \to \mathbb{Z}^+$, it holds that

$$x_\ell + j_m\alpha \in I_n \setminus \mathcal{I}_n,\ 0 \leq \ell \leq p.$$

In other words, we always have $|x + j_m\alpha - C_{n+1}| > e^{-\delta q_{N+n-1}}$. Thus for a suitable choice of δ, we are able to apply Lemma 4 to concatenate the sequence

$$A_{j_1}(x_\ell), \ldots, A_{j_{m+1}-j_m}(x_\ell + j_m\alpha), \ldots, A_{j_p-j_{p-1}}(x_\ell + j_{p-1}\alpha). \tag{77}$$

We may choose large λ and N such that $C\sum_{n\geq 0} \eta_{N+n} < \frac{\varepsilon}{4}$. Hence, $\lambda_{N+n} > \lambda^{1-\frac{\varepsilon}{4}}$ for all $n \geq 0$, where $\{\lambda_{N+n}\}_{n\geq 0}$ with $\lambda_N = \lambda$ are from Theorem 3. Then the analysis above implies the following facts.

First, for $1 \leq m \leq p$, it holds

$$\|A_{j_m-j_{m-1}}(x_\ell + j_{m-1}\alpha)\| \geq \lambda_{N+n}^{r_n} \geq \lambda_{N+n}^{q_{N+n-1}} \geq \lambda^{(1-\frac{\varepsilon}{4})q_{N+n-1}}. \tag{78}$$

Secondly, for $1 \leq m \leq p$,

$$|x + j_m\alpha - C_{n+1}| > e^{-\delta q_{N+n-1}}. \tag{79}$$

Then we claim for $1 \leq m \leq p-1$, it holds that

$$|s_{r_n}(x_\ell + j_m\alpha) - u_{r_n}(x_\ell + j_m\alpha)| = |g_{n+1}(x_\ell + j_m\alpha)| > ce^{-3\delta q_{N+n-1}}. \tag{80}$$

In case $(n+1)_{\mathrm{I}}$ and $(n+1)_{\mathrm{II}}$, (80) follows directly from (79) since there is no extra minimal points of $g_{n+1} : I_n \to \mathbb{RP}^1$. In the resonance case $(n+1)_{\mathrm{III}}$, we need to worry about the possibility that

$$|x_\ell + j_m\alpha - c'_{n+1,2}| < ce^{-\delta q_{N+n-1}}. \tag{81}$$

However, either $|c'_{n+1,2} - c_{n+1,1}| \ll e^{-\delta q_{N+n-1}}$ so that (81) contradicts with (79). Or by (65), for some $k < q_{N+n-1}$,

$$|c'_{n+1,2} + k\alpha - c_{n+1,2}| < \lambda_{N+n}^{-\frac{1}{30}r_n} \ll e^{-\delta q_{N+n-1}},$$

which together with (81) implies

$$|x_\ell + (j_m + k)\alpha - c_{n+1,2}| \le e^{-\delta q_{N+n-1}}.$$

This again contradicts with our choice of $x \notin \mathcal{B}_n$ since $j_m + k < j_p \le i - \ell$ for $1 \le m \le p-1$.

Finally, by choosing n large, we can of course assume that for each $1 \le m \le p$,

$$p < R_n^{2\tau} = q_{N+n-1}^{2\tau C} < \lambda^{\frac{1}{2}(1-\frac{\varepsilon}{4})q_{N+n-1}} < \|A_{j_m - j_{m-1}}(x_\ell + j_{m-1}\alpha)\|^{\frac{1}{2}}. \tag{82}$$

Now by suitable choice of δ, e.g. $\delta < c\varepsilon$, we may assume

$$Ce^{3\delta q_{N+n-1}} < \lambda^{\frac{\varepsilon}{12}(1-\frac{\varepsilon}{4})q_{N+n-1}}.$$

Thus, by (78), (80) and (82), we can apply Lemma 4 and Remark 3 to the sequence (77) and obtain

$$\begin{aligned}
\|A_{j_p}(x_\ell)\| &\ge \left(\prod_{m=1}^{p} \|A_{j_m - j_{m-1}}(x_\ell + j_{m-1}\alpha)\|\right)^{1-\frac{\varepsilon}{12}} \\
&\ge \lambda^{(1-\frac{\varepsilon}{4})(1-\frac{\varepsilon}{12})j_p} \\
&\ge \lambda^{(1-\frac{\varepsilon}{3})j_p}.
\end{aligned}$$

Notice we have $\ell < R_n$, $i - \ell - j_p < R_n$ and $i \ge R_n^2$. Hence,

$$j_p \ge i - \ell - R_n > i - 2R_n.$$

By choosing n large, we also get $\frac{4R_n}{i} < \frac{4}{R_n} < \frac{\varepsilon}{6}$. Thus, it holds that

$$\begin{aligned}
\|A_i(x)\| &= \|A_{i-\ell-j_p}(x_\ell + j_p\alpha) \cdot A_{j_p}(x_\ell) \cdot A_\ell(x)\| \\
&\ge \|A_{i-\ell-j_p}(x_\ell + j_p\alpha)\|^{-1} \cdot \|A_{j_p}(x_\ell)\| \cdot \|A_\ell(x)\|^{-1} \\
&\ge (C\lambda)^{-2R_n} \cdot \lambda^{(1-\frac{\varepsilon}{3})j_p} \\
&\ge (C\lambda)^{-2R_n} \cdot \lambda^{(1-\frac{\varepsilon}{3})(i-2R_n)} \\
&\ge \lambda^{(1-\frac{\varepsilon}{3})i} \cdot \lambda^{-4R_n} \\
&\ge \lambda^{(1-\frac{\varepsilon}{2})i},
\end{aligned}$$

concluding the proof of Lemma 10. □

5.2. **Uniform positivity: proof of Theorem 1.** Theorem 1 is an easy consequence of Lemma 10. For the given $\varepsilon > 0$, let $\lambda_0(v, \alpha, \varepsilon) = \lambda_2(v, \alpha, \varepsilon)$, which is from Lemma 10. By (73), we clearly have

$$\bigcup_{n \geq n_0} [R_n^2, R_n^{2\tau}] = [R_{n_0}, +\infty). \tag{83}$$

Set $i_0 = R_{n_0}^2$. By Lemma 10 and (83), we have for each $\lambda > \lambda_0$ and each $i \geq i_0$,

$$\|A_i(x)\| > \lambda^{(1-\frac{\varepsilon}{2})i} \text{ for each } x \notin \bigcup_{n \geq n_0} \mathcal{B}_n. \tag{84}$$

Clearly, by (75), for large n_0, it holds that

$$\mathrm{Leb}\left(\bigcup_{n \geq n_0} \mathcal{B}_n\right) < \sum_{n \geq n_0} e^{-\frac{\delta}{2} q_{N+n-1}} < C e^{-\frac{\delta}{2} q_{N+n_0-1}} \ll 1. \tag{85}$$

By Kingman's Subadditive Ergodic Theorem, (84) and (85) clearly imply

$$L(E, \lambda) > (1 - \frac{\varepsilon}{2}) \log \lambda > (1 - \varepsilon) \log \lambda \tag{86}$$

for all $(E, \lambda) \in \mathbb{R} \times [\lambda_0, \infty)$, concluding the proof of Theorem 1.

5.3. **LDT and continuity: proof of Theorem 2.** The following version of LDT is an easy consequence of Lemma 10 and Theorem 1.

Theorem 4. *Let v, α be as in Theorem 1. Then for each $\varepsilon > 0$, there exists $\lambda_1 = \lambda_1(v, \alpha, \varepsilon)$, $\delta = \delta(\varepsilon) > 0$, $0 < \sigma = \sigma(\alpha) < 1$, and $i_0 = i_0(\alpha, \varepsilon) \in \mathbb{Z}^+$, such that for each $\lambda > \lambda_1$ and each $i \geq i_0$, it holds that*

$$\mathrm{Leb}\{x \in \mathbb{R}/\mathbb{Z} | \left|\frac{1}{i} \log \|A_i(x)\| - L(E)\right| > \varepsilon \log \lambda\} < e^{-\frac{1}{2}\delta i^\sigma}. \tag{87}$$

Proof. Let $\lambda_1(v, \alpha, \varepsilon) = \lambda_2(v, \alpha, \varepsilon)$, which is from Lemma 10. Then we have (71), (76) and (86). Let $\delta = \delta(\varepsilon)$ and $n_0 = n_0(\alpha, \varepsilon) \in \mathbb{Z}^+$ be as in Lemma 10. Let $\sigma = \sigma(\alpha)$ be as in (75).

By (83), for each $i \geq i_0 = R_{n_0}^2$, we may assume $i \in [R_n^2, R_n^{2\tau}]$ for a some $n \geq n_0$. Then for each $x \in (\mathbb{R}/\mathbb{Z}) \setminus \mathcal{B}_n$, we have

$$(1 - \frac{\varepsilon}{2}) \log \lambda - (1 + \frac{\varepsilon}{2}) \log \lambda \leq \frac{1}{i} \log \|A_i(x)\| - L(E) \leq (1 + \frac{\varepsilon}{2}) \log \lambda - (1 - \frac{\varepsilon}{2}) \log \lambda.$$

This clearly implies that

$$\left|\frac{1}{i} \log \|A_i(x)\| - L(E)\right| \leq \varepsilon \log \lambda. \tag{88}$$

Thus, $\left|\frac{1}{i} \log \|A_i(x)\| - L(E)\right| > \varepsilon \log \lambda$ implies that $x \in \mathcal{B}_n$, which together with (75) clearly imply our Theorem 4. $\square$

Once we have Theorem 4, the weak Hölder continuity of Lyapunov exponents and of IDS follows essentailly from the approach that is developed by Goldstein-Schlag [GoSc]. For the purpose of completeness, we include the proof here. First, we state the following lemma from [GoSc], which is called "Avalanche Principle".

Lemma 11. *Let $E^{(1)}, \dots, E^{(n)}$ be a finite sequence in $\mathrm{SL}(2,\mathbb{R})$ satisfying the following condition*

$$\min_{1\le j\le n} \|E^{(j)}\| \ge \mu \ge n, \tag{89}$$

$$\max_{1\le j<n} \left|\log \|E^{(j+1)}\| + \log \|E^{(j)} - \log \|E^{(j+1)}E^{(j)}\|\right| < \frac{1}{2}\log\mu \tag{90}$$

then

$$\left|\log \|E^{(n)}\dots E^{(1)}\| + \sum_{2}^{n-1} \log \|E^{(j)}\| - \sum_{1}^{n-1} \log \|E^{(j+1)}E^{(j)}\|\right| \le C\frac{n}{\mu}. \tag{91}$$

See [GoSc, Proposition 2.2] for a proof of Lemma 11. It is probably interesting to point out that there is some intrinsic relation between the C^0 version of Lemma 4 (see Remark 3) and Lemma 11. In fact, though taking different point of views, both of them deal with long finite concatenation of $\mathrm{SL}(2,\mathbb{R})$ matrices, and the conditions assumed in both lemmas are similar. Let

$$L_n(E) = \frac{1}{n}\int_{x\in\mathbb{R}/\mathbb{Z}} \log \|A_n^{(E-v)}(x)\| dx.$$

Then, combining Lemma 11 and Theorem 4, we get the following Lemma. Let us fix some ε small. For instance, $\varepsilon = \frac{1}{100}$ will be enough for our purpose. Then we may also replace δ in Theorem 4 by c since we may set $\delta = c\varepsilon = \frac{c}{100}$.

Lemma 12. *Let v, α, λ be as in Theorem 4, then for all $n \in \mathbb{Z}_+$ large enough and for all $E \in [\lambda \inf v - 2, \lambda \sup v + 2]$,*

$$|L(E) + L_n(E) - 2L_{2n}(E)| < Ce^{-cn^{\sigma}}, \tag{92}$$

where $c, C > 0$ depend on v, α, λ, and $0 < \sigma < 1$ on α.

Proof. Apply Lemma 11 to

$$n \le e^{\frac{\delta}{10}\ell^{\sigma}},\ E^{(j)} = A_\ell(x + (j-1)\ell\alpha).$$

From (87), with $i = \ell$, there is an exceptional set $\Omega \subset \mathbb{R}/\mathbb{Z}$ with

$$\mathrm{Leb}(\Omega) < e^{-\frac{1}{4}\delta\ell^{\sigma}}$$

such that if $x \notin \Omega$ and $j = 0, \dots, n$ then

$$(1-\varepsilon)\log\lambda < \frac{1}{\ell}\log\|A_\ell(x + j\ell\alpha)\|,\ \frac{1}{2\ell}\log\|A_{2\ell}(x + j\ell\alpha)\| < \log\lambda.$$

Hence for $x \notin \Omega$

$$\left|\log \|E^{(j+1)}\| + \log \|E^{(j)} - \log \|E^{(j+1)}E^{(j)}\|\right| \le 2\varepsilon\log\lambda < \frac{1-\varepsilon}{2}\log\lambda.$$

Thus taking

$$\mu = e^{\frac{(1-\varepsilon)\ell}{2}}\lambda,$$

conditions (89) and (90) of Lemma 11 are clearly fulfilled. For $x \notin \Omega$ the conclusion (91) is that

$$\left|\log \|A_{\ell n}(x)\| + \sum_{2}^{n-1} \log \|A_\ell(x + (j-1)\ell\alpha)\| - \sum_{1}^{n-1} \log \|A_{2\ell}(x + (j-1)\ell\alpha)\|\right| \le C\frac{n}{\mu}.$$

Divide the above inequality by ℓn and integrate it in $x \in \mathbb{R}/\mathbb{Z}$. Splitting the integration as $(\mathbb{R}/\mathbb{Z}) \setminus \Omega$ and Ω, we get

$$\left| L_{\ell n}(E) + \frac{n-2}{n} L_\ell(E) - \frac{2(n-1)}{n} L_{2\ell}(E) \right| < C\left(\mu^{-1}\ell^{-1} + \mathrm{Leb}(\Omega)\right) < Ce^{-\frac{1}{4}\delta\ell^\sigma}.$$

Note here C depends on λ. Hence we obtain

$$|L_{\ell n}(E) + L_\ell(E) - 2L_{2\ell}(E)| < Ce^{-\frac{1}{4}\delta\ell^\sigma}. \tag{93}$$

Applying (93) to $n = n_1 = \ell$ large enough and $\log n_2 \sim n_1^\sigma$ yields

$$|L_{n_2}(E) + L_{n_1}(E) - 2L_{2n_1}(E)| < Ce^{-\frac{1}{4}\delta n_1^\sigma}. \tag{94}$$

Applying (93) to $n = n_1 = \ell$ and $2n_2$ yields

$$|L_{2n_2}(E) + L_{n_1}(E) - 2L_{2n_1}(E)| < Ce^{-\frac{1}{4}\delta n_1^\sigma}.$$

Therefore, we get

$$|L_{2n_2}(E) - L_{n_2}(E)| < Ce^{-\frac{1}{4}\delta n_1^\sigma}. \tag{95}$$

Clearly, in (94) and (95), we may replace n_1 and n_2 by any n_s and $\log n_{s+1} \sim n_s^\sigma$. Thus we obtain

$$\begin{aligned} |L(E) - L_{n_2}(E)| &\le \sum_{s\ge 2} |L_{n_{s+1}}(E) - L_{n_s}(E)| \\ &\le 2\sum_{s\ge 2}\left(|L_{2n_s}(E) - L_{n_s}(E)| + Ce^{-\frac{1}{4}\delta n_s^\sigma}\right) \\ &< 4\sum_{s\ge 1} Ce^{-\frac{1}{4}\delta n_s^\sigma} < Ce^{-\frac{1}{4}\delta n^\sigma}. \end{aligned}$$

Thus, replacing $L_{n_2}(E)$ by $L(E)$ in (94) yields (92). □

Then we are already to prove Theorem 2

Proof. (**Proof of Theorem 2**) Clearly, we only need to show (6) for E' and E in the statement of Theorem 2 that are sufficiently close to each other. It is straightforward computation to see that $|L_n(E) - L_n(E')| < C^n|E - E'|$, where C depends on λ and v. So by (92), for all large n, it holds that

$$|L(E) - L(E')| < C^n|E - E'| + Ce^{-cn^\sigma}.$$

Then, by choosing

$$n = \left[\frac{1}{2\log C}\log\frac{1}{|E - E'|}\right],$$

we clearly get Lyapunov exponents part of (6)

$$|L(E) - L(E')| < Ce^{-c(\log|E_1 - E_2|^{-1})^\sigma}. \tag{96}$$

This says that Lyapunov exponents is weak Hölder continuous with respect to the energy. Now for the IDS part of (6), by (96) and the discussion following (4), we obtain for $E, E' \in \mathbb{R}$,

$$|N(E) - N(E')| < Ce^{-c(\log|E - E'|^{-1})^\sigma},$$

concluding the proof of Theorem 2. □

Appendix A. Proof of Lemmas in Section 2

A.1. **Proof of Lemma 1.** For $B \in \mathrm{SL}(2,\mathbb{R})$, it is a standard result that we can decompose it as $B = U_1\sqrt{B^tB}$, where $U_1 \in \mathrm{SO}(2,\mathbb{R})$ and $\sqrt{B^tB}$ is a positive symmetric matrix. We can further decompose $\sqrt{B^tB}$ as $\sqrt{B^tB} = U_2\Lambda U_2^t$, where $U_2 \in \mathrm{SO}(2,\mathbb{R})$ and $\Lambda = \begin{pmatrix} \|B\| & 0 \\ 0 & \|B\|^{-1} \end{pmatrix}$, thus $B = U_1U_2\Lambda U_2^t$.

Consider a map $B \in C^r(\mathbb{R}/\mathbb{Z}, SL(2,\mathbb{R}))$ for some $r \geq 1$. Then, it can be decomposed as

$$B(x) = U_1(x)U_2(x)\Lambda(x)U_2^t(x).$$

By Lemma 10 of [Z1], $U_1(x)$, $U_2(x)$ and $\Lambda(x)$ are C^r in x as long as $B(x)$ does not touch $\mathrm{SO}(2,\mathbb{R})$. Hence, we have

$$(U_1U_2)^t(x)B(x)(U_1U_2)(x-\alpha) = \Lambda(x)U(x),$$

where $U(x) = U_2^t(x)(U_1U_2)(x-\alpha) \in \mathrm{SO}(2,\mathbb{R})$. Let $c(x,t)$ be the upper left element of $U(x,t)$.

Now let us come back to the Schrödinger cocycles, we first use a simple trick to avoid that $A^{(E-\lambda v)}$ can always touch $\mathrm{SO}(2,\mathbb{R})$ for $t = \frac{E}{\lambda} \in v(\mathbb{R}/\mathbb{Z})$, which leads to the discontinuity of the polar decomposition. We instead consider $A^{(t,\lambda)} = TA^{(E-\lambda v))}T^{-1}$, where

$$T = \begin{pmatrix} \sqrt{\lambda}^{-1} & 0 \\ 0 & \sqrt{\lambda} \end{pmatrix}$$

This obviously does not change the dynamics. Thus

$$A(x) = A^{(t,\lambda)}(x) = \begin{pmatrix} \lambda[t-v(x)] & -\lambda^{-1} \\ \lambda & 0 \end{pmatrix}.$$

Let $r(x,t) = t - v(x)$. Then $r(x,t)$ is uniformly bounded on $\mathbb{R}/\mathbb{Z} \times \mathcal{I}$, where $\mathcal{I} \subset \mathbb{R}$ is any compact interval. If we set

$$a = a(x,t,\lambda) = r^2 + 1 + \frac{1}{\lambda^4} + \sqrt{(r^2+1+\frac{1}{\lambda^4})^2 - \frac{4}{\lambda^4}},$$

then obviously $\left|\frac{\partial^m a}{\partial^m x}\right|^{\pm}$ are uniformly bounded for all $(x,t,\lambda) \in \mathbb{R}/\mathbb{Z} \times \mathcal{I} \times [c,\infty)$ and $m = 0,1,2$. Then a direct computation shows that $\|A\| = \lambda\sqrt{\frac{a}{2}}$. Thus, it is clear that (9) holds for large λ.

A direct computation shows

$$U_2 = \frac{1}{\sqrt{(a-\frac{2}{\lambda^4})^2 + \frac{4}{\lambda^4}r(x)^2}} \begin{pmatrix} a - \frac{2}{\lambda^4} & \frac{2}{\lambda^2}r(x) \\ -\frac{2}{\lambda^2}r(x) & a - \frac{2}{\lambda^4} \end{pmatrix}$$

For simplicity let

$$f(x,t,\lambda) = \left(\sqrt{(a-\frac{2}{\lambda^4})^2 + \frac{4}{\lambda^4}r(x)^2}\right)^{-1}.$$

Thus we get that the corresponding upper-left element of U is

$$c(x,t,\lambda,\alpha) = c_4\left\{r(x-\alpha) - \frac{2r(x)}{\lambda^2 a(x)} + \frac{2r(x-\alpha)}{\lambda^4 a(x)} - \frac{4r(x)}{\lambda^6 a(x-\alpha)a(x)}\right\},$$

where

$$c_4 = \sqrt{\frac{2}{a(x-\alpha)}} f(x)f(x-\alpha)a(x)a(x-\alpha).$$

Hence, we have $c(x,t,\infty,\alpha) = \frac{t-v(x-\alpha)}{\sqrt{(t-v(x-\alpha))^2+1}}$. Furthermore, it is not difficult to see that for any fixed α,

$$c(x,t,\lambda,\alpha) \to c(x,t,\infty,\alpha) \text{ in } C^2(\mathbb{R}/\mathbb{Z}\times\mathcal{I},\mathbb{R}) \text{ as } \lambda\to\infty.$$

Indeed, it is easy to see this reduces to the convergence of $a(x,t,\lambda)$ to $a(x,t,\infty)$ in C^2 topology, which is immediate.

In the proof of Theorem 3, it is clear that the only important thing about the matrix $U = R_{\theta(x)}$ is the C^2 shape of the function θ, see, for example, Corollary 5. Thus, we could replace $c(x,t,\lambda,\alpha)$ by $c(x,t,\infty,\alpha)$ for large λ since they are sufficiently close in C^2 norm. After a translation, we can of course replace $v(x-\alpha)$ by $v(x)$. This completes the proof.

A.2. **Proof of Lemma 2.**

Proof. It suffices to consider $s(x)$ since all estimates for $u(x)$ follow from the same way. By the polar decomposition procedure we have that $s(x) = \frac{\pi}{2} + \theta_2(x)$, where $\theta_2(x)$ is the eigen-direction of $E^t(x)E(x)$ corresponding to the eigenvalue $\|E(x)\|^2$. For simplicity, let us omit the dependence on x in the following computation. A direct computation show that

$$E^tE = \begin{pmatrix} e_1^2e_2^2\cos^2\theta + e_1^2e_2^{-2}\sin^2\theta & (e_2^{-2}-e_2^2)\sin\theta\cos\theta \\ (e_2^{-2}-e_2^2)\sin\theta\cos\theta & e_1^{-2}e_2^{-2}\cos^2\theta + e_2^2e_1^{-2}\sin^2\theta \end{pmatrix}. \tag{97}$$

First note we have

$$e_3^2(x) + e_3^{-2}(x) = e_1^2e_2^2\cos^2\theta + e_1^2e_2^{-2}\sin^2\theta + e_1^{-2}e_2^{-2}\cos^2\theta + e_2^2e_1^{-2}\sin^2\theta. \tag{98}$$

Thus the results for $e_3(x)$ are quite clear since we always have $e_3(x)\gg 1$ $\forall x\in I$ in each case as stated in Lemma 2.

Now let's focus on $s(x)$. Let $a = e_1e_2$ and $b = \frac{e_1}{e_2}$. Then it is easily calculated that

$$\tan s(x) = -\cot\theta_2 = \frac{w}{u},$$

where we have that

$$w = 2(e_2^2 - e_2^{-2})\sin\theta\cos\theta$$

and

$$u = \sqrt{[(a^2+a^{-2})\cos^2\theta + (b^2+b^{-2})\sin^2\theta]^2 - 4} - [(a^2-a^{-2})\cos^2\theta + (b^2-b^{-2})\sin^2\theta].$$

Let $U = [(a^2-a^{-2})\cos^2\theta + (b^2-b^{-2})\sin^2\theta]$, then it is easy to see that

$$u = \sqrt{U^2+w^2} - U.$$

Thus it is straightforward that

$$\tan s(x) = \frac{w}{\sqrt{U^2+w^2}-U} = \frac{\sqrt{U^2+w^2}+U}{w}$$

To simplify the above formula, we divide it into the following cases. In the following discussion, we will also consider the case $\frac{\sqrt{U^2+w^2}+U}{w} = \pm\infty$ since they correspond to $s(x) = \pm\frac{\pi}{2}$.

(1) $e_1 \geq e_2 \gg 1$. Obviously, $U > 0$. Then we further subdivide this case into the following two cases.

(a) $w \geq 0$. Then we have

$$\tan s(x) = \frac{\sqrt{U^2 + w^2} + U}{w} = \sqrt{\frac{U^2}{w^2} + 1} + \frac{U}{w}.$$

(b) If $w < 0$, Then we have

$$\tan s(x) = -\frac{\sqrt{U^2 + w^2} + U}{-w} = -\left(\sqrt{\frac{U^2}{(-w)^2} + 1} + \frac{U}{-w}\right).$$

(2) $e_2 > e_1 \gg 1$. We again subdivide this case into the following cases.

(a) $w \geq 0$ and $U \geq 0$. Then we get that

$$\tan s(x) = \frac{\sqrt{U^2 + w^2} + U}{w} = \sqrt{\frac{U^2}{w^2} + 1} + \frac{U}{w}.$$

(b) $w < 0$ and $U \geq 0$. Then we get that

$$\tan s(x) = -\frac{\sqrt{U^2 + w^2} + U}{-w} = -\left(\sqrt{\frac{U^2}{w^2} + 1} + \frac{U}{-w}\right).$$

(c) $w \geq 0$ and $U < 0$. Then we have

$$\tan s(x) = \frac{w}{\sqrt{U^2 + w^2} - U} = \frac{1}{\sqrt{\frac{U^2}{w^2} + 1} + \frac{-U}{w}}.$$

(d) $w < 0$ and $U < 0$. Then we have

$$\tan s(x) = -\frac{-w}{\sqrt{U^2 + w^2} - U} = -\frac{1}{\sqrt{\frac{U^2}{w^2} + 1} + \frac{U}{w}}.$$

Now to further simplify these formula, we need to simplify $\frac{U}{w}$, and to understand how the simplification will affect the derivative estimate of $s(x)$. First, we note the following facts. Consider f as a function of x, then it holds that

$$\frac{d^m \tan^{-1}\left[\sqrt{f^2 + 1} + f\right]}{dx^m} = \frac{1}{2}\frac{d^m \tan^{-1} f}{dx^m}. \tag{99}$$

Indeed, a direction computation shows that

$$\frac{d \tan^{-1}\left[\sqrt{f^2 + 1} + f\right]}{dx} = \frac{f'}{2(1 + f^2)} = \frac{1}{2}\frac{d \tan^{-1} f}{dx}, \tag{100}$$

thus (99) follows. Since we will also consider the second order derivative, we get

$$\frac{d^2 \tan^{-1}\left[\sqrt{f^2 + 1} + f\right]}{dx^2} = \frac{1}{2}\frac{d^2 \tan^{-1} f}{dx^2} = \frac{f''}{2(1 + f^2)} - \frac{f(f')^2}{(1 + f^2)^2}. \tag{101}$$

The case $\cot^{-1}[\sqrt{f^2+1}+f]$ is similar. Indeed, $\sqrt{f^2+1}-f=\frac{1}{\sqrt{f^2+1}+f}$ implies

$$\frac{d^m\cot^{-1}\left[\sqrt{f^2+1}+f\right]}{dx^m}=\frac{d^m\tan^{-1}\left(\frac{1}{\sqrt{f^2+1}+f}\right)}{dx^m}$$
$$=\frac{d^m\tan^{-1}\left[\sqrt{f^2+1}-f\right]}{d^mx}$$
$$=\frac{1}{2}\frac{d^m\tan^{-1}(-f)}{dx^m}.$$

Now we take a close look at $\frac{U}{w}$, which is

$$\frac{U}{w}=\frac{(a^2-a^{-2})\cos^2\theta+(b^2-b^{-2})\sin^2\theta}{2(e_2^2-e_2^{-2})\sin\theta\cos\theta}$$
$$=\frac{(e_1e_2)^2-(e_1e_2)^{-2}}{2(e_2^2-e_2^{-2})}\cot\theta+\frac{(e_1e_2^{-1})^2-(e_1e_2^{-1})^{-2}}{2(e_2^2-e_2^{-2})}\tan\theta.$$

Now we first consider the case of $e_0\gg 1$ and $|\theta-\frac{\pi}{2}|\gg e_0^{-1}$. Thus, we have $|\cot\theta|,|\cos\theta|>ce_0^{-1}$. These in turn imply $U>0$ and $|e_1^2\cot\theta|\gg 1$. Then a direct computation shows that

$$\frac{U}{w}\precsim_2\frac{1}{2}e_1^2\cot\theta. \tag{102}$$

Since $\left|\frac{U}{w}\right|>c\left|\frac{1}{2}e_1^2\cot\theta\right|\gg 1$, we have

$$\pm\left(\sqrt{\frac{U^2}{w^2}+1}+\frac{U}{\pm w}\right)\precsim_0 2\frac{U}{w}, \tag{103}$$

where $\frac{U}{\pm w}>0$. Apply (100)–(103) to the cases (1), (2)(a) and (2)(b), we obtain (11).

The case $e_1(x)\gg e_2(x)\gg 1$ can be discussed similarly since we again have $U>0$. One only note that in this case we have

$$\frac{U}{w}\precsim_2\frac{1}{2}e_1^2\cot\theta+\frac{1}{2}e_1^2e_2^{-4}\tan\theta.$$

Finally, we consider the case $e_2(x)\gg e_1(x)\gg 1$. In this case, we claim that it always holds

$$\tan s(x)\precsim_2 e_1^2\cot\theta. \tag{104}$$

Given (104), (12) clearly follows from (100)-(101). Now we prove (104). For simplicity, define $G(y)=\sqrt{y^2+1}+y$. Then It is clear that

$$G\left[\frac{1}{2}(y-\frac{1}{y})\right]=\begin{cases}y, & y\geq 0;\\ \frac{1}{y}, & y<0.\end{cases}$$

Then for $U\geq 0$, it suffices to consider the case (2)(a). First note that in this case, we approximately have $e_1^2\cot\theta>1$. Then we may rewrite $\tan s(x)$ as

$$\tan s(x)=e_1^2\cot\theta+\left[G\left(\frac{U}{w}\right)-G\left(\frac{1}{2}e_1^2\cot\theta-\frac{1}{2e_1^2\cot\theta}\right)\right]. \tag{105}$$

On the other hand, It is clear that in this case

$$e_1^2 \cot\theta + \left[\frac{U}{w} - \left(\frac{1}{2}e_1^2\cot\theta - \frac{1}{2e_1^2\cot\theta}\right)\right] \precsim_2 e_1^2 \cot\theta. \tag{106}$$

Thus (105) and (106) together with the straightforward facts $\left\|\frac{d^m G}{dy^m}\right\|_\infty < 2$ for $m = 1, 2$ clearly imply (104).

If $U < 0$, it suffices to consider the case (2)(c). Note in this case, we approximately have $0 \le e_1^2 \cot\theta < 1$. Then we may rewrite $\cot s(x)$ as

$$\cot s(x) = \frac{1}{e_1^2\cot\theta} + \left[G\left(\frac{U}{-w}\right) - G\left(\frac{1}{2e_1^2\cot\theta} - \frac{1}{2}e_1^2\cot\theta\right)\right].$$

Thus by the same argument as in case (2)(a), we get

$$\cot s(x) \precsim_2 \frac{1}{e_1^2\cot\theta},$$

which clearly implies (104). This completes the proof. □

A.3. **Proof of Lemma 3 and 4.**

Proof. (**Proof of Lemma 3**) It is enough to consider only s since u can be treated similarly. By our assumption $\left|\theta - \frac{\pi}{2}\right| > ce_0^{-\eta}$, (11) can be applied in this case. Thus to estimate $\left|\frac{d^m s}{dx^m}\right|, m = 0, 1, 2$, it suffices to estimate each non-splittable term of

$$\frac{d^m \tan^{-1}(e_1^2\cot\theta)}{dx^m}.$$

Thus, from the case $m = 0$, we get

$$\left|s - \frac{\pi}{2}\right|_I < Ce_1^{-(2-\eta)}. \tag{107}$$

From $m = 1$, we get

$$\left|\frac{ds}{dx}\right| < \frac{C}{1 + e_1^4\cot^2\theta}\left(|2e_1e_1'\cot\theta| + |e_1^2\theta'\cot^2\theta| + |e_1^2\theta'|\right).$$

By our assumption, it is easy to see that we get

$$\left|\frac{ds}{dx}\right| < C\left|\frac{2e_1'}{e_1^3\cot\theta}\right| + C\left|\frac{\theta'}{e_1^2}\right| + C\left|\frac{\theta'}{e_1^2\cot^2\theta}\right|.$$

Hence, we have

$$\left|\frac{ds}{dx}\right|_I < Ce_1^{-(2-2\eta)} + Ce_1^{-(2-\eta)} + Ce_1^{-(2-3\eta)} < Ce_1^{-(2-3\eta)}. \tag{108}$$

For $m = 2$, we get

$$\frac{d^2\tan^{-1}(e_1^2\cot\theta)}{dx^2} = \frac{1}{1 + (e_1^2\cot\theta)^2}\frac{d^2(e_1^2\cot\theta)}{dx^2} - \frac{2e_1^2\cot\theta}{[1 + (e_1^2\cot\theta)^2]^2}\left[\frac{d(e_1^2\cot\theta)}{dx}\right]^2.$$

Thus we obtain

$$\left|\frac{d^2s}{dx^2}\right| < C\left|\frac{2e_1''}{e_1^3\cot\theta}\right| + \left|\frac{4e_1'\theta'}{e_1^3\cot^2\theta}\right| + \left|\frac{6(e_1')^2}{e_1^4\cot\theta}\right| + \left|\frac{2(\theta')^2}{e_1^2\cot^3\theta}\right| + \left|\frac{2\theta''}{e_1^2\cot^2\theta}\right|.$$

Hence, we obtain

$$\left|\frac{d^2s}{dx^2}\right|_I < Ce_1^{-(2-3\eta)} + Ce_1^{-(2-4\eta)} + Ce_1^{(2-5\eta)} < Ce_1^{-(2-5\eta)}. \tag{109}$$

Clearly, (107)–(109) imply (14).

Now we consider $e_3(x)$. Without loss of generality, assume $\cos\theta > 0$. Then by (11), it suffices to estimate each non-splittable term of

$$\frac{d^m(e_1e_2\cos\theta)}{dx^m} \text{ for } m = 1, 2. \tag{110}$$

Hence, we have

$$\left|\frac{de_3}{dx}\right| \le C\left(|e_1'e_2\cos\theta| + |e_1e_2'\cos\theta| + |e_1e_2\theta'\sin\theta|\right).$$

It is clear that it is dominated by the first two terms. Hence, we have

$$\left|\frac{de_3}{dx}\right| < C\left|e_1e_2^{1+\eta}\cos\theta\right| + C\left|e_1^{1+\eta}e_1\cos\theta\right| < C(e_2e_1\cos\theta)^{1+\eta} < Ce_3^{1+\eta}.$$

Then a straightforward calculation shows that

$$\begin{aligned}\left|\frac{d^2e_3}{dx^2}\right| <& C|e_1e_2''\cos\theta| + C|e_1''e_2\cos\theta| + Ce_1'e_2'\cos\theta\\ <& Ce_1e_2^{1+2\eta}\cos\theta + Ce_1^{1+2\eta}e_2\cos\theta + Ce_1^{1+\eta}e_2^{1+\eta}\cos\theta\\ <& C(e_1e_2\cos\theta)^{1+2\eta}\\ <& Ce_3^{1+2\eta},\end{aligned}$$

which completes the proof. □

Proof. (**Proof of Lemma 4**).

By (20), we only need to show (17) and (18).

We will proceed by induction. In case of the product of two matrices, i.e. $k = 2$, by (16), Lemma 3 can be applied to $E_2(\ell) = E(\ell + 1) \cdot E(\ell)$. Thus we get for all $0 \le \ell \le n - 2$,

$$\|s(\ell) - s_2(\ell)\|_{C^2},\ \|u(\ell+1) - u_2(\ell+2)\|_{C^2} < C\lambda_0^{-(2-5\eta)};$$
$$\left|\frac{d^m\lambda_2(\ell)}{d^mx}\right| < [\lambda_2(\ell)]^{1+m\eta},\ m = 1, 2.$$

We want to make sure that the error terms

$$\|s(\ell) - s_k(\ell)\|_{C^2} \text{ and } \|u(\ell + k - 1) - u_k(\ell + k)\|_{C^2}$$

do not accumulate as k gets large. So we assume for $k \le n - 1$ and for all possible ℓ,

$$\|s_k(\ell) - s_{k-1}(\ell)\|_{C^2},\ \|u_k(n-\ell) - u_{k-1}(n-\ell)\|_{C^2} < C\lambda_0^{-(2-5\eta)(k-1)(1-\eta)}; \tag{111}$$

$$\left|\frac{d^m\lambda_k(\ell)}{d^mx}\right| < [\lambda_k(\ell)]^{1+m\eta},\ m = 1, 2. \tag{112}$$

Clearly, (111) implies for $\ell = 0, 1$,

$$\begin{aligned}&\|s(\ell) - s_{n-1}(\ell)\|_{C^2},\ \|u(n-\ell-1) - u_{n-1}(n-\ell)\|_{C^2} \\ &< C\lambda_0^{-(2-5\eta)} + \sum_{j=2}^{n-2}\lambda_0^{-(2-5\eta)(1-\eta)j}\\ &< C\lambda_0^{-(2-5\eta)}.\end{aligned} \tag{113}$$

(113) in turn implies for $m = 1, 2$ and $\ell = 0, 1$

$$\left|\frac{d^m u_{n-1}(n-\ell)}{dx^m}\right|, \left|\frac{d^m s_{n-1}(\ell)}{dx^m}\right| < C\lambda_0^{\eta}. \tag{114}$$

On the other hand, combining our conditions

$$\|u(n-2) - s(n-1)\|_{C^2}, \ \|u(0) - s(1)\|_{C^2} > c\lambda_0^{-\eta}$$

with (113), we obtain

$$|u_{n-1}(n-1) - s(n-1)|, \ |s_{n-1}(1) - u(0)| > c\lambda_0^{-\eta} \tag{115}$$

Now, (20), (112), (114) and (115) imply that we can apply Lemma 3 to the product

$$E_n(0) = E(n-1) \cdot E_{n-1}(0) = E_{n-1}(1) \cdot E(0),$$

which yields (18) and

$$\|s_{n-1}(0) - s_n(0)\|_{C^2}, \ \|u_n(n) - u_{n-1}(n)\|_{C^2} < C\lambda_0^{-(1-\eta)(n-1)(2-5\eta)}. \tag{116}$$

By the induction assumption (111) and the same argument we obtain (113), (116) clearly implies (17). □

A.4. **Proof of Lemma 5 and 6.**

Proof. (**Proof of Lemma 5**). For first two estimate, which are (21) and (22), it again suffices to consider s since u can be done similarly. Note the difference between the current situation and the one for (14) is that, here it is possible that $\theta = \frac{\pi}{2}$. Now for (21), we apply the first estimate of (12). For $m = 1$, similar to the proof of (108), we get

$$\left|\frac{ds}{dx}\right| < \frac{C}{1 + e_1^4 \cot^2\theta}\left(|2e_1 e_1' \cot\theta| + \left|e_1^2\theta' \cot^2\theta\right| + \left|e_1^2\theta'\right|\right) < Ce_1^{2+\eta},$$

where the worst case happens when $\theta = \frac{\pi}{2}$ and hence $\cot\theta = 0$. For $m = 2$, similar to the proof of (109), we get

$$\left|\frac{d^2 s}{dx^2}\right| < C\left|\frac{e_1^6(\theta')^2\cot\theta}{1 + e_1^8\cot^4\theta}\right|_{\cot\theta\approx e_1^{-2}} < Ce_1^{4+2\eta}.$$

For (22), we apply the first part of (13). In other words, it suffices to estimate each non-splittable term of

$$\frac{d^m \tan^{-1}(f)}{dx^m},$$

where $f(x) = e_1^2\cot\theta + e_1^2 e_2^{-4}\tan\theta$. Note $|f| \geq e_1^2 e_2^{-2} \geq e_1^{2(1-\beta)}$. The worst case again happens when $\theta = \frac{\pi}{2}$ or $\tan\theta = \infty$. Note $e_3 > e_1 e_2^{-1}$ and $e_2 \leq e_1^{\beta}$ for some $\beta \ll 1$. Then for $m = 1$, we get

$$\left|\frac{ds}{dx}\right| < Ce_1^{-2}e_2^4 \cdot |\theta'| < C(e_1 e_2^{-1})^{\frac{3}{2}} < Ce_3^{-\frac{3}{2}}.$$

For $m = 2$, we get

$$\left|\frac{d^2 s}{dx^2}\right| < Ce_1' e_1^{-3} e_2^4 \cdot |\theta'| < C(e_1 e_2^{-1})^{\frac{3}{2}} < Ce_3^{-\frac{3}{2}}.$$

Now let us consider the function $e_3(x)$. Clearly, here we can apply (12). In other words, it suffices to estimate each non-splittable term of

$$\frac{d^m\sqrt{e_1^2e_2^2\cos^2\theta+e_1^{-2}e_2^2\sin^2\theta}}{dx^m}.$$

The difference between the case in question and the case in Lemma 3 is that it is possible that $\cos\theta = 0$. In fact, this is the the worst case in the sense that the derivatives of e_3 may get large with respect to e_3 itself. Thus we only need to consider the case that e_3 is dominated by the second term. So we must have $\sin\theta \approx \pm 1$. Without loss of generality, assume $\sin\theta > 0$. Thus, we essentially only need to estimate each non-splittable term of

$$\frac{d^m[e_1^{-1}e_2\sin\theta]}{dx^m}.$$

Thus we get

$$\frac{de_3}{dx} < C\left(\left|\frac{e_1'}{e_1^2}e_2\sin\theta\right| + \left|e_1^{-1}e_2'\sin\theta\right| + \left|e_1^{-1}e_2\cdot\theta'\cos\theta\right|\right)$$

By our assumption, it is easy to see that $\frac{de_3}{dx}$ is dominated by the second term. Thus we have

$$\left|\frac{de_3}{dx}\right| < Ce_1^{-1}e_2'|\sin\theta| < Ce_1^{-1}e_2^{1+\eta}.$$

Now we need to find a γ_1 so that

$$\left|\frac{de_3}{dx}\right| < Ce_1^{-1}e_2^{1+\eta} < C(e_1^{-1}e_2)^{1+\gamma_1} < Ce_3^{1+\gamma_1},$$

leading to $e_2^{\gamma_1-\eta} > e_1^{\gamma_1}$.

Then it is easy to see that it is enough to choose $\gamma_1 = \eta + 2\eta\beta$. Similarly, we get

$$\left|\frac{d^2e_3}{dx^2}\right| < Ce_1^{-1}e_2''\sin\theta < Ce_1^{-1}e_2^{1+2\eta}.$$

By choosing $\gamma_2 = 2\eta + 4\eta\beta$, we have

$$\left|\frac{d^2e_3}{dx^2}\right| < Ce_1^{-1}e_2^{1+2\eta} < C(e_1^{-1}e_2)^{1+\gamma_2} < Ce_3^{1+\gamma_2}.$$

Clearly, this concludes the proof. □

Proof. (**Proof of Lemma 6**) Recall we have

$$f = \tan^{-1}(l^2[\tan f_1(x)]) - \frac{\pi}{2} + f_2, \tag{117}$$

where f_1 is of type I_+ and f_2 of type I_-. Furthermore, $f_1(0) = 0$ and $f_2(d) = 0$ with $d \geq 0$. A direct computation shows that

$$\frac{df}{dx}(x) = \frac{2ll'\tan[f_1(x)]}{1+l^4\tan^2[f_1(x)]} + \frac{l^2+l^2\tan^2[f_1(x)]}{1+l^4\tan^2[f_1(x)]}\cdot\frac{df_1}{dx}(x) + \frac{df_2}{dx}(x).$$

We may simplify $\frac{df}{dx}$ as

$$\frac{df}{dx}(x) = \frac{l^2}{1+l^4\tan^2(f_1)}\cdot\frac{df_1}{dx} + \frac{df_2}{dx} \tag{118}$$

since the other terms are negligible in the following estimate. Similarly, we may take in the following estimate

$$\frac{d^2 f}{dx^2} = \frac{l^2}{1+l^4\tan^2(f_1)}\cdot\frac{d^2 f_1}{dx^2} + \frac{d^2 f_2}{dx^2} - \frac{l^6\tan(f_1)}{1+l^8\tan^4(f_1)}\cdot\left(\frac{df_1}{dx}\right)^2. \tag{119}$$

(117), (118) and (119) clearly imply that, $\|f-f_2\|_{C^2} < Cl^{-\frac{3}{2}}$ on $I\setminus B(0,Cl^{-\frac{1}{4}})$. In the following discussion, let r_j, $1\le j\le 9$ be numbers such that $cr^2\le r_j\le Cr^{-2}$.

Now by (118) and the definition of type of f_1 and f_2, it is straightforward calculation to see that there exist two points x_3 and x_4 with

$$x_3 = -r_1 l^{-1} \text{ and } x_4 = r_2 l^{-1}$$

such that

$$f \text{ increases from } r_3 l^{-1} + cr^3 - \pi \text{ to } cr^3 - r_4 l^{-1} \text{ on } [x_3, x_4].$$

In case $d<\frac{r}{3}$, x_3 and x_4 are exactly the two solutions of $\frac{df}{dx}=0$. In case of $d\ge\frac{r}{3}$, depending on the sign of the derivative of f_2, they might be the solutions of $\frac{df}{dx}=0$ or not. In any case, we may still choose such a pair of x_3 and x_4.

Then we first consider the case $d\ge\frac{r}{3}$. By (118), the following clearly holds:

Thus there is a zero contained in $B(0,Cl^{-\frac{3}{4}})$, say x_1. Hence, $|x_1|<Cl^{-\frac{3}{4}}$. We must have $B(d,\frac{r}{4})\cap B(0,Cl^{-\frac{1}{4}})=\varnothing$ since $d\ge\frac{r}{3}$. Hence, there is a zero contained in $B(d,Cl^{-\frac{3}{4}})$, say x_2. This implies that

$$B(x_2,\frac{r}{4})\cap B(x_1,Cl^{-\frac{1}{4}})=\varnothing \text{ and } \|f-f_2\|_{C^2}<Cl^{-\frac{3}{2}} \text{ on } B(x_2,\frac{r}{4}).$$

Now we have

$$\|f-f_2\|_{C^2}<Cl^{-\frac{3}{2}} \text{ on } I\setminus B(x_1,Cl^{-\frac{1}{4}}), \text{ and } |f_2|>cr^3 \text{ on } I\setminus B(x_2,\frac{r}{4}).$$

This clearly implies that

$$|f(x)|>cr^3 \text{ for all } x\notin B(x_1,Cl^{-\frac{1}{2}})\cup B(x_2,\frac{r}{4}).$$

If $d<\frac{r}{3}$, similar to the case $d\ge\frac{r}{3}$, we get

$$|f(x)|>cr^3 \text{ for all } x\notin B(0,\frac{r}{6})\cup B(d,\frac{r}{6}).$$

Now let us focus the set $B(0,\frac{r}{6})\cup B(d,\frac{r}{6})$. To find x_1 and x_2, note f is strictly increasing on $[x_3,x_4]=[-r_1l^{-1},r_2l^{-1}]$. Then, by (117), we have

$$f(x)=\begin{cases} r_3l^{-1}+r_5(d+r_1l^{-1})-\pi, & x=-r_1l^{-1}\\ -r_4l^{-1}+r_6(d-r_2l^{-1}), & x=r_2l^{-1}.\end{cases} \tag{120}$$

Note $f(Cl^{-2})=\tan^{-1}(Cr_7)-\frac{\pi}{2}+f_2(Cl^{-2})<0$. So to solve the possible equation $f(x)=0$, we only need to consider $x>Cl^{-2}$. Then we may write $f(x)$ as

$$f(x)=\tan^{-1}(l^2[\tan f_1(x)])-\frac{\pi}{2}+f_2(x)=-cl^{-2}(r_8x)^{-1}+cr_9(d-x).$$

If $d > c(l\sqrt{r_8 r_9})^{-1}$, we get two solutions $x_{1,2} = \frac{d \pm \sqrt{d^2 - c(r_8 r_9 l^2)^{-1}}}{2}$ of $f(x) = 0$. If $0 \leq d \leq c(l\sqrt{r_8 r_9})^{-1}$, we get

$$\begin{aligned} f(x) &= -cl^{-2}(r_8 x)^{-1} + cr_9(d - x) = -cr_8^{-1} l^{-2} x^{-1} - cr_9 x + cr_9 d \\ &\leq -c\sqrt{\frac{r_9}{r_8}} \cdot l^{-1} + cr_9 d. \end{aligned}$$

The equality holds when $x_{1,2} = c(l\sqrt{r_8 r_9})^{-1}$. In any case, we get $|x_1| < Cl^{-\frac{3}{4}}$ and $|x_2 - d| < Cl^{-\frac{3}{4}}$. Hence, we get

$$|f(x)| > cr^3 \text{ for all } x \notin B(x_1, \frac{r}{6}) \cup B(x_2, \frac{r}{6})$$

in this case. From the discussion above, we clearly get the bifurcation as d varying.

By choosing

$$\eta_0 = \max\{r_1, r_2\},\ \eta_1 = r_3,\ \eta_2 = c\sqrt{r_8 r_9}^{-1},\ \eta_3 = c\sqrt{\frac{r_9}{r_8}} \text{ and } \eta_4 = cr_9,$$

we get the corresponding estimates in Lemma 6.

From the graph of f and the bifurcation procedure, it is also clear that

$$\text{either } 0 < x_1 \leq x_2 < d \text{ if } f(x_1) = f(x_2) = 0; \text{ or } x_1 = x_2 \text{ if } f(x_1) = f(x_2) \neq 0.$$

Moreover, whenever $x_1 = x_2$, we must have $x_1 = x_4 = x_2$ and $\frac{df}{dx}(x_4) = 0$.

Finally, to estimate $|\frac{d^2 f}{dx^2}|$ when $|\frac{df}{dx}| < r^2$. By (118), we only need to take care of the case when $l^{-\frac{5}{2}} < \tan^2[f_1(x)] < l^{-\frac{3}{2}}$. By (119), in this case, it is clear that $|\frac{d^2 g_2}{dx^2}|$ is dominated by

$$\frac{l^6 \tan[f_1(x)]}{1 + l^8 \tan^4[f_1(x)]} \cdot [\frac{df_1}{dx}(x)]^2,$$

which is of order at least

$$l^{\frac{1}{4}}[\frac{df_1}{dx}(x)]^2 > cl^{\frac{1}{4}} r^4 \gg c \tag{121}$$

since $l \gg r^{-1}$. This completes the proof. □

Appendix B. Applications

It is clear that the proof the Theorem 3, hence of Theorem 1 and 2, can be applied to any one parameter family of cocycle maps $B \in C^2(\mathcal{J} \times \mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ such that we could get started with our induction as case $(1)_{\mathrm{I}}$ and $(2)_{\mathrm{II}}$ in Section 3.1. Here $\mathcal{J} \subset \mathbb{R}$ is any compact interval of parameters. In particular, consider

$$B^{(t,\lambda)} = \Lambda(x) \circ R_{\psi(x,t)} = \begin{pmatrix} \lambda(x) & 0 \\ 0 & \lambda^{-1}(x) \end{pmatrix} \cdot \begin{pmatrix} \cos\psi(x,t) & -\sin\psi(x,t) \\ \sin\psi(x,t) & \cos\psi(x,t) \end{pmatrix}$$

with $\psi(t,x) \in C^2(\mathcal{J} \times \mathbb{R}/\mathbb{Z}, \mathbb{R})$, $\lambda(x) \in C^2(\mathbb{R}/\mathbb{Z}, \mathbb{R})$. Assume $\lambda(x)$ and $\psi(t,x)$ satisfying the following conditions.

First, $\lambda(x) > \lambda \gg 1$ and $\left|\frac{d^m \lambda(x)}{dx^m}\right| < C\lambda$ for each $x \in \mathbb{R}/\mathbb{Z}$ and $m = 0, 1, 2$. For each t, $\psi(t; \mathbb{R}/\mathbb{Z}) \subset [0, \pi)$ in $\mathbb{RP}^1$.

Secondly, For each $t \in \mathcal{J}$, we have that the set

$$\mathcal{C}_t := \{x : \psi(x,t) = \frac{\pi}{2}\} = \{c_{t,1}, c_{t,2}\} \text{ with the possibility that } c_{t,1} = c_{t,2}.$$

Finally, there exists a $r > 0$ such that if we consider the interval $I(t) = I_1(t) \cup I_2(t)$ with $I_j(t) = B(c_{j,t}, r)$, $j = 1, 2$, then we have the following.

- If $I_1(t) \cap I_2(t) = \varnothing$, then $\psi(t; \cdot)$ is of type I on $I_j(t)$, $j = 1, 2$. Moreover, if $\psi(t; \cdot)$ is of type I_- on $I_1(t)$ then it is of type I_+ on $I_2(t)$, vice versa.
- If $I_1(t) \cap I_2(t) \neq \varnothing$, then $\psi(t; \cdot)$ is of type II on $I(t)$.

Let $L(\alpha, B^{(t,\lambda)})$ be the Lyapunov exponents of the dynamical systems $(\alpha, B^{(t,\lambda)})$. Then we have that the follow corollary of the proof of Theorem 1 and 3.

Corollary 5. *For the given $B^{(t,\lambda)}$ as above, for each $\alpha \in DC_\tau$ with $\tau > 2$ and each $\varepsilon > 0$, there exists a $\lambda_0 = \lambda_0(\alpha, B, \varepsilon)$ such that*

$$L(\alpha, B^{(t,\lambda)}) > (1 - \varepsilon) \log \lambda$$

for all $(t, \lambda) \in \mathcal{J} \times [\lambda_0, \infty)$. Moreover, for any fixed $\lambda > \lambda_0$ and for all $t, t' \in \mathcal{J}$, it holds that

$$|L(\alpha, B^{(t,\lambda)}) - L(\alpha, B^{(t',\lambda)})| < Ce^{-c(\log |t-t'|^{-1})^\sigma},$$

where $c, C > 0$ depend on α, ψ, λ, and $0 < \sigma < 1$ on α.

Clearly, Corollary 2 is a direct consequence of Lemma 1 and Corollary 5. We may also apply Corollary 5 to the Szegő cocycles which arise naturally in the study of orthogonal polynomial on the unit circle. See [Z1] for a brief introduction. For detailed information, see [Sim1] and [Sim2]. In particular, the cocycle map $A^{(E,f)} : \mathbb{R}/\mathbb{Z} \to \mathrm{SU}(1, 1)$ is given by

$$A^{(E,f)}(x) = (1 - |f(x)|^2)^{-1/2} \begin{pmatrix} \sqrt{E} & \frac{-\overline{f(x)}}{\sqrt{E}} \\ f(x)\sqrt{E} & \frac{1}{\sqrt{E}} \end{pmatrix}, \tag{122}$$

where $E \in \partial\mathbb{D}$, $\mathbb{D}$ is the open unit disk in complex plane $\mathbb{C}$, and $f : \mathbb{R}/\mathbb{Z} \to \mathbb{D}$ is a measurable function satisfying

$$\int_X \ln(1 - |f|) d\mu > -\infty.$$

$\mathrm{SU}(1, 1)$ is the subgroup of $\mathrm{SL}(2, \mathbb{C})$ preserving the unit disk in $\mathbb{CP}^1 = \mathbb{C} \cup \{\infty\}$ under Möbius transformations. It is conjugate in $\mathrm{SL}(2, \mathbb{C})$ to $\mathrm{SL}(2, \mathbb{R})$ via

$$Q = \frac{-1}{1 + i} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \in \mathbb{U}(2).$$

In other words, $Q^* \mathrm{SU}(1, 1) Q = \mathrm{SL}(2, \mathbb{R})$. Now consider a function $\theta \in C^2(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ such that $\theta(\mathbb{R}/\mathbb{Z}) \subset [0, \frac{1}{2})$ and for some *Diophantine* α, $\theta(x) - \theta(x - \alpha)$ is of the same type of function with v in Theorem 1.

One easy example is that $\theta(x) = \frac{1}{2}\cos(x)$, of which $\theta(x) - \theta(x - \alpha)$ is of the same type of function with v for all irrational α. Then, we have the follow corollary of Corollary 5.

Corollary 6. *Let $f = \lambda e^{2\pi i[\theta(x)+kx]}$, $0 < \lambda < 1$, $k \in \mathbb{Z}$ with θ satisfying the above conditions. Let α be a Diophantine number such that $\theta(x) - \theta(x - \alpha)$ is of the same type of function with v in Theorem 1. Then for each Diophantine α and each $\varepsilon > 0$, there exists a $\lambda_0 = \lambda_0(\theta, \alpha, \varepsilon) \in (0, 1)$ such that*

$$L(\alpha, A^{(E,f)}) > -\frac{1}{2}(1 - \varepsilon) \log(1 - \lambda)$$

for all $(E,\lambda) \in \partial D \times [\lambda_0, 1)$. *Moreover, for any fixed* $\lambda \in [\lambda_0, 1)$ *and for all* $E, E' \in \partial\mathbb{D}$, *it holds that*

$$|L(\alpha, A^{(E,f)}) - L(\alpha, A^{(E',f)})| < Ce^{-c(\log|E-E'|^{-1})^{\sigma}},$$

where $c, C > 0$ *depend on* α, θ, λ, *and* $0 < \sigma < 1$ *on* α.

Proof. Transform $\mathrm{SU}(1,1)$ to $\mathrm{SL}(2,\mathbb{R})$, set $E = e^{2\pi t}$ for $0 \le t < 1$ and do the polar decomposition. We see that the cocycle map (122) can be transformed into the following form

$$A^{(\lambda,t)} = \begin{pmatrix} \sqrt{\frac{1+\lambda}{1-\lambda}} & 0 \\ 0 & \sqrt{\frac{1-\lambda}{1+\lambda}} \end{pmatrix} \cdot R_{\psi(x,t)},$$

where $\psi(x,t) = \pi[\theta(x) - \theta(x-\alpha) + k\alpha + t]$. Clearly, $\psi(x,t)$ satisfies all the conditions of Corollary 5. This concludes the proof. $\square$

Remark 9. [Z1, Theorem A] *constructed analytic Szegő cocycles with uniformly positive Lyapunov exponents, which answers a question proposed in* [Sim2, Section 10.16] *and* [DaKr, Section 3]. *Clearly, Corollary 6 is a smooth version, which to the best of our knowledge, is the first example of this kind.*

Finally, let us mention the following application of Theorem 1. Denote by $\Sigma(H_{\alpha,v,x})$ the spectrum of the Schrödinger operator $H_{\alpha,v,x}$. Let

$$\Sigma_{\alpha,v} = \bigcup_{x \in \mathbb{R}/\mathbb{Z}} \Sigma(H_{\alpha,v,x}).$$

Define $L_+(\alpha, v) = \{E : L(\alpha, v, E) > 0\}$ with $L(\alpha, v, E)$ the associated Lyapunov exponent. Recently, Jitomirskaya and Mavi proved the following result in [JiMav].

Proposition 2. *For each irrational* α, *there exists a sequence of rationals* $\frac{p_n}{q_n} \to \alpha$ *such that for any potential* $v \in C^{\gamma}(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ *with* $\gamma > \frac{1}{2}$,

$$\lim_{n\to\infty} \Sigma_{\frac{p_n}{q_n},v} \cap L_+(\alpha, v) = \Sigma_{\alpha,v} \cap L_+(\alpha, v).$$

Moreover, in the Diophantine *case, the sequence* $\frac{p_n}{q_n}$ *is the full sequence of continued fraction approximants of* α.

Here, the sequence of bounded measurable sets $B_n \subset \mathbb{R}$ converges to B in the following sense.

$$\limsup_{n\to\infty} B_n = \liminf_{n\to\infty} B_n = B. \tag{123}$$

Note $\lim_{n\to\infty} \Sigma_{\alpha_n,v_n} = \Sigma_{\alpha,v}$ in Hausdorff metric if $\lim_{n\to\infty}(\alpha_n, v_n) = (\alpha, v)$ in $(\mathbb{R}/\mathbb{Z}, |\cdot|_{\infty}) \times C^0(\mathbb{R}/\mathbb{Z}, \mathbb{R})$, see [ADZ, Lemma 12]. Thus, the convergence in the sense of (123) is more delicate. In particular, it implies that

$$\lim_{n\to\infty} \mathrm{Leb}(B_n) = \mathrm{Leb}(B).$$

Thus, we have the following immediate corollary of Theorem 1 and Proposition 2.

Corollary 7. *Let* α *and* v *be given in Theorem 1. Let* $\frac{p_n}{q_n}$ *be the sequence of continued fraction approximants of* α. *Then there exists* $\lambda_0 > 0$ *such that for any* $\lambda > \lambda_0$, *it holds that*

$$\lim_{n\to\infty} \Sigma_{\frac{p_n}{q_n},\lambda v} = \Sigma_{\alpha,\lambda v},$$

which implies that

$$\lim_{n\to\infty} \mathrm{Leb}(\Sigma_{\frac{p_n}{q_n},\lambda v}) = \mathrm{Leb}(\Sigma_{\alpha,\lambda v}).$$

References

[A1] A. Avila, Density of positive Lyapunov exponents for SL(2; $\mathbb{R}$)-cocycles, *J. Amer. Math. Soc.* **24**(2011), 999-1014.

[A2] A. Avila, Global theory of one-frequency Schrödinger operators I: Stratified analyticity of the Lyapunov exponent and the boundary of nonuniform hyperbolicity, arXiv:0905.3902v1(2009).

[A3] A. Avila, Global theory of one-frequency Schrödinger operators II: Acriticality and finiteness of phase transitions for typical potentials, Preprint.

[A4] A. Avila, On the spectrum and Lyapunov exponent of limit periodic Schrödinger operators, *Comm. Math. Phys.* **288**(2009), 907-918.

[AD] A. Avila and D. Damanik, Generic singular spectrum for ergodic Schrödinger operators, *Duke Math. J.* **130**(2005), 393-400.

[ADZ] A. Avila, D. Damanik and Z. Zhang, Singular density of states measure for subshift and quasi-periodic Schrödinger operators, *Comm. Math. Phys.* **330**(2014), 469-498.

[AK] A. Avila and R. Krikorian, Monotonic cocycles, arXiv:1310.0703v1.

[AvJi] A. Avila and S. Jitomirskaya, The Ten Martini Problem, *Ann. Math.* **170**(2009), 303-342.

[AvJiSa] A. Avila, S. Jitomirskaya and C. Sadel, Complex one-frequency cocycles, *J. Eur. Math. Soc.* **16**(9)(2014), 1915-1935.

[BeCa] M. Benedicks and L. Carleson, The dynamics of the Hénon map, *Ann. Math.* **133**(1991), 73-169.

[Bj1] K. Bjerklöv, The dynamics of a class of quasi-periodic Schrödinger cocycles, arXiv:1311.5394v1.

[Bj2] K. Bjerklöv, Positive Lyapunov exponent and minimality for a class of one dimensional quasi-periodic Schrödinger equations, *Ergod. Th. Dynam. Sys.* **25**(2005), 1015-1045.

[Boc1] J. Bochi, Discontinuity of the Lyapunov exponent for non-hyperbolic cocycles, unpublished, 1999.

[Boc2] J. Bochi, Genericity of zero Lyapunov exponents, *Ergodic Theory Dynam. Systems* **22**(6)(2002), 1667-1696.

[Bou1] J. Bourgain, Positive Lyapunov exponents for most energies, *Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics* **1745**:37-66(Springer, Berlin, 2000).

[Bou2] J. Bourgain, On the spectrum of lattice Schrödinger operators with deterministic potential, *J. Anal. Math.* **87**(2002), 37-75.

[Bou3] J. Bourgain, Positivity and continuity of the Lyapunov exponent for shifts on $\mathbb{T}^d$ with arbitrary frequency vector and real analytic potential, *J. Anal. Math.* **96** (2005), 313-355.

[BoGo] J. Bourgain and M. Goldstein, On nonperturbative localization with quasi-periodic potential, *Ann. of Math.* **152** (2000), 835-879.

[BoGoSc] J. Bourgain, M. Goldstein and W. Schlag, Anderson localization for Schrödinger operators on $\mathbb{Z}$ with potentials given by skew-shift, *Commun. Math. Physics* **220**(3)(2001), 583-620.

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

[BoSc] J. Bourgain and W. Schlag, Anderson localization for Schrödinger operators on $\mathbb{Z}$ with strongly mixing potentials, *Commun. Math. Phys.* **215**(2000), 143-175.

[Cha] J. Chan, Method of variations of potential of quasi-periodic Schrödinger equations, *Geom. Funct. Anal.* **17**(2008), 1416-1478.

[ChuSp] V. Chulaevsky and T. Spencer, Positive Lyapunov exponents for a class of deterministic potentials, *Comm. Math. Phys.* **168**(1995), 455-466.

[CrSim] W. Craig and B. Simon, Subharmonicity of the Lyaponov index, *Duke Math. J.* **50**(1983), 551-560.

[DuKl] P. Duarte and S.Klein, Continuity of the Lyapunov exponents for quasiperiodic cocycles, *Comm. Math. Phys.* **332**(3)(2014), 1113-1166.

[DaKr] D. Damanik and H. Krüger, Almost Periodic Szegő cocycles with uniformly positive Lyapunov exponents, *J. Approx. Theory* **61**(2009), 813-818.

[E] L.H. Eliasson, Discrete one-dimensional quasi-periodic Schrödinger operators with pure point spectrum, *Acta Math.* **179**(2) (1997), 153-196.

[Fur] A. Furman, On the multiplicative ergodic theorem for the uniquely ergodic systems, *Ann. Inst. Henri Poincaré* **33**(1997), 797-815.

[FrSpWi] J. Fröhlich, T. Spencer and P. Wittwer, Localization for a class of one-dimensional quasi-periodic Schrödinger operators, *Comm. Math. Phys.* **132**(1990), 5-25.

[Fu] H. Furstenberg, Noncommuting random products, *Trans. Amer. Math. Soc.* **108**(1963), 377-428.

[GoSc] M. Goldstein and W. Schlag, Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions, *Annals of Math.* **154**(2001), 155-203.

[H] M. Herman, Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2, *Comment. Math. Helv.* **58**(1983), 453-502.

[I] K. Ishii, Localization of eigenstates and transport phenomena in one-dimensional disordered systems, *Suppl. Prog. Theoret. Phys.* **53** (1973), 77-138.

[JiKoSch] S. Jitomirskaya, D. Koslover and M. Schulteis, Continuity of the Lyapunov exponent for general analytic quasiperiodic cocycles, *Ergodic theory and dynamical systems* **29**(2009), 1881-1905.

[JiMar1] S. Jitomirskaya and C. Marx, Continuity of the Lyapunov exponent for analytic quasi-periodic cocycles with singularities, *Journal of Fixed Point Theory and Applications* **10**(2011), 129-146.

[JiMar2] S. Jitomirskaya and C. Marx, Analytic quasi-perodic cocycles with singularities and the Lyapunov exponent of extended Harper's model, *Commun. Math. Phys.* **316**(1)(2012), 237-267.

[JiMav] S. Jitomirskaya and R. Mavi, Continuity of the measure of the spectrum for quasiperiodic Schrödinger operators with rough potentials, *Comm. Math. Phys.* **325**(2)(2014), 585-601.

[Jo] R. Johnson, Exponential dichotomy, rotation number, and linear differential operators with bounded coefficients, *J. Differential Equations* **61**(1986), 54–78.

[Kn] O. Knill, The upper Lyapunov exponent of $SL(2;\mathbb{R})$ cocycles: Discontinuity and the problem of positivity, Lecture notes in Math. **1486**, Lyapunov exponents (Oberwolfach, 1990) 86-97, 1991.

[Kl] S. Klein, Anderson localization for the discrete one-dimensional quasi-periodic Schrödinger operator with potential defined by a Gevrey-class function, *J. Funct. Anal.* **218**(2005), 255-292.

[Ko1] S. Kotani, Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators, *Stochastic Analysis*(Katata/Kyoto, 1982), (North-Holland Math. Library 32, North-Holland, Amsterdam)(1984), 225-247.

[Ko2] S. Kotani, Jacobi matrices with random potentials taking finitely many values, *Rev. Math. Phys.* **1** (1989), 129-133.

[Kr1] H. Krüger, An explicit example of a skew-shift Schröinger operator with positive Lyapunov exponent at small coupling, preprint.

[Kr2] H. Krüger, Multiscale analysis for ergodic Schröinger operators and positivity of Lyapunov exponents, *J. Anal. Math.* **115**(2011), 343-387.

[Ma1] R. Ma$\tilde{n}\acute{e}$, Oseledec's theorem from the generic viewpoint, In Proceedings of the ICM (Warsaw, 1983), 1269-1276, PWN, Warsaw, 1984.

[Ma2] R. Ma$\tilde{n}\acute{e}$, The Lyapunov exponents of generic area preserving diffeomorphisms, In International Conference on Dynamical Systems (Montevideo, 1995), 110-119, Pitman Res. Notes Math. **362**, Longman, Harlow, 1996.

[P] L. A. Pastur, Spectral properties of disordered systems in one-body approximation, *Comm. Math. Phys.* **75** (1980), 179-196.

[Sc] W. Schlag, Regularity and convergence rates for the Lyapunov exponents of linear co-cycles, *J. Mod. Dyn.* **7**(4)(2013), 619-637.

[Sim1] B. Simon, Orthogonal Polynomials on the Unit Circle(Part 1): Classical Theory, Colloquium Publications 54, (American mathematical Society, Providence, 2005).

[Sim2] B. Simon, Orthogonal Polynomials on the Unit Circle(Part 2): Spectral Theory, Colloquium Publications 54, (American mathematical Society, Providence, 2005).

[Sin] Ya. G. Sinai, Anderson localization for one-dimensional difference Schrödinger operator with quasiperiodic potential. *J. Statist. Phys.* **46**(1987), 861-909.

[SoSp] E. Sorets and T. Spencer, Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials, *Commun. Math. Phys.*142(1991), 543-566.

[T] J. Thouvenot, An example of discontinuity in the computation of the Lyapunov exponents, *Proc. Stekolov Inst. Math.* **216**(1997), 366-369.

[WaYo1] Y. Wang and J. You, Examples of discontinuity of Lyapunov exponent in smooth quasi-periodic cocycles, *Duke Math. J.* **162**(2013), 2363-2412.

[WaYo2] Y. Wang and J. You, Examples of Non-openness of Smooth Quasi-Periodic Cocycles with Positive Lyapunov Exponent, preprint (2014).

[WaZ] Y. Wang and Z. Zhang, Cantor spectrum for a class of C^2 quasiperiodic Schrödinger operators, arXiv:1410.0101(2014).

[Y] L. Young, Lyapunov exponents for some quasi-periodic cocycles, *Ergodic Theory and Dynamical Systems*, **17**(1997), 483-504.

[Z1] Z. Zhang, Positive Lyapunov exponents for quasiperiodic Szegő cocycles, *Nonlinearity* **25**(2012), 1771-1797.

[Z2] Z. Zhang, Resolvent set of Schrödinger operators and uniform hyperbolicity, arXiv:1305.4226v2(2013).

Department of Mathematics, Nanjing University, Nanjing 210093, China
E-mail address: `yiqianw@nju.edu.cn`

Department of Mathematics, Northwestern University, Evanston, IL 60208, USA
E-mail address: `zhenghe@math.northwestern.edu`