Acta Math., 215 (2015), 1–54
DOI: 10.1007/s11511-015-0128-7

Global theory of one-frequency Schrödinger operators

by

Artur Avila

Université Paris Diderot
Paris, France

and

Instituto Nacional de Matemática Pura e Aplicada
Rio de Janeiro, Brasil

1. Introduction

This work is concerned with the dynamics of one-frequency SL(2) cocycles, and has two distinct aspects: the analysis, from a new point of view, of the dependence of the Lyapunov exponent with respect to parameters, and the study of the boundary of non-uniform hyperbolicity. But our underlying motivation is to build a global theory of one-frequency Schrödinger operators with general analytic potentials, so we will start from there.

1.1. One-frequency Schrödinger operators

A one-dimensional quasiperiodic Schrödinger operator with one-frequency analytic potential $H=H_{\alpha,v}\colon \ell^2(\mathbb{Z})\to\ell^2(\mathbb{Z})$ is given by

$$(Hu)_n=u_{n+1}+u_{n-1}+v(n\alpha)u_n, \tag{1}$$

where $v\colon \mathbb{R}/\mathbb{Z}\to\mathbb{R}$ is an analytic function (the potential), and $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ is the frequency. We denote by $\Sigma=\Sigma_{\alpha,v}$ the spectrum of H. Despite many recent advances ([BG], [GS1], [B], [BJ1], [BJ2], [AK1], [GS2], [GS3], [AJ], [AFK], [A2]) key aspects of an authentic global theory of such operators have been missing. Namely, progress has been made

I am grateful to Svetlana Jitomirskaya and David Damanik for several detailed comments which greatly improved the exposition. This work was partially conducted during the period the author served as a Clay Research Fellow. This work has been supported by the ERC Starting Grant "Quasiperiodic" and by the Balzan project of Jacob Palis.

mainly into the understanding of the behavior in regions of the spectrum belonging to two regimes with (at least some of the) behavior caracteristic, respectively, of “large” and “small” potentials. But the transition between the two regimes has been considerably harder to understand.

1.1.1. The almost Mathieu operator

Until now, there has been only one case where the analysis has genuinely been carried out at a global level. The almost Mathieu operator, $v(x)=2\lambda\cos 2\pi(\theta+x)$, is a highly symmetric model for which coupling strengths λ and λ^{-1} can be related through the Fourier transform (Aubry duality). Due to this unique feature, it has been possible to establish that the transition happens precisely at the (self-dual) critical coupling $|\lambda|=1$. In the *subcritical regime* $|\lambda|<1$ all energies in the spectrum behave as for small potentials, while in the *supercritical regime* $|\lambda|>1$ all energies in the spectrum behave as for large potentials. Hence typical almost Mathieu operators fall entirely in one regime or the other. Related to this simple phase transition picture is the fundamental spectral result of [J], which implies that the spectral measure of a typical almost Mathieu operator has no singular continuous components (it is either typically atomic for $|\lambda|>1$ or typically absolutely continuous for $|\lambda|<1$).

One precise way to distinguish the subcritical and the supercritical regime for the almost Mathieu operator is by means of the Lyapunov exponent. Recall that for $E\in\mathbb{R}$, a formal solution $u\in\mathbb{C}^{\mathbb{Z}}$ of $Hu=Eu$ can be reconstructed from its values at two consecutive points by application of n-step transfer matrices:

$$A_n(k\alpha)\begin{pmatrix} u_k \\ u_{k-1} \end{pmatrix}=\begin{pmatrix} u_{k+n} \\ u_{k+n-1} \end{pmatrix}. \tag{2}$$

The $A_n:\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{R})$, $n\in\mathbb{Z}$, are analytic functions defined on the same band of analyticity as v and are given in terms of

$$A=\begin{pmatrix} E-v & -1 \\ 1 & 0 \end{pmatrix}$$

by $A_0(\cdot)=\mathrm{id}$ and, for $n\geqslant 1$, by

$$A_n(\cdot)=A(\cdot+(n-1)\alpha)\ldots A(\cdot) \quad \text{and} \quad A_{-n}(\cdot)=A_n(\cdot-n\alpha)^{-1}. \tag{3}$$

The Lyapunov exponent at energy E is denoted by $L(E)$ and given by

$$\lim_{n\to\infty}\frac{1}{n}\int_{\mathbb{R}/\mathbb{Z}}\log\|A_n(x)\|\,dx\geqslant 0. \tag{4}$$

It follows from the Aubry–André formula (proved by Bourgain–Jitomirskaya [BJ1]) that $L(E)=\max\{0,\log|\lambda|\}$ for $E\in\Sigma_{\alpha,v}$. Thus the supercritical regime can be distinguished by the positivity of the Lyapunov exponent: supercritical just means *non-uniformly hyperbolic* in dynamical systems terminology.(1)

How to distinguish subcritical energies from critical ones (since both have zero Lyapunov exponent)? One way could be in terms of their stability: critical energies are in the boundary of the supercritical regime, while subcritical ones are far away. Another, more intrinsic way, consists of looking at the complex extensions of the A_n: it can be shown (by a combination of [J] and [JKS]) that for subcritical energies we have a uniform subexponential bound $\log\|A_n(z)\|=o(n)$ through a band $|\operatorname{Im} z|<\delta(\lambda)$, while for critical energies this is not the case (this follows from [H]). (See also Appendix A for a rederivation of both facts in the spirit of this paper.)

1.1.2. The general case

This work is not concerned with the almost Mathieu operator, whose global theory was constructed around duality and many remarkable exact computations. Still, what we know about it provides a powerful hint about what one can expect from the general theory. By analogy, we can always classify energies in the spectrum of an operator $H_{\alpha,v}$ as supercritical, subcritical or critical in terms of the growth behavior of (complex extensions of) the corresponding transfer matrices A_n. More precisely, $E\in\Sigma_{\alpha,v}$ is said to be

(1) *supercritical* if $\sup_{x\in\mathbb{R}/\mathbb{Z}}\|A_n(x)\|$ grows exponentially,

(2) *subcritical* if there is a uniform subexponential bound on the growth of $\|A_n(z)\|$ through some band $|\operatorname{Im} z|<\varepsilon$, and

(3) *critical* otherwise.

That large potentials fall into the supercritical regime then follows from [SS] and that small potentials fall into the subcritical one is a consequence of [BJ1] and [BJ2]. However, differently from the almost Mathieu case the coexistence of regimes is known to be possible [Bj1]. Thus subcriticality, criticality and supercriticality are not, in general, a property of a whole operator, but of individual energies.

Beyond the local problems of describing precisely the behavior at the supercritical and subcritical regimes, a proper global theory should certainly explain how the phase transition between them occurs, and how this critical set of energies affects the spectral analysis of H.

(1) Recalling that uniform hyperbolicity is well known to characterize the complement of the spectrum.

It was deliberately implied in the discussion above that the non-critical regimes are stable (with respect to perturbations of the energy, potential or frequency), but the critical one is not. Stability of non-uniform hyperbolicity was known (continuity of the Lyapunov exponent [BJ1]), while the stability of the subcritical regime and the instability of the critical regime are obtained here. The stability of the subcritical regime implies that the critical set contains the boundary of the supercritical regime. By a more delicate argument, we will show that any critical energy can be made supercritical under an arbitrarily small perturbation of the potential, and thus identifying the critical set with the boundary of non-uniform hyperbolicity (see Theorem 15).([2])

While a given potential may display both subcritical and supercritical energies (and such coexistence is clearly robust under perturbations of both the potential and the frequency), in order to go from one regime to the other it may not be necessary to pass through the critical regime. This is because the spectrum may be a Cantor set (this is actually what one usually expects), and the transition could thus happen through a gap. In this paper we show that this is the prevalent behavior. Let us say that H is *acritical* if no energy $E\in\Sigma$ is critical.

MAIN THEOREM. *Let α be irrational. Then for a* (*measure-theoretically*) *typical* $v\in C^\omega(\mathbb{R}/\mathbb{Z},\mathbb{R})$, *the operator $H_{\alpha,v}$ is acritical.*

The main theorem yields a precise description of the basic structure of the spectrum of typical operators with respect to the behavior of the Lyapunov exponent. Indeed the stability of the non-critical regimes immediately yields:

(1) Acriticality is stable with respect to perturbations of both the frequency and the potential, that is, the set of $(\alpha,v)\in(\mathbb{R}\setminus\mathbb{Q})\times C^\omega(\mathbb{R}/\mathbb{Z},\mathbb{R})$ such that $H_{\alpha,v}$ is acritical is open. Moreover, the supercritical and subcritical parts of the spectrum define compact sets that depend continuously (in the Hausdorff topology) on (α,v).

(2) As a consequence, acritical operators have the nicest behavior from the point of view of bifurcations: There is at most a finite number of alternances of regime as one moves through the spectrum Σ in the following sense: there are $k\geqslant 1$ and points $a_1<b_1<...<a_k<b_k$ in the spectrum such that $\Sigma\subset\bigcup_{i=1}^k[a_i,b_i]$ and energies alternate between supercritical and subcritical along the sequence $\{\Sigma\cap[a_i,b_i]\}_{i=1}^k$.

(3) Another consequence is *spectral uniformity* through both subcritical and supercritical regimes: There exists $\varepsilon>0$ such that whenever E is supercritical we have $L(E)\geqslant\varepsilon$ (by continuity of the Lyapunov exponent [BJ1]), and when E is subcritical we have uniform subexponential growth of $\|A_n(z)\|$ through the band $|\mathrm{Im}\, z|<\varepsilon$ (again by continuity

([2]) We should note that our work leaves open the question of whether the critical set coincides with the boundary of the subcritical regime as well.

of the Lyapunov exponent, together with a key result obtained here, the quantization of the acceleration).

As we will show in Appendix B, the number of phase transitions can be arbitrarily large.

In developing the work presented here, we were guided by the hope that typical one-frequency operators have nice spectral properties. Particularly, we conjectured early on that typically the spectral measures should have no singular continuous component (which, if present, would be responsible for the most exotic behavior from the point of view of quantum dynamics). We will next describe how our main theorem relates to the goal of establishing a more precise version of this conjecture.

1.2. The spectral dichotomy program

The main theorem reduces the spectral theory of a typical one-frequency Schrödinger operator H to the separate local theories of (uniform) supercriticality and subcriticality. It is thus a key step in our program to establish the *spectral dichotomy*, the decomposition of a typical operator as a direct sum of operators with the spectral type of large-like and small-like operators. Below we comment briefly on the current state of the local theories.

The supercritical theory has been intensively developed in [BG] and [GS1]–[GS3]. As far as the spectral type is concerned, perhaps the key result is that, up to a typical perturbation of the frequency, Anderson localization (pure point spectrum with exponentially decaying eigenfunctions) holds through the supercritical regime. It is important to emphasize that these developments superseded several early results depending on suitable largeness conditions on the potentials, and that the change of focus towards the Lyapunov exponent can be in large part attributed to [J].

The concept of subcriticality has evolved more recently. The development of the corresponding local theory originally centered on the concept of *almost reducibility*, which by definition generalizes the scope of applicability of the theory of small potentials (which is well understood by Kolmogorov–Arnold–Moser theory and localization-duality methods). In particular, it was shown ([AJ], [A1], [A2]) that almost reducibility implies absolute continuity of spectral measures. In [AJ] the vanishing of the Lyapunov exponent in a band was suggested to be the sought after mirror condition to positivity of the Lyapunov exponent. More specifically, it was conjectured to be equivalent to almost reducibility (in the spectrum). Proving this *almost reducibility conjecture* would at once provide an almost complete understanding of subcriticality, and partial results were obtained in [A1] and [A2].

We have recently proved the almost reduciblity conjecture (for all frequencies) [A3]. Together with this work, it implies in particular that typical one-frequency operators have only point spectrum in the supercritical region, and absolutely continuous spectrum in the subcritical region.(3) See §2.1.1 and [A3] for a more detailed account of spectral consequences.

1.3. Prevalence

Let us explain in more detail the notion of typical that appears in the main theorem. Since in infinite-dimensional settings one lacks a translation-invariant measure, it is common to replace the notion of *almost every* by *prevalence*: one fixes some probability measure μ of compact support (a set of admissible perturbations w), and declare a property to be *typical* if it is satisfied for almost every perturbation $v+w$ of every starting condition v. In finite-dimensional vector spaces, prevalence implies full Lebesgue measure.

In our case, we have quite a bit of flexibility for the choice of μ. For instance, though we do want to be able to perturb all Fourier coefficients, we may impose arbitrarily strong restrictions on high Fourier mode perturbations. For definiteness, we will set $\Delta=\mathbb{D}^{\mathbb{N}}$ endowed with the probability measure μ given by the product of normalized Lebesgue measure. Given an arbitrary function $\varepsilon\colon\mathbb{N}\to\mathbb{R}_+$ which decays exponentially fast (the particular choice is quite irrelevant for us), we associate a probability measure μ_ε with compact support on $C^\omega(\mathbb{R}/\mathbb{Z},\mathbb{R})$ by push forward of μ under the map

$$\{t_m\}_{m\in\mathbb{N}}\longmapsto\sum_{m\geqslant 1}\varepsilon(m)2\operatorname{Re}[t_m e^{2\pi imx}].$$

In other words, we will establish the main theorem by showing that, for any $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and every $v\in C^\omega(\mathbb{R}/\mathbb{Z},\mathbb{R})$, the operator $H_{\alpha,v+w}$ is acritical for μ_ε-almost every w.

Remark 1. (1) The notion of prevalence is usually formulated for separable Banach spaces (see [HSY]). Our result does imply prevalence of acriticality in any Banach space of analytic potentials which is continuously and densely embedded in C^ω.

(2) The notion of prevalence (or rather, the corresponding smallness notion called shyness in [HSY]) was first introduced in [C], i.e., the complement of a prevalent set in a Banach space is what is called a *Haar-null set*. There is a stronger notion of smallness (and thus a corresponding stronger notion of typical) which is induced by the family of non-degenerate Gauss measures in a Banach space: Gauss-null sets.(4) In a Banach

(3) Let us emphasize that here typical refers to the entire parameter space, including both the potential and the frequency. Indeed it is known that for frequencies that are very well approximated by rational numbers the supercritical regime can only support singular continuous spectrum.

(4) The author learned this notion from Assaf Naor.

space, a Borel set which has zero probability with respect to any affine embedding of the Hilbert cube (endowed with the natural product measure) which is non-degenerate (i.e., not contained in a proper closed affine subspace) is Gauss-null, see [BL, §6.2]. While we have considered in the description above a particular family of embeddings of $\mathbb{D}^{\mathbb{N}}$, it is transparent from the proof that an arbitrary non-degenerate embedding of the Hilbert cube would work equally well, so acritical potentials are also typical in this stronger sense.

1.4. Main structure of the proof

The proof of the main theorem breaks into two rather distinctive parts (originally presented separately in two preprints which have been merged in this version) corresponding to §2 and §3. The second part uses the concepts and results developed in the first part, so it cannot be read independently.

The main goal of the first part is to establish the following result. Let $C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$ be the real Banach space of analytic functions $\mathbb{R}/\mathbb{Z}\to\mathbb{R}$ admitting a holomorphic extension to $|\mathrm{Im}\, z|<\delta$ which is continuous up to the boundary.

THEOREM 1. *For any $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, the set of potentials and energies (v,E) such that E is a critical energy for $H_{\alpha,v}$ is contained in a countable union of codimension-one analytic submanifolds of $C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})\times\mathbb{R}$.*(5)

Note that this immediately implies that a typical operator H will have at most countably many critical energies. It also shows the instability of the critical regime.

The second part of the proof consists of studying how large the (fractal) critical set is within the subvarieties provided by Theorem 1. We consider finite-dimensional families of pairs (v,E) depending on a large number of parameters, intersecting transversally the subvariety(6) and show that, within such families, the critical set has zero Lebesgue measure inside the subvarieties.

It is perhaps helpful to make a parallel with the almost Mathieu operator, whose potential depends (essentially) on a single parameter, the coupling constant $\lambda>0$ (the so-called phase parameter is inessential for the discussion here). For definiteness, let us fix the frequency $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, so that the parameter space becomes the (λ,E) half-plane. Then the critical set lies in the subvariety given by the equation $\lambda=1$. The fact that the critical set has zero Lebesgue measure within this locus corresponds precisely to the

(5) A codimension-1 analytic submanifold is a (not-necessarily closed) set X given locally (near any point of X) as the zero set of an analytic submersion $C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})\to\mathbb{R}$.

(6) The large number of parameters is necessary to check this transversality condition and another more delicate one that we will discuss in a moment.

Hofstadter conjecture that the spectrum of the critical almost Mathieu operator has zero Lebesgue measure.

For the almost Mathieu operator the equation $\lambda=1$ is of course given by Aubry duality (it is the self-duality condition). The corresponding equations implied in the statement of Theorem 1 have a very different nature: ultimately, we want to define it simply as the boundary of the set $\{E:L(E)>0\}$, which seems at first somewhat foolish given the known irregularity of the Lyapunov exponent L (this will be discussed extensively in §2). What has made this approach at all reasonable was our discovery that the restrictions of L to appropriate (fractal) sets admit nice analytic extensions.

As for the Hofstadter conjecture, there had been many earlier advances based on exact computations for the almost Mathieu operator. It was eventually fully solved in [AK1] by a more abstract approach, based on renormalization. It is this approach that we are able to generalize here. Given our previous work [AK1], [AK2], the main difficulty involves the verification of an appropriate transversality condition for the subvarieties in question. This transversality is necessary to verify a certain monotonicity condition with respect to an appropriate parameter (for the almost Mathieu operator, the monotonicity with respect to the parameter E within the curve $\lambda=1$ is elementary).

2. Part I: Stratified analyticity of the Lyapunov exponent

As discussed above, the Lyapunov exponent L is fundamental in the understanding of the spectral properties of H. It is also closely connected with another important quantity, the *integrated density of states* (i.d.s.), denoted by N. As the Lyapunov exponent, the i.d.s. is a function of the energy. While the Lyapunov exponent measures the asymptotic average growth/decay of solutions (not necessary in ℓ^2) of the equation $Hu=Eu$, the integrated density of states gives the asymptotic distribution of eigenvalues of H restricted to large boxes. The two are related by the Thouless formula:

$$L(E)=\int \log|E'-E|\,dN(E'). \tag{5}$$

Much work has been dedicated to the regularity properties of L and N. For quite general reasons, the integrated density of states is a continuous non-decreasing function onto $[0,1]$, which is constant outside the spectrum. Notice that this is not enough to conclude continuity of the Lyapunov exponent from the Thouless formula. Other regularity properties (such as Hölder) do pass from N to L and vice-versa. This being said, our focus here is primarily on the Lyapunov exponent on its own.

It is easy to see that the Lyapunov exponent is real-analytic outside the spectrum. Beyond that, however, there are obvious limitations to its regularity. For a constant potential, say $v=0$, the Lyapunov exponent is given by $\max\{0, \log\frac{1}{2}(E+\sqrt{E^2-4})\}$, so it is only $\frac{1}{2}$-Hölder continuous. With Diophantine frequencies and small potentials, the generic situation is to have Cantor spectrum with countably many square root singularities at the endpoints of gaps [E]. For small potential and generic frequencies, it is possible to show that the Lyapunov exponent escapes any fixed continuity modulus (such as Hölder), and also it is not of bounded variation. More delicately, Bourgain [B] has observed that in the case of the critical almost Mathieu operator the Lyapunov exponent need not be Hölder continuous even for Diophantine frequencies (another instance of complications arising at the boundary of non-uniform hyperbolicity). Though a surprising result, analytic regularity was obtained in a related but non-Schrödinger context [AK2]. However, the negative results described above seemed to impose serious limitations on the amount of regularity one should even try to look for in the Schrödinger case.

As for positive results, a key development was the proof by Goldstein–Schlag [GS1] that the Lyapunov exponent is Hölder continuous for Diophantine frequencies in the regime where the Lyapunov exponent is positive. Later Bourgain–Jitomirskaya [BJ1] proved that the Lyapunov exponent is continuous for all irrational frequencies, and this result played a fundamental role in the recent theory of the almost Mathieu operator. More delicate estimates on the Hölder regularity for Diophantine frequencies remained an important topic of the local theories (see [GS2] and [AJ]).

There is however one important case where, in a different sense, much stronger regularity holds. For small analytic potentials, it follows from the work of Bourgain-Jitomirskaya ([BJ1], [BJ2]; see also [AJ]) that the Lyapunov exponent is zero (and hence constant) in the spectrum. In general, however, the Lyapunov exponent need not be constant in the spectrum. In fact, there are examples where the Lyapunov exponent vanishes in part of the spectrum and is positive in some other part [Bj2]. Particularly in the positive Lyapunov exponent regime it would seem unreasonable, given the negative results outlined above, to expect much more regularity. In fact, from a dynamical systems perspective, it would be natural to expect bad behavior in this setting, since when the Lyapunov exponent is positive, the associated dynamical system in the two-torus presents strange attractors with very complicated dependence of the parameters [Bj1].

In this respect, the almost Mathieu operator would seem to behave quite oddly. As we have seen, by the Aubry–André formula, the Lyapunov exponent is always constant in the spectrum. Moreover, this constant is just a simple expression of the coupling $\max\{0, \log\lambda\}$ (in particular, it is positive in the supercritical regime $\lambda>1$). It remains true that the Lyapunov exponent displays wild oscillations just outside the spectrum, so this is not inconsistent with the negative results discussed above.

However, for a long time, the general feeling has been that this just reinforces the special status of the almost Mathieu operator (and its remarkable but specific symmetry, Aubry duality, relating the supercritical and the subcritical regimes), and such a phenomenon would seem to have little to do with the case of general potentials. This general feeling is wrong, as the following sample result shows.

Example theorem. *Let $\lambda>1$ and let w be any real-analytic function. For $\varepsilon\in\mathbb{R}$, let $v(x)=2\lambda\cos 2\pi x+\varepsilon w(x)$. Then, for ε small enough and for every $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, the Lyapunov exponent restricted to the spectrum is a positive real-analytic function.*

Of course by a real-analytic function on a set we just mean the restriction of some real-analytic function defined on an open neighborhood.

For an arbitrary real-analytic potential, the situation is just slightly lengthier to describe. Let X be a topological space. A *stratification* of X is a strictly decreasing finite or countable sequence of closed sets $X=X_0\supset X_1\supset\dots$ such that $\bigcap_i X_i=\varnothing$. We call $X_i\setminus X_{i+1}$ the ith *stratum* of the stratification.

Let now X be a subset of a real-analytic manifold, and let $f\colon X\to\mathbb{R}$ be a continuous function. We say that f is *C^r-stratified* if there exists a stratification such that the restriction of f to each stratum is C^r.

Theorem 2. (Stratified analyticity in the energy) *Let $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and v be any real-analytic function. Then the Lyapunov exponent is a C^ω-stratified function of the energy.*

We will see that in this theorem the stratification starts with $X_1=\Sigma_{\alpha,v}$, which is compact, so the stratification is finite.

Nothing restricts us to look only at the energy as a parameter. For instance, in the case of the almost Mathieu operator, the Lyapunov exponent (restricted to the spectrum) is real-analytic also in the coupling constant, except at $\lambda=1$.

Theorem 3. (Stratified analyticity in the potential) *Let $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and X be a real-analytic manifold. Also let v_λ, $\lambda\in X$, be a real-analytic family of real-analytic potentials. Then the Lyapunov exponent is a C^ω-stratified function of both λ and E.*

It is quite clear how this result opens up for the analysis of the boundary of non-uniform hyperbolicity, since parameters corresponding to the vanishing of the Lyapunov exponent are contained in the set of solutions of equations (in infinitely many variables) with analytic coefficients. Of course, one still has to analyze the nature of the equations one gets, guaranteeing the non-vanishing of the coefficients. Indeed, in the subcritical regime, the coefficients do vanish. In what follows, we will work out suitable expressions for the Lyapunov exponent, restricted to strata, which will allow us to show the non-vanishing outside the subcritical regime.

In the case of the almost Mathieu operator, there is no dependence of the Lyapunov exponent on the frequency parameter. In general, Bourgain–Jitomirskaya [BJ1] proved that the Lyapunov exponent is a continuous function of $\alpha\in\mathbb{R}\setminus\mathbb{Q}$. This is a very subtle result, as the continuity is not in general uniform in α. We will show that the Lyapunov exponent is in fact C^∞-stratified as a function of $\alpha\in\mathbb{R}\setminus\mathbb{Q}$.

THEOREM 4. *Let X be a real-analytic manifold, and v_λ, $\lambda\in X$, be a real-analytic family of real-analytic potentials. Then the Lyapunov exponent is a C^∞-stratified function of $(\alpha,\lambda,E)\in(\mathbb{R}\setminus\mathbb{Q})\times X\times\mathbb{R}$.*

With v as in the example theorem, the Lyapunov exponent is actually C^∞ as a function of α and E in the spectrum.

2.1. Lyapunov exponents of $\mathrm{SL}(2,\mathbb{C})$ cocycles

In the dynamical systems approach, which we follow here, the understanding of the Schrödinger operator is obtained through the detailed description of a certain family of dynamical systems.

A (one-frequency, analytic) quasiperiodic $\mathrm{SL}(2,\mathbb{C})$ cocycle is a pair (α,A), where $\alpha\in\mathbb{R}$ and $A:\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{C})$ is analytic, understood as defining a linear skew-product acting on $\mathbb{R}/\mathbb{Z}\times\mathbb{C}^2$ by $(x,w)\mapsto(x+\alpha,A(x)\cdot w)$. The iterates of the cocycle have the form $(n\alpha,A_n)$ where A_n is given by (3). The Lyapunov exponent $L(\alpha,A)$ of the cocycle (α,A) is given by the left hand side of (4). We say that (α,A) is *uniformly hyperbolic* if there exist analytic functions $u,s:\mathbb{R}/\mathbb{Z}\to\mathbb{P}\mathbb{C}^2$, called the *unstable* and *stable directions*, and $n\geqslant1$ such that

(i) for every $x\in\mathbb{R}/\mathbb{Z}$ we have $A(x)\cdot u(x)=u(x+\alpha)$ and $A(x)\cdot s(x)=s(x+\alpha)$,

(ii) for every unit vector $w\in s(x)$ we have $\|A_n(x)\cdot w\|<1$,

(iii) for every unit vector $w\in u(x)$ we have $\|A_n(x)\cdot w\|>1$.[7]

The unstable and stable directions are uniquely characterized by these properties, and clearly $u(x)\neq s(x)$ for every $x\in\mathbb{R}/\mathbb{Z}$. It is also clear that, if (α,A) is uniformly hyperbolic, then $L(\alpha,A)>0$.

If $L(\alpha,A)>0$ but (α,A) is not uniformly hyperbolic, we will say that (α,A) is *non-uniformly hyperbolic*.

---

(7) This is one of several equivalent definitions of uniform hyperbolicity in this context, for instance the unstable and stable directions could have been assumed to be merely continuous (since analyticity is automatic).

Uniform hyperbolicity is a stable property: the set $\mathcal{UH}\subset\mathbb{R}\times C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ of uniformly hyperbolic cocycles is open. Moreover, it implies that the Lyapunov exponent is well behaved, in the sense that the restriction of $(\alpha,A)\mapsto L(\alpha,A)$ to $\mathcal{UH}$ is a C^∞ function of both variables,([8]) and it is a pluriharmonic function of the second variable.([9]) In fact regularity properties of the Lyapunov exponent are a consequence of the regularity of the unstable and stable directions, which depend smoothly on both variables (by normally hyperbolic theory [HPS]) and holomorphically on the second variable (by a simple normality argument).

On the other hand, a variation of [BJ1] (see [JKS]) gives that $(\alpha,A)\mapsto L(\alpha,A)$ is continuous as a function on $(\mathbb{R}\setminus\mathbb{Q})\times C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$. It is important to notice (and in fact, fundamental in what follows) that the Lyapunov exponent is not continuous on $\mathbb{R}\times C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$. In the remainder of this section we will restrict our attention (except otherwise noted) to cocycles with irrational frequencies.

The most important examples are Schrödinger cocycles $A^{(v)}$, determined by a real-analytic function v by

$$A^{(v)}=\begin{pmatrix} v & -1\\ 1 & 0\end{pmatrix}.$$

In this notation, the Lyapunov exponent at energy E for the operator $H_{\alpha,v}$ becomes $L(E)=L(\alpha,A^{(E-v)})$. One of the most basic aspects of the connection between spectral and dynamical properties is that $E\notin\Sigma_{\alpha,v}$ if and only if $(\alpha,A^{(E-v)})$ is uniformly hyperbolic. Thus the analyticity of $E\mapsto L(E)$ outside of the spectrum just translates a general property of uniformly hyperbolic cocycles.

If $A\in C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ admits a holomorphic extension to $|\mathrm{Im}\, z|<\delta$, then for $|\varepsilon|<\delta$ we can define $A_\varepsilon\in C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ by $A_\varepsilon(x)=A(x+i\varepsilon)$. The Lyapunov exponent $L(\alpha,A_\varepsilon)$ is easily seen to be a convex function of ε. Thus we may define a function

$$\omega(\alpha,A)=\lim_{\varepsilon\to 0+}\frac{1}{2\pi\varepsilon}(L(\alpha,A_\varepsilon)-L(\alpha,A)), \tag{6}$$

called *acceleration*. By the convexity and the continuity of the Lyapunov exponent, the acceleration is an upper semicontinuous function in $(\mathbb{R}\setminus\mathbb{Q})\times C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$.

Our starting point is the following result.

THEOREM 5. (Acceleration is quantized) *The acceleration of an* $\mathrm{SL}(2,\mathbb{C})$ *cocycle with irrational frequency is always an integer.*

([8]) Since $\mathcal{UH}$ is not a Banach manifold, it might seem important to be precise about what notion of smoothness is used here. This issue can be avoided by enlarging the setting to include C^∞ non-analytic cocycles (say by considering a Gevrey condition), so that we end up with a Banach manifold.

([9]) This means that, in addition to being continuous, given any family $\lambda\mapsto A^{(\lambda)}\in\mathcal{UH}$, $\lambda\in\mathbb{D}$, which is holomorphic (in the sense that it is continuous and for every $x\in\mathbb{R}/\mathbb{Z}$ the map $\lambda\mapsto A^{(\lambda)}(x)$ is holomorphic), the map $\lambda\mapsto L(\alpha,A^{(\lambda)})$ is harmonic.

Remark 2. It is easy to see that quantization does not extend to rational frequencies, see Remark 7.

This result allows us to break the parameter space into suitable pieces restricted to which we can study the behavior of the Lyapunov exponent.

Quantization implies that $\varepsilon \mapsto L(\alpha, A_\varepsilon)$ is a piecewise affine function of ε. Knowing this, it makes sense to introduce the following definition.

Definition 3. We say that $(\alpha, A) \in (\mathbb{R} \setminus \mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2, \mathbb{C}))$ is *regular* if $L(\alpha, A_\varepsilon)$ is affine for ε in a neighborhood of 0.

Remark 4. If A takes values in $\mathrm{SL}(2, \mathbb{R})$ then $\varepsilon \mapsto L(\alpha, A_\varepsilon)$ is an even function. By convexity, we have $\omega(\alpha, A) \geqslant 0$. Further, if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then (α, A) is regular if and only if $\omega(\alpha, A)=0$.

Clearly regularity is an open condition in $(\mathbb{R} \setminus \mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2, \mathbb{C}))$.

It is natural to assume that regularity has important consequences for the dynamics. Indeed, we have been able to completely characterize the dynamics of regular cocycles with positive Lyapunov exponent, which is the other cornerstone of this section.

THEOREM 6. *Let* $(\alpha, A) \in (\mathbb{R} \setminus \mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2, \mathbb{C}))$. *Assume that* $L(\alpha, A)>0$. *Then* (α, A) *is regular if and only if* (α, A) *is uniformly hyperbolic.*

One striking consequence is the following result.

COROLLARY 7. *For any* $(\alpha, A) \in (\mathbb{R} \setminus \mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2, \mathbb{C}))$, *there exists* $\varepsilon_0>0$ *such that either*

(1) $L(\alpha, A_\varepsilon)=0$ *(and* $\omega(\alpha, A)=0$*) for every* $0<\varepsilon<\varepsilon_0$, *or*

(2) (α, A_ε) *is uniformly hyperbolic for every* $0<\varepsilon<\varepsilon_0$.

Proof. Since $\varepsilon \mapsto L(\alpha, A_\varepsilon)$ is piecewise affine, it must be affine on $(0, \varepsilon_0)$ for $\varepsilon_0>0$ sufficiently small, and thus (α, A_ε) is regular for every $0<\varepsilon<\varepsilon_0$.

Since the Lyapunov exponent is non-negative, if $L(\alpha, A_\varepsilon)>0$ for some $0<\varepsilon<\varepsilon_0$, then $L(\alpha, A_\varepsilon)>0$ for every $0<\varepsilon<\varepsilon_0$. The result follows from the previous theorem. □

This result plays an important role in §3 (and also in [A3]), since it allows us to consider the dynamics of non-regular $\mathrm{SL}(2, \mathbb{R})$-cocycles as a non-tangential limit of better behaved (uniformly hyperbolic) cocycles.

2.1.1. Almost reducibility

The case of regular cocycles with zero Lyapunov exponent is the topic of the almost reducibility conjecture, which we already discussed in the introduction. For completeness, let us state it precisely here.

Conjecture 1. (Almost reducibility conjecture) Let the frequency α be irrational and let $A \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$. If (α, A) is regular and $L(\alpha, A)=0$ then (α, A) is *almost reducible*: There exist $\delta>0$, a constant $A_* \in \mathrm{SL}(2,\mathbb{R})$ and a sequence of analytic maps $B_n \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{PSL}(2,\mathbb{R}))$ such that A and the maps B_n extend holomorphically to the band $|\mathrm{Im}\, x|<\delta$ and $\sup_{|\mathrm{Im}\, x|<\delta} \|B_n(x+\alpha)A(x)B_n(x)^{-1}-A_*\| \to 0$ as $n\to\infty$.

This conjecture was first made in [AJ], and it can be generalized to $\mathrm{SL}(2,\mathbb{C})$-valued cocycles in the obvious way. What makes it so central is that almost reducibility was analyzed in much detail in recent works, see [AJ], [A1], [AFK] and [A2], so a proof would immediately give a very fine picture of the subcritical regime. In particular, coupled with the results of this paper concerning the critical regime and the results of Bourgain, Goldstein and Schlag in the supercritical regime, the almost reducibility conjecture allows one to conclude the typical absence of singular continuous spectrum:

(1) The almost reducibility conjecture implies that the subcritical regime can only support absolutely continuous spectrum [A2];

(2) [BG] implies that pure point spectrum is typical throughout the supercritical regime;([10])

(3) The main theorem implies that typically the critical regime does not appear at all.([11])

Previous to this work (see [A2]), we had already established the almost reducibility conjecture when α is exponentially well approximated by rational numbers (so that if p_n/q_n are the continued fraction approximants, we have $\limsup_{n\to\infty}(\log q_{n+1})/q_n>0$). Coupled with [AJ] and [A1], this proved the almost reducibility conjecture in the case of the almost Mathieu operator. A complete proof of the almost reducibility conjecture was obtained after this work was completed (see [A3]), and the results obtained here about regular cocycles with positive Lyapunov exponent play an important role.

2.1.2. Stratified regularity: proof of Theorems 2–4

We now turn to the deduction of regularity properties of the Lyapunov exponent from Theorems 5 and 6.

([10]) More precisely, for every fixed potential, and for almost every frequency, the spectrum is pure point with exponentially decaying eigenfunctions throughout the region of the spectrum where the Lyapunov exponent is positive.

([11]) In fact, here it is enough to use Theorem 1. We just sketch the argument. For fixed frequency, Theorem 1 implies that a typical potential admits at most countably many critical energies. Considering phase changes $v_\theta(x)=v(x+\theta)$, which do not change the critical set, we see that for almost every θ the critical set, being a fixed countable set, cannot carry any spectral weight (otherwise the average over θ of the spectral measures would have atoms, but this average has a continuous distribution, namely the integrated density of states [AS]).

For $\delta>0$, denote by $C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ the set of all $A\in C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ which admit a bounded holomorphic extension to $|\mathrm{Im}\, z|<\delta$ that is continuous up to the boundary. This set is naturally endowed with a complex Banach manifold structure.

For $j\neq 0$, let $\Omega_{\delta,j}$ be the set of all $(\alpha,A)\in\mathbb{R}\times C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ such that there exists $0<\delta'<\delta$ such that $(\alpha,A_{\delta'})\in\mathcal{UH}$ and $\omega(\alpha,A_{\delta'})=j$. Also, let $L_{\delta,j}:\Omega_{\delta,j}\to\mathbb{R}$ be given by

$$L_{\delta,j}(\alpha,A)=L(\alpha,A_{\delta'})-2\pi j\delta'. \tag{7}$$

If $0<\delta'<\delta''<\delta$, then $\omega(\alpha,A_{\delta'})=\omega(\alpha,A_{\delta''})=j$ implies that

$$L(\alpha,A_{\delta'})=L(\alpha,A_{\delta''})-2\pi j(\delta''-\delta'),$$

and thus we see that $L_{\delta,j}$ is well defined.

Proposition 5. *The set $\Omega_{\delta,j}$ is open and $(\alpha,A)\mapsto L_{\delta,j}(\alpha,A)$ is a C^∞ function, pluriharmonic in the second variable. Moreover, if $(\alpha,A)\in(\mathbb{R}\setminus\mathbb{Q})\times C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{C}))$ has acceleration j, then $(\alpha,A)\in\Omega_{\delta,j}$ and $L(\alpha,A)=L_{\delta,j}(\alpha,A)$.*

Proof. The first part follows from the openness of $\mathcal{UH}$ and the regularity of the Lyapunov exponent restricted to $\mathcal{UH}$. For the second part, we use Corollary 7 and upper semicontinuity of the Lyapunov exponent to conclude that $\omega(\alpha,A)=j$ implies that $(\alpha,A_{\delta'})\in\mathcal{UH}$ and has acceleration j for every δ' sufficiently small, which also gives $L(\alpha,A)=L(\alpha,A_{\delta'})-2\pi j\delta'$. □

We can now explain the proofs of Theorems 2–4. For definiteness, we will consider Theorem 3, the argument is exactly the same for the other theorems. Define a stratification of the parameter space $\mathbb{X}=\mathbb{R}\times X$: let $\mathbb{X}_0=\mathbb{X}$, $\mathbb{X}_1\subset\mathbb{X}_0$ be the set of (E,λ) such that $(\alpha,A^{(E-v_\lambda)})$ is not uniformly hyperbolic, and $\mathbb{X}_j\subset\mathbb{X}_1$, for $j\geqslant 2$, be the set of (E,λ) such that $\omega(\alpha,A^{(E-v_\lambda)})\geqslant j-1$.

Since uniform hyperbolicity is open and the acceleration is upper semicontinuous, each $\mathbb{X}_j$ is closed, so this is indeed a stratification. Since the 0-th stratum $\mathbb{X}_0\setminus\mathbb{X}_1$ corresponds to uniformly hyperbolic cocycles, the Lyapunov exponent is analytic there.

By quantization, the jth stratum, for $j\geqslant 2$, corresponds to cocycles which are not uniformly hyperbolic and have acceleration $j-1$. For each (E,λ_0) in such a stratum, choose $\delta>0$ such that $\lambda\mapsto A^{(v_\lambda)}$ is an analytic function in a neighborhood of λ_0, taking values in $C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{R}))$. The analyticity of the Lyapunov exponent restricted to the stratum is then a consequence of Proposition 5.

As for a parameter (E,λ) in the first stratum $\mathbb{X}_1\setminus\mathbb{X}_2$, quantization implies that $(\alpha,A^{(E-v_\lambda)})$ has non-positive acceleration, so by Remark 4, $(\alpha,A^{(E-v_\lambda)})$ must be regular with zero acceleration. Since it is not uniformly hyperbolic, Theorem 6 implies that $L(\alpha,A^{(E-v_\lambda)})=0$. Thus the Lyapunov exponent is in fact identically 0 in the first stratum.

2.1.3. Codimensionality of critical cocycles

Non-regular cocycles split into two groups, the ones with positive Lyapunov exponent (non-uniformly hyperbolic cocycles), and the ones with zero Lyapunov exponent, which we call *critical cocycles.*(12)

As discussed before, the first group has recently been extensively studied ([BG], [GS1]–[GS3]). However, very little is known about the second one.

Though our methods provide no new information on the dynamics of critical cocycles, they are perfectly adapted to show that critical cocycles are rare. This is somewhat surprising, since in dynamical systems it is rarely the case that the success of parameter exclusion precedes a detailed control of the dynamics.

Of course, for $\mathrm{SL}(2,\mathbb{C})$ cocycles, our previous results already show that critical cocycles are rare in certain one-parameter families, since for every (α, A) and any $\delta \neq 0$ small enough, (α, A_δ) is regular, and hence not critical. But for our applications we are mostly concerned with $\mathrm{SL}(2,\mathbb{R})$-valued cocycles, and even more specifically with Schrödinger cocycles.

If $(\alpha, A) \in (\mathbb{R} \setminus \mathbb{Q}) \times C^\omega_\delta(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$ is critical with acceleration j, then $(\alpha, A) \in \Omega_{\delta,j}$ and $L_{\delta,j}=0$. Moreover, if A is $\mathrm{SL}(2,\mathbb{R})$-valued, criticality implies that the acceleration is positive (see Remark 4). So the locus of critical $\mathrm{SL}(2,\mathbb{R})$-valued cocycles is covered by countably many analytic sets $L^{-1}_{\delta,j}(0)$. Thus the main remaining issue is to show that the functions $L_{\delta,j}$ are non-degenerate.

Theorem 8. *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $\delta > 0$ and $j > 0$. If $v_* \in C^\omega_\delta(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ and $\omega(\alpha, A^{(v_*)}) = j$, then $v \mapsto L_{\delta,j}(\alpha, A^{(v)})$ is a submersion in a neighborhood of v_*.*

This theorem immediately implies Theorem 1.

We are also able to show non-degeneracy in the case of non-Schrödinger cocycles, see Remark 13. Although the derivative of $L_{\delta,j}$ may vanish, this forces the dynamics to be particularly nice, and it can be shown that the second derivative is non-vanishing.

2.1.4. Further comments

As mentioned before, it follows from the combination of [BJ1] and [BJ2] that the Lyapunov exponent is zero in the spectrum, provided the potential is sufficiently small, independently of the frequency. This is a very surprising result from the dynamical point of view.

---

(12) As explained before, this terminology is consistent with the almost Mathieu operator terminology. It turns out that if $v(x) = 2\lambda \cos 2\pi(\theta + x)$, $\lambda \in \mathbb{R}$, then $(\alpha, A^{(E-v)})$ is critical if and only if $\lambda = 1$ and $E \in \Sigma_{\alpha,v}$.

For instance, fix some non-constant small v and consider α close to 0. Then the spectrum is close, in the Hausdorff topology, to the interval $[\inf v-2, \sup v+2]$. However, if $E\notin[\sup v-2, \inf v+2]$, we have

$$\lim_{n\to\infty}\lim_{\alpha\to 0}\frac{1}{n}\int_{\mathbb{R}/\mathbb{Z}}\log\|A^{(E-v)}(x+(n-1)\alpha)\ldots A^{(E-v)}(x)\|\,dx>0. \tag{8}$$

At first it might seem that as $\alpha\to 0$ the dynamics of $(\alpha, A^{(E-v)})$ become increasingly complicated and we should expect the behavior of large potentials (with positive Lyapunov exponents by [SS]).[13] Somehow, delicate cancellation between expansion and contraction takes place precisely at the spectrum and kills the Lyapunov exponent.

Bourgain's and Jitomirskya's result that the Lyapunov exponent must be zero on the spectrum in this situation involves duality and localization arguments which are far from the dynamical point of view. Our work provides a different explanation for it, and extends it from $\mathrm{SL}(2,\mathbb{R})$-cocycles to $\mathrm{SL}(2,\mathbb{C})$-cocycles. Indeed, quantization implies that all cocycles near constant have zero acceleration. Thus they are all regular. Thus if A is close to constant and (α, A) has a positive Lyapunov exponent then it must be uniformly hyperbolic.

We stress that while this argument explains why constant cocycles are far from non-uniform hyperbolicity, localization methods remain crucial to the understanding of several aspects of the dynamics of cocycles close to a constant one, at least in the Diophantine regime.

Let us finally make a few remarks and pose questions about the actual values taken by the acceleration.

(1) If the coefficients of A are trigonometric polynomials of degree at most n, then $|\omega(\alpha, A)|\leqslant n$ by convexity (since $L(\alpha, A_\varepsilon)\leqslant\sup_{x\in\mathbb{R}/\mathbb{Z}}\log\|A(x+\varepsilon i)\|\leqslant 2\pi n\varepsilon+O(1)$).

(2) On the other hand, if $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, $|\lambda|\geqslant 1$ and $n\in\mathbb{N}$, then for $v(x)=2\lambda\cos 2\pi nx$ we have that $\omega(\alpha, A^{(E-v)})=n$ for every $E\in\Sigma_{\alpha,v}$. In the case $n=1$ (the almost Mathieu operator), this is shown in Appendix A. The general case reduces to the case $n=1$, since for any $A\in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$ and $n\in\mathbb{N}$ we have that $L(n\alpha, A(x))=L(\alpha, A(nx))$, which implies that $n\omega(n\alpha, A(x))=\omega(\alpha, A(nx))$.

(3) If $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and A takes values in $\mathrm{SO}(2,\mathbb{R})$, the acceleration is easily seen to be the norm of the topological degree of A. The results of [AK2] imply that this also holds for premonotonic cocycles which include small $\mathrm{SL}(2,\mathbb{R})$ perturbations of $\mathrm{SO}(2,\mathbb{R})$-valued cocycles with non-zero topological degree.

([13]) In fact, the Lyapunov exponent function converges in the L^1-sense, as $\alpha\to 0$, to a continuous function, which is positive outside $[\sup v-2, \inf v+2]$ (see the argument of [AD]). This is compatible with the fact that the edges of the spectrum (located in two small intervals of size $\sup v-\inf v$) become increasingly thinner (in measure) as $\alpha\to 0$.

(4) It seems plausible that the norm of the topological degree is always a lower bound for the acceleration of $\mathrm{SL}(2,\mathbb{R})$ cocycles. In case of non-zero degree, is this bound achieved precisely by premonotonic cocycles?

(5) Consider a typical perturbation of the potential $2\lambda\cos 2\pi nx$, $\lambda>1$. Do energies with any fixed acceleration $1\leqslant k\leqslant n$ form a set of positive measure? It seems promising to use the Benedicks–Carleson method of Young [Y] to address aspects of this question ($k=n$ and large λ, allowing exclusion of a small set of frequencies). One is also tempted to relate the acceleration to the number of critical points for the dynamics (which can be identified when her method works). Collisions between a few critical points might provide a mechanism for the appearance of energies with intermediate acceleration.

2.1.5. Outline of the remaining of this section

The outstanding issues are the proofs of Theorems 5, 6 and 8 (besides the proof of the example theorem, which will be left for Appendix A).

We first address quantization (Theorem 5) in §2.2. The proof uses periodic approximation. A Fourier series estimate shows that, as the denominators of the approximations grow, the quantization becomes more and more pronounced. The result then follows by continuity of the Lyapunov exponent [JKS].

Next we show, in §2.3, that regularity with positive Lyapunov exponent implies uniform hyperbolicity (the hard part of Theorem 6). The proof again proceeds by periodic approximation. We first notice that the Fourier series estimate implies that periodic approximants are uniformly hyperbolic, and hence have unstable and stable directions. If we can show that we can take an analytic limit of those directions, then the uniform hyperbolicity of (α, A) will follow. A simple normality argument shows that we only need to prove that the invariant directions do not get too close as the denominators grow. We show (by direct computation) that if they would get too close, then the derivative of the Lyapunov exponent would be relatively large with respect to perturbations of some Fourier modes of the potential. This contradicts a macroscopic bound on the derivative which comes from pluriharmonicity.

We then show, in §2.4, the non-vanishing of the derivative of the canonical analytic extension of the Lyapunov exponent $L_{\delta,j}$ (Theorem 8). Under the hypothesis that $\omega(\alpha, A^{(v_*)})=j>0$, we get that $(\alpha, A^{(v_*)}_{\delta'})\in\mathcal{UH}$ for $0<\delta'<\delta_0$ ($0<\delta_0<\delta$ small), so we can define holomorphic invariant directions u and s, over $0<\operatorname{Im} z<\delta_0$. Using the explicit expressions for the derivative of the Lyapunov exponent in terms of the unstable and stable directions u and s, derived in §2.3, we conclude that the vanishing of the derivative would imply a symmetry of Fourier coefficients (of a suitable expression involving u

and s), which is enough to conclude that u and s can be analytically continued through $\operatorname{Im} z=0$. This implies that $(\alpha, A^{(v_*)})$ is conjugate to a cocycle of rotations, which implies that its acceleration is zero, contradicting the hypothesis.

Remark 6. We should point out that, after this work was completed, a different approach to some of the results in this section was developed in [AJS]. While their approach is more readily generalizable to higher-dimensional cocycles, the route taken here gives some extra information which will be crucial in the proof of the main theorem (and also in the proof of the almost reducibility conjecture given in [A3]).

2.2. Quantization of acceleration: proof of Theorem 5

We will use the continuity properties of the Lyapunov exponent (particularly, with respect to the frequency) obtained in [BJ1] and [JKS].(14)

THEOREM 9. ([JKS]) *The map* $(\alpha, A)\mapsto L(\alpha, A)$, $(\alpha, A)\in\mathbb{R}\times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$, *is continuous at every* (α, A) *with* $\alpha\in\mathbb{R}\setminus\mathbb{Q}$.

This result is very delicate, since the restriction of $\alpha\mapsto L(\alpha, A)$ to $\mathbb{R}\setminus\mathbb{Q}$ is not, in general, uniformly continuous.

Notice that if p/q is a rational number, then there exists a simple expression for the Lyapunov exponent $L(p/q, A)$:

$$L\left(\frac{p}{q}, A\right) = \frac{1}{q}\int_{\mathbb{R}/\mathbb{Z}} \log \varrho(A_{(p/q)}(x))\, dx, \tag{9}$$

where $A_{(p/q)}(x)=A(x+(q-1)p/q)\, ...\, A(x)$ and $\varrho(B)$ is the spectral radius of an $\mathrm{SL}(2,\mathbb{C})$ matrix $\varrho(B)=\lim_{n\to\infty}\|B^n\|^{1/n}$. A key observation is that if p and q are coprime, then the trace $\operatorname{tr} A_{(p/q)}(x)$ is a $1/q$-periodic function of x. This follows from the relation

$$A(x)A_{(p/q)}(x) = A_{(p/q)}\left(x+\frac{p}{q}\right)A(x), \tag{10}$$

expressing the fact that $A_{(p/q)}(x)$ and $A_{(p/q)}(x+p/q)$ are conjugate in $\mathrm{SL}(2,\mathbb{C})$, and hence $A_{(p/q)}(x)$ is conjugate to $A_{(p/q)}(x+kp/q)$ for any $k\in\mathbb{Z}$.

Fix $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and $A\in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$ and let p_n/q_n be a sequence of rational numbers (p_n and q_n coprime) approaching α (not necessarily continued fraction approximants).

(14) Bourgain and Jitomirskaya actually restricted considerations to the case of Schrödinger (in particular $\mathrm{SL}(2,\mathbb{R})$-valued) cocycles. Their result was generalized to the $\mathrm{SL}(2,\mathbb{C})$ case in the work of Jitomirskaya, Koslover and Schulteis [JKS].

Let $\varepsilon>0$ and $C>0$ be such that A admits a bounded extension to $|\mathrm{Im}\, z|<\varepsilon$ with $\sup_{|\mathrm{Im}\, z|<\varepsilon}\|A(z)\|<C$. Since $\mathrm{tr} A_{(p_n/q_n)}$ is $(1/q_n)$-periodic,

$$\mathrm{tr} A_{(p_n/q_n)}(x)=\sum_{k\in\mathbb{Z}} a_{k,n} e^{2\pi i k q_n x}, \tag{11}$$

with $a_{k,n}\leqslant 2C^{q_n}e^{-2\pi k q_n \varepsilon}$.

Fix $0<\varepsilon'<\varepsilon$. Choosing k_0 sufficiently large, we get

$$\mathrm{tr} A_{(p_n/q_n)}(x)=\sum_{|k|\leqslant k_0} a_{k,n} e^{2\pi i k q_n x}+O(e^{-q_n}), \quad |\mathrm{Im}\, x|<\varepsilon', \tag{12}$$

for n large. Since $\max\bigl\{1,\tfrac{1}{2}|\mathrm{tr}|\bigr\}\leqslant\varrho\leqslant\max\{1,|\mathrm{tr}|\}$, it follows that

$$L\biggl(\frac{p_n}{q_n},A_\delta\biggr)=\max_{k\leqslant|k_0|}\max\biggl\{\frac{\log|a_{k,n}|}{q_n}-2\pi k\delta,0\biggr\}+o(1), \quad \delta<\varepsilon'. \tag{13}$$

Thus, for large n, the function $\delta\mapsto L(p_n/q_n,A_\delta)$ is close, for $|\delta|<\varepsilon'$, to a convex piecewise linear function with slopes in $\{-2\pi k_0,...,2\pi k_0\}$. By Theorem 9, these functions converge uniformly on compact subsets of $|\delta|<\varepsilon$ to $\delta\mapsto L(\alpha,A_\delta)$. It follows that $\delta\mapsto L(\alpha,A_\delta)$ is a convex piecewise linear function of $|\delta|<\varepsilon'$, with slopes in $\{-2\pi k_0,...,2\pi k_0\}$, so $\omega(\alpha,A)\in\mathbb{Z}$. This completes the proof of Theorem 5.

Remark 7. Consider say

$$A(x)=\begin{pmatrix} e^{\lambda(x)} & 0\\ 0 & e^{-\lambda(x)}\end{pmatrix},$$

with $\lambda(x)=e^{2\pi i q_0 x}$ for some $q_0>0$. Then $L(\alpha,A_\varepsilon)=(2/\pi)e^{-2\pi q_0\varepsilon}$ if $\alpha=p/q$ for some q dividing q_0, and $L(\alpha,A_\varepsilon)=0$ otherwise. This gives an example of both the discontinuity of the Lyapunov exponent and the lack of quantization of acceleration at rationals.

If we had chosen λ as a more typical function of zero average, we would get discontinuity of the Lyapunov exponent and lack of quantization at all rationals, both becoming increasingly less pronounced as the denominators grow.

2.3. Characterization of uniform hyperbolicity: proof of Theorem 6

Since the Lyapunov exponent is a C^∞ function in $\mathcal{UH}$, the "if" part of the proof is obvious from quantization. In order to prove the "only if" direction, we will first show the uniform hyperbolicity of periodic approximants and then show that uniform hyperbolicity persists in the limit. To do this last part, we will use an explicit formula for the derivative of the Lyapunov exponent (at a fixed frequency) in $\mathcal{UH}$.

2.3.1. Uniform hyperbolicity of approximants

LEMMA 8. *Let* $(\alpha, A) \in (\mathbb{R}\setminus\mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ *and assume that* (α, A) *is regular with positive Lyapunov exponent. If* p/q *is close to* α *and* $\tilde{A}$ *is close to* A *then* $(p/q, \tilde{A})$ *is uniformly hyperbolic.*

Proof. Let us show that if $p_n/q_n \to \alpha$ and $A^{(n)} \to A$ then there exists $\varepsilon''>0$ such that

$$\frac{1}{q_n} \log \varrho(A^{(n)}_{(p_n/q_n)}(x)) = L(\alpha, A_{\operatorname{Im} x}) + o(1), \quad |\operatorname{Im} x| < \varepsilon'', \tag{14}$$

which implies the result. In fact this estimate is just a slight adaptation of what we did in §2.2.

Since $A^{(n)} \to A$ and (α, A) is regular, we may choose $\varepsilon>0$ such that (α, A_δ) is regular for $|\delta|<\varepsilon$, and a sequence $A_n \in C^\omega_\varepsilon(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$ such that $A_n \to A$ uniformly in $|\operatorname{Im} z|<\varepsilon$ as $n\to\infty$.

Choose $\varepsilon''<\varepsilon'<\varepsilon$. We have seen in §2.2 that there exists k_0 such that

$$\operatorname{tr} A^{(n)}_{(p_n/q_n)}(x) = \sum_{|k|\leqslant k_0} a_{k,n} e^{2\pi i k q_n x} + O(e^{-q_n}), \qquad |\operatorname{Im} x| < \varepsilon', \tag{15}$$

$$L\left(\frac{p_n}{q_n}, A^{(n)}_\delta\right) = \max_{k\leqslant |k_0|} \max\left\{\frac{\log|a_{k,n}|}{q_n} - 2\pi k\delta, 0\right\} + o(1), \qquad |\delta| < \varepsilon'. \tag{16}$$

By Theorem 9, $L(p_n/q_n, A^{(n)}_\delta) \to L(\alpha, A_\delta)$ uniformly on compact subsets of $|\delta|<\varepsilon$, so we may rewrite (16) as

$$L(\alpha, A^{(n)}_\delta) = \max_{k\leqslant |k_0|} \max\left\{\frac{\log|a_{k,n}|}{q_n} - 2\pi k\delta, 0\right\} + o(1), \quad |\delta| < \varepsilon'. \tag{17}$$

Since the left-hand side in (17) is an affine positive function over $|\delta|<\varepsilon$, with slope $2\pi\omega(\alpha, A)$, it follows that $|\omega(\alpha, A)| \leqslant k_0$ and

$$L(\alpha, A_\delta) = \frac{\log|a_{-\omega(\alpha,A),n}|}{q_n} + 2\pi\omega(\alpha, A)\delta + o(1), \quad |\delta| < \varepsilon''. \tag{18}$$

Morever, if $|j|\leqslant k_0$ is such that $j\neq -\omega(\alpha, A)$, we have

$$\frac{\log|a_{j,n}|}{q_n} - 2\pi j\delta + 2\pi(\varepsilon' - \varepsilon'') \leqslant L(\alpha, A_\delta) + o(1), \quad |\delta| < \varepsilon''. \tag{19}$$

Together, (15), (18) and (19) imply (14), as desired. □

2.3.2. Derivative of the Lyapunov exponent at uniformly hyperbolic cocycles

Fix $(\alpha, A)\in\mathcal{UH}$ and let $u, s: \mathbb{R}/\mathbb{Z}\to\mathbb{PC}^2$ be the unstable and stable directions. Also let $B: \mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{C})$ be analytic with column vectors in the directions of $u(x)$ and $s(x)$. Then

$$B(x+\alpha)^{-1}A(x)B(x) = \begin{pmatrix} \lambda(x) & 0 \\ 0 & \lambda(x)^{-1} \end{pmatrix} = D(x). \tag{20}$$

Obviously $L(\alpha, A)=L(\alpha, D)=\int_{\mathbb{R}/\mathbb{Z}} \operatorname{Re}\log\lambda(x)\,dx$.([15])

Write

$$B(x) = \begin{pmatrix} a(x) & b(x) \\ c(x) & d(x) \end{pmatrix}.$$

We note that, although the definition of B involves arbitrary choices, it is clear that

$$q_1(x) = a(x)d(x)+b(x)c(x), \quad q_2(x) = c(x)d(x) \quad \text{and} \quad q_3(x) = -b(x)a(x)$$

depend only on (α, A). We will call q_i, $i=1,2,3$, the *coefficients of the derivative of the Lyapunov exponent*, for reasons that will be clear in a moment.

LEMMA 9. *Let $(\alpha, A)\in\mathcal{UH}$ and let $q_1, q_2, q_3: \mathbb{R}/\mathbb{Z}\to\mathbb{C}$ be the coefficients of the derivative of the Lyapunov exponent. Let $w: \mathbb{R}/\mathbb{Z}\to\mathfrak{sl}(2,\mathbb{C})$ be analytic, and write*

$$w = \begin{pmatrix} w_1 & w_2 \\ w_3 & -w_1 \end{pmatrix}.$$

Then

$$\frac{d}{dt}L(\alpha, Ae^{tw}) = \operatorname{Re}\int_{\mathbb{R}/\mathbb{Z}} \sum_{i=1}^{3} q_i(x)w_i(x)\,dx, \quad \text{at } t=0. \tag{21}$$

Proof. Write $B(x+p/q)^{-1}A(x)e^{tw(x)}B(x)=D^t(x)$. We notice that

$$D(x)^{-1}\frac{d}{dt}D^t(x) = B(x)^{-1}w(x)B(x), \quad \text{at } t=0, \tag{22}$$

and

$$\sum_{i=1}^{3} q_i(x)w_i(x) = \text{u.l.c. of } B(x)^{-1}w(x)B(x), \tag{23}$$

where u.l.c. stands for the upper left coefficient.

Suppose first that α is a rational number p/q. Then

$$\frac{d}{dt}L\left(\frac{p}{q}, Ae^{tw}\right) = \frac{1}{q}\frac{d}{dt}\int_{\mathbb{R}/\mathbb{Z}} \log\varrho(D^t_{(p/q)}(x))\,dx, \tag{24}$$

([15]) Notice that the quantization of the acceleration in the uniformly hyperbolic case follows immediately from this expression (the integer arising being the number of turns $\lambda(x)$ does around 0).

so it is enough to show that

$$\frac{d}{dt}\log \varrho(D^t_{(p/q)}(x)) = \operatorname{Re}\sum_{j=0}^{q-1}\sum_{i=1}^{3} q_i\left(x+j\frac{p}{q}\right)w_i\left(x+j\frac{p}{q}\right), \quad \text{at } t=0. \tag{25}$$

Since $D_{(p/q)}(x)$ is diagonal and its u.l.c. has norm bigger than 1,

$$\frac{d}{dt}\log \varrho(D^t_{(p/q)}(x)) = \operatorname{Re} \text{ u.l.c. of } D_{(p/q)}(x)^{-1}\frac{d}{dt}D^t_{(p/q)}(x), \quad \text{at } t=0. \tag{26}$$

Writing $D_{[j]}(x)=D(x+(j-1)p/q)\ldots D(x)$, and using (22), we see that

$$\begin{aligned} &D_{(p/q)}(x)^{-1}\frac{d}{dt}D^t_{(p/q)}(x) \\ &\quad=\sum_{j=0}^{q-1} D_{[j]}(x)^{-1}B\left(x+j\frac{p}{q}\right)^{-1}w\left(x+j\frac{p}{q}\right)B\left(x+j\frac{p}{q}\right)D_{[j]}(x), \quad \text{at } t=0. \end{aligned} \tag{27}$$

Since the $D_{[j]}$ are diagonal,

$$\begin{aligned} &\text{u.l.c. of } D_{[j]}(x)^{-1}B\left(x+j\frac{p}{q}\right)^{-1}w\left(x+j\frac{p}{q}\right)B\left(x+j\frac{p}{q}\right)D_{[j]}(x) \\ &\qquad\qquad = \text{u.l.c. of } B\left(x+j\frac{p}{q}\right)^{-1}w\left(x+j\frac{p}{q}\right)B\left(x+j\frac{p}{q}\right). \end{aligned} \tag{28}$$

Putting together (23), (26), (27) and (28), we get (25).

Since the Lyapunov exponent is C^∞ in $\mathcal{UH}$, the validity of the formula in the rational case yields the irrational case by approximation. □

2.3.3. Proof of Theorem 6

Let $(\alpha, A)\in(\mathbb{R}\setminus\mathbb{Q})\times C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$ be such that (α, A) is regular. Then there exists $\varepsilon>0$ such that $L(\alpha, A_\delta)$ is regular for $|\delta|<\varepsilon$.

Fix $0<\varepsilon'<\varepsilon$ and choose a sequence of rationals $p_n/q_n\to\alpha$. By Lemma 8, if n is large then $(p_n/q_n, A_\delta)$ is uniformly hyperbolic for $\delta<\varepsilon'$. So one can define functions $u_n(x)$ and $s_n(x)$, with values in $\mathbb{PC}^2$, corresponding to the eigendirections of $A_{(p_n/q_n)}(x)$ with the largest and smallest eigenvalues. Our strategy will be to show that the sequences $u_n(x)$ and $s_n(x)$ converge uniformly (in a band) to functions $u(x)$ and $s(x)$.

The coefficients of the derivative of $L(p_n/q_n, A)$ will be denoted by q_i^n, $i=1,2,3$. The basic idea now is that if $q_2^n(x)$ and $q_3^n(x)$ are bounded, then it follows directly from the definitions that the angle between $u_n(x)$ and $s_n(x)$ is not too small, and this is enough to guarantee convergence. On the other hand, the derivative of the Lyapunov exponent is under control by pluriharmonicity, which yields the desired bound on the coefficients.

There are various ways to proceed here, and we will just estimate the Fourier coefficients of the q_j^n, $j=2,3$. Write

$$\zeta_{n,k,j} = \int_{\mathbb{R}/\mathbb{Z}} q_j^n(x) e^{-2\pi ikx}\, dx. \tag{29}$$

LEMMA 10. *There exist $C>0$ and $\gamma>0$ such that, for every n sufficiently large,*

$$|\zeta_{n,k,j}| \leqslant Ce^{-\gamma|k|}, \quad j=2,3, \quad k\in\mathbb{Z}. \tag{30}$$

Proof. Choose $0<\gamma<2\pi\varepsilon'$. Then, for each fixed large n, we have $|\zeta_{n,k,j}|\leqslant C_n e^{-\gamma|k|}$ (since q_j^n extend to $|\operatorname{Im} z|<\varepsilon'$). If the result was false, then there would exist sequences $n_l\in\mathbb{N}$, $k_l\in\mathbb{Z}$ and $j_l\in\{2,3\}$ such that $n_l\to\infty$ and, for each l, $|\zeta_{n_l,k_l,j_l}|>le^{-\gamma|k_l|}$. We may assume that j_l is a constant and that either $k_l>0$ for all l or $k_l\leqslant 0$ for all l.

For simplicity, we will assume that $j_l=2$ and $k_l\leqslant 0$ for all l. Let

$$w_{(l)}(x) = \frac{|\zeta_{n_l,k_l,2}|}{\zeta_{n_l,k_l,2}} e^{\gamma|k_l|} \begin{pmatrix} 0 & e^{-2\pi ik_l x} \\ 0 & 0 \end{pmatrix} \tag{31}$$

and choose $\gamma<\gamma'<2\pi\varepsilon'$. Setting $\tilde{A}(x)=A(x+i\gamma'/2\pi)$ and $\widetilde{w}_{(l)}(x)=w_{(l)}(x+i\gamma'/2\pi)$, we get

$$\frac{d}{dt} L\left(\frac{p_{n_l}}{q_{n_l}}, \tilde{A}e^{t\widetilde{w}_{(l)}}\right) = e^{\gamma|k_l|}|\zeta_{n_l,k_l,2}| \geqslant l, \tag{32}$$

since the coefficients of the derivative at $(p_{n_l}/q_{n_l}, \tilde{A})$ are $\tilde{q}_j^{n_l}(x)=q_j^{n_l}(x+i\gamma'/2\pi)$.

Note that $\widetilde{w}_{(l)}$ admits a holomorphic extension bounded by 1 on $|\operatorname{Im} z|<(\gamma'-\gamma)/2\pi$. Since $(\alpha,\tilde{A})$ is regular with positive Lyapunov exponent, it follows from Lemma 8 that there exists $r>0$ such that, for every large l, $(p_{n_l}/q_{n_l}, \tilde{A}e^{t\widetilde{w}_{(l)}})$ is uniformly hyperbolic for complex t with $|t|<r$. In particular, the functions $t\mapsto L(p_{n_l}/q_{n_l}, \tilde{A}e^{t\widetilde{w}_{(l)}})$ are harmonic on $|t|<r$ for large l. These functions are also clearly uniformly bounded. Harmonicity gives then that the derivative at $t=0$ is uniformly bounded as well. This contradicts (32). □

LEMMA 11. *For every $C>0$ there exists $\varepsilon>0$ such that, if $a,b,c,d\in\mathbb{C}$ are such that $ad-bc=1$ and the angle between the complex lines through $\binom{a}{c}$ and $\binom{b}{d}$ is less than ε, then $\max\{|ab|,|cd|\}>C$.*

Proof. This is a straightforward computation. □

Lemma 10 implies that there exists $\gamma>0$ such that q_2^n and q_3^n are uniformly bounded, as $n\to\infty$, on $|\operatorname{Im} x|<\gamma$. By Lemma 11, this implies that there exists $\eta>0$ such that the angle between $u_n(x)$ and $s_n(x)$ is at least η, whenever n is large and $|\operatorname{Im} x|<\gamma$. We are in position to apply a normality argument.

LEMMA 12. *Let $u_n(x)$ and $s_n(x)$ be holomorphic functions defined in some complex manifold, with values in $\mathbb{PC}^2$. If the angle between $u_n(x)$ and $s_n(x)$ is bounded away from 0 uniformly in x and n, then $u_n(x)$ and $s_n(x)$ form normal families. Moreover, the limits of u_n and s_n (taken along the same subsequence) are holomorphic functions such that $u(x)\neq s(x)$ for every x.*

Proof. We may identify $\mathbb{PC}^2$ with the Riemann sphere. Write $\phi_n(x)=u_n(x)/s_n(x)$. Then $\phi_n(x)$ avoids a neighborhood of 1, and hence it forms a normal family. Let us now take a sequence along which ϕ_n converges, and let us show that $u_n(x)$ and $s_n(x)$ form normal families. This is a local problem, so we may work in a neighborhood of a point z. If $\lim_{n\to\infty}\phi_n(z)\neq\infty$ then, for every large n, ϕ_n must be uniformly bounded in a neighborhood of z, so u_n and $1/s_n$ must also be bounded. If $\lim_{n\to\infty}\phi_n(z)=\infty$ then, for every large n, $1/\phi_n$ must be uniformly bounded in a neighborhood of z, so s_n and $1/u_n$ must be bounded. In either case we conclude that s_n and u_n are normal in a neighborhood of z.

The last statement is obvious by pointwise convergence. □

Let $u(x)$ and $s(x)$ be limits of $u_n(x)$ and $s_n(x)$ over $|\operatorname{Im} x|<\gamma$, taken along the same subsequence. Then $A(x)\cdot u(x)=u(x+\alpha)$, $A(x)\cdot s(x)=s(x+\alpha)$ and $u(x)\neq s(x)$. Since $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and $L(\alpha,A)>0$, this implies that $(\alpha,A)\in\mathcal{UH}$, which completes the proof of Theorem 6.

2.4. Local non-triviality of the Lyapunov function in strata: Proof of Theorem 8

Let δ, j and v_* be as in the statement of Theorem 8. Notice that $(\alpha,A^{(v_*)})\notin\mathcal{UH}$, since otherwise we would have $j=\omega(\alpha,A^{(v_*)})=0$.

Let $0<\varepsilon_0<\delta$ be such that $(\alpha,A_\varepsilon^{(v_*)})\in\mathcal{UH}$ and $\omega(\alpha,A_\varepsilon^{(v_*)})=j$ for $0<\varepsilon<\varepsilon_0$. By definition, for every $0<\varepsilon<\varepsilon_0$ we have $L_{\delta,j}(\alpha,A^{(v)})=L(\alpha,A_\varepsilon^{(v)})-2\pi j\varepsilon$ for $v\in C^\omega_{\varepsilon_0}(\mathbb{R}/\mathbb{Z},\mathbb{R})$ in a neighborhood of v_*.

Let $u,s\colon\{x:0<\operatorname{Im} x<\varepsilon_0\}\to\mathbb{PC}^2$ be such that $x\mapsto u(x+i\varepsilon)$ and $x\mapsto s(x+i\varepsilon)$ are the unstable and stable directions of $(\alpha,A_\varepsilon^{(v_*)})$. Also, let $q_1,q_2,q_3\colon\{x:0<\operatorname{Im} x<\varepsilon_0\}\to\mathbb{C}$ be such that $x\mapsto q_j(x+i\varepsilon)$ is the jth coefficient of the derivative of $(\alpha,A_\varepsilon^{(v_*)})$.

Notice that $A^{(v_*+w)}=A^{(v_*)}e^{\widetilde{w}}$, where

$$\widetilde{w}(x)=\begin{pmatrix}0 & 0\\ -w(x) & 0\end{pmatrix}.$$

Thus the derivative of $w\mapsto L_{\delta,j}(\alpha,A^{(v_*+tw)})$ at $t=0$ is

$$-\operatorname{Re}\int_{\mathbb{R}/\mathbb{Z}} w(x+i\varepsilon)q_3(x+i\varepsilon)\,dx. \tag{33}$$

If the result is false, then (33) must vanish for every $w\in C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$. Testing this with w of the form $a\cos 2\pi kx+b\sin 2\pi kx$, with $a,b\in\mathbb{R}$ and $k\in\mathbb{Z}$, we see that the kth Fourier coefficient of q_3 must be minus the complex conjugate of the $(-k)$-th Fourier coefficient of q_3 for every $k\in\mathbb{Z}$. Since the Fourier series converges when $0<\operatorname{Im} x<\varepsilon_0$, this implies that it actually converges when $|\operatorname{Im} x|<\varepsilon_0$, and over $\mathbb{R}/\mathbb{Z}$ it defines a purely imaginary function. Thus $q_3(x)$ extends analytically through $|\operatorname{Im} x|=0$, and hence

$$q_2(x)=c(x)d(x)=a(x-\alpha)b(x-\alpha)=-q_3(x-\alpha)$$

(the middle equality holding due to the Schrödinger form) also does.

Identifying $\mathbb{PC}^2$ with the Riemann sphere in the usual way (the line through $\binom{z}{w}$ corresponding to z/w), we get $q_2=1/(u-s)$ and $-q_3=us/(u-s)$. These formulas allow us to analytically continue u and s through $\operatorname{Im} x=0$.(16) Since q_2 and q_3 are purely imaginary when $\operatorname{Im} x=0$, we conclude that u and s are complex conjugate directions in $\mathbb{PC}^2$ when $\operatorname{Im} x=0$.(17) Note that if $u(x)=s(x)$ for some x then $u(x+n\alpha)=s(x+n\alpha)$ for every n and thus, by analytic continuation, $u=s$ everywhere, which is impossible. So when $\operatorname{Im} x=0$, u and s are in fact distinct complex conjugate directions, and in particular they are not real.

Let $B(x)\in\mathrm{SL}(2,\mathbb{R})$ be the unique upper triangular matrix with positive diagonal coefficients taking $u(x)$ to $\pm i$ (and hence $s(x)$ to $\mp i$). Then $B:\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{R})$ is analytic. Define $A(x)=B(x+\alpha)A^{(v_*)}(x)B(x)^{-1}$. Since $A^{(v_*)}(x)$ takes $u(x)$ and $s(x)$ to $u(x+\alpha)$ and $s(x+\alpha)$, we conclude that $B(x+\alpha)A^{(v_*)}(x)B(x)^{-1}\in\mathrm{SO}(2,\mathbb{R})$. As $x\mapsto A^{(v_*)}(x)$ is homotopic to a constant as a function $\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{R})$, $x\mapsto B(x+\alpha)A^{(v_*)}(x)B(x)^{-1}\in\mathrm{SL}(2,\mathbb{R})$ is homotopic to a constant as a function $\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{R})$, and thus also as a function $\mathbb{R}/\mathbb{Z}\to\mathrm{SO}(2,\mathbb{R})$. It follows that there exists an analytic function $\phi:\mathbb{R}/\mathbb{Z}\to\mathbb{R}$ such that $B(x+\alpha)A^{(v_*)}(x)B(x)^{-1}=A(x)$, where $A(x)$ is the rotation by the angle $2\pi\phi(x)$.

---

(16) Indeed, if q_2 is identically vanishing then we set either $u=\infty$ and $s=-q_3$, or $s=\infty$ and $u=q_3$ (to match the previous definition when $\operatorname{Im} z>0$). If q_2 is not identically vanishing but $1-4q_2q_3$ vanishes identically, then we can take $u=-s=1/2q_2$. Assume now that there exists $x\in\mathbb{R}/\mathbb{Z}$ such that $q_2(x)\neq0$ and $1-4q_2(x)q_3(x)\neq0$. Then we can define $u(z)$ and $-s(z)$ in a small open square Q of side $2r$ centered at x (with sides parallel to the coordinate axis) as the distinct solutions of the equation

$$w^2-\frac{1}{q_2}w+\frac{q_3}{q_2}=0$$

(to match the previous definition when $\operatorname{Im} z>0$). We then spread it using the dynamics to $|\operatorname{Im} z|<r$ in order to have $A_k(z)\cdot u(z)=u(z+k\alpha)$ and $A_k(z)\cdot s(z)=s(z+k\alpha)$ for every $z\in Q$ and $k\in\mathbb{Z}$. This gives well-defined analytic functions because whenever z and $z+n\alpha\in Q$ we have $A_n(z)\cdot u(z)=u(z+n\alpha)$ and $A_n(z)\cdot s(z)=s(z+n\alpha)$ (this is clear when $\operatorname{Im} z>0$ and holds in general by analytic continuation).

(17) In principle, $\operatorname{Re} q_2=\operatorname{Re} q_3=0$ is also compatible with values of u and s such that $\operatorname{Re} u=\operatorname{Re} s=0$ (where we set $\operatorname{Re}\infty=0$ for simplicity). But if $\operatorname{Re} u=0$ then $\operatorname{Re}(A\cdot u)=\operatorname{Re} v_*$, so if $u\neq\bar{s}$ somewhere, then v_* must be identically vanishing. This contradicts $\omega(\alpha,A^{(v_*)})\neq0$.

Obviously this relation implies that $L(\alpha, A_\varepsilon^{(v_*)})=L(\alpha, A_\varepsilon)$ for $\varepsilon>0$ small. If we show that $L(\alpha, A_\varepsilon)=0$ for $\varepsilon>0$ small, we can conclude that $\omega(\alpha, A^{(v_*)})=0$, which contradicts the hypothesis.

To see that $L(\alpha, A_\varepsilon)=0$, we notice that, for $n\geqslant 1$, $A_n(x)$ is the rotation by the angle

$$\sum_{k=0}^{n-1} \phi(x+k\alpha).$$

Thus

$$\frac{1}{n}\int_{\mathbb{R}/\mathbb{Z}} \log\|A_n(x+i\varepsilon)\|\,dx = 2\pi\int_{\mathbb{R}/\mathbb{Z}} \frac{1}{n}\left|\sum_{k=0}^{n-1} \operatorname{Im}\phi(x+k\alpha+i\varepsilon)\right| dx. \tag{34}$$

Since $x\mapsto x+\alpha$ is uniquely ergodic with respect to Lebesgue measure, the integrand of the right-hand side converges uniformly, as $n\to\infty$, to

$$\left|\int_{\mathbb{R}/\mathbb{Z}} \operatorname{Im}\phi(x+i\varepsilon)\,dx\right| = \left|\int_{\mathbb{R}/\mathbb{Z}} \operatorname{Im}\phi(x)\,dx\right| = 0.$$

Thus the limit of the right-hand side, which is $L(\alpha, A_\varepsilon)$ by definition, is zero as well.

Remark 13. The analysis of the function $A\mapsto L_{\delta,j}(\alpha, A)$ from the real Banach manifold $C^\omega_\delta(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ (of maps in $C^\omega_\delta(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{C}))$ taking values in $\mathrm{SL}(2,\mathbb{R})$) to $\mathbb{R}$, with $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and near some $\tilde{A}$ with $\omega(\alpha, \tilde{A})>0$, can be carried out as above with one important difference.

The argument above allows one to establish that, if $L_{\delta,j}$ is not a local submersion, then the coefficients of the derivative q_2 and q_3 extend from some half band $0<\operatorname{Im} x<\varepsilon_0$ to a full band $|\operatorname{Im} x|<\varepsilon_0$.[18] This again leads to the conclusion that there exists $B\colon\mathbb{R}/\mathbb{Z}\to \mathrm{SL}(2,\mathbb{R})$ analytic such that $A(x)=B(x+\alpha)\tilde{A}(x)B(x)^{-1}$ takes values in $\mathrm{SO}(2,\mathbb{R})$. But now there are two cases:

(1) $x\mapsto\tilde{A}(x)$ is homotopic to a constant. In this case, the above argument goes through and one concludes that $\omega(\alpha, \tilde{A})=0$, yielding a contradiction.

(2) $x\mapsto\tilde{A}(x)$ is not homotopic to a constant. In this case, there is no contradiction, and the reader is invited to check that, if $\tilde{A}(x)$ is the rotation by the angle $2\pi x$, then indeed the derivative of $L_{\delta,j}$ vanishes, but $\omega(\alpha, \tilde{A})=1$.

The analysis of the second case has been carried out by different means in [AK2], where it is shown that the Lyapunov exponent is real-analytic near cocycles with values in $\mathrm{SO}(2,\mathbb{R})$ provided they are not homotopic to a constant. We should emphasize that this result is obtained for any number of frequencies, which is certainly beyond the scope of the techniques we develop in this paper.

---

[18] Though one lacks the symmetry between q_2 and q_3 exploited above, we just separately evaluate the extensions of q_2 and q_3, as we are not constrained to consider just perturbations of a specific form.

Interpreting their results (in the one-frequency case) from our new point of view, [AK2] shows that all real perturbations of $(\alpha, \tilde{A})$ have the same acceleration, namely the absolute value of the topological degree of $\tilde{A}$ as a map $\mathbb{R}/\mathbb{Z} \to \mathrm{SL}(2,\mathbb{R})$. Real analyticity implies that the derivative of the Lyapunov exponent then is forced to vanish whenever the Lyapunov exponent is zero. In [AK2] it is shown that the second derivative is non-zero. The locus of zero exponents can be shown to intersect a neighborhood of $\tilde{A}$ in $C^\omega_\delta(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ in an analytic submanifold of codimension $4|\omega(\alpha, \tilde{A})|$.

Thus our result implies that among cocycles which are not homotopic to constants (and with a given irrational frequency), the locus of zero exponents is contained in a countable union of positive codimension submanifolds of $C^\omega_\delta(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$.

3. Part II: Acriticality for typical operators

The main goal of this section is to prove the main theorem. As we have done so far, our analysis of the operator $H_{\alpha,v}$ will be based on the dynamics of the associated family of Schrödinger cocycles.

Given $H_{\alpha,v}$, we define the acceleration ω at energy $E \in \mathbb{R}$ by $\omega(E) = \omega(\alpha, A)$, where $A = A^{(E-v)}$. Then E is critical if and only if $L(E) = 0$ and $\omega(E) > 0$ (see Remark 4). If E is critical with acceleration k, we call it a *critical point of degree* k.

Our basic plan is to show that critical points of maximal degree $k \geqslant 1$ can be destroyed by a small typical perturbation by trigonometric polynomials of some large degree. This may give rise to many critical points of degree $\leqslant k-1$, but by iterating this process we will eventually get rid of all of them.

More formally, let $\mathcal{A}_k$, $k \geqslant 0$, be the set of all $(\alpha, v) \in (\mathbb{R} \setminus \mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ such that $H_{\alpha,v}$ has only critical points of degree at most k. Hence $\mathcal{A}_k$ forms an increasing sequence of open sets with $\bigcup_{k \geqslant 0} \mathcal{A}_k = (\mathbb{R} \setminus \mathbb{Q}) \times C^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ and $\mathcal{A}_0$ is the set of all (α, v) such that $H_{\alpha,v}$ is acritical.[19] Let $\mathcal{P}^n \subset C^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ be the space of trigonometric polynomials with degree at most n, and let $\mathcal{P}^n_0 \subset \mathcal{P}^n$ be the subspace of zero-average functions. Also, for $\varepsilon > 0$, let $\mathcal{P}^n(\varepsilon) \subset \mathcal{P}^n$ and $\mathcal{P}^n_0(\varepsilon) \subset \mathcal{P}^n_0$ be the corresponding ε-balls with respect to the C^0 norm.

Our main estimate is the following.

---

[19] Note that $\mathcal{A}_k \setminus \mathcal{A}_{k-1}$ is actually non-empty for every $k \geqslant 1$, as it contains (α, v) with $v(x) = 2\cos 2\pi k x$ for every $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ (see (2) in §2.1.4).

PROPOSITION 14. *For every* $(\alpha,v)\in\mathcal{A}_k$ *there exist* $n\geqslant 1$ *and* $\varepsilon>0$ *such that for almost every* $w\in\mathcal{P}_0^n(\varepsilon)$ *we have* $(\alpha,v+w)\in\mathcal{A}_{\max\{0,k-1\}}$.[20]

The main theorem follows immediately from this estimate.

The proof of Proposition 14 will take the remainder of this section. In §3.1 we apply the renormalization and the generalized Kotani theory developed in [AK1] and [AK2] to study the critical set within the subvarieties obtained in Theorem 1. This shows that Proposition 14 can be concluded if we can locally subfoliate the subvarieties by premonotonic curves. The existence of such a local subfoliation can be deduced by locating, for any critical cocycle, an appropriate signed vector tangent to the subvariety. This is obtained in Theorem 10. The proof, by contradiction, is given in §3.3. We concentrate on studying the dynamics when the derivative of the Lyapunov exponent behaves as a measure. (If this is not the case, a special kind of signed vector, a so-called directed vector, can be constructed right away.) We show in Theorem 16, proved in §3.4, that this is very constraining of the dynamics, and can be broken by a small conjugacy of the cocycle (the directed vectors are not conjugacy invariant), which implies the result since signed vectors are conjugacy invariant. Actually the conjugacy initially produces a cocycle which is not in Schrödinger form, so we must produce an additional conjugacy to put it into Schrödinger form. This is achieved by Theorem 17, proved in §3.5.

We note that another application of Theorem 10 is that the critical regime is the boundary of the supercritical regime. This is proved in §3.2 (which may be skipped in a first reading, since its content is not needed for the proof of the main theorem).

3.1. Parameter exclusion argument

From now on, $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ is fixed.

In §2 we proved that critical cocycles have codimension one among all cocycles. Earlier, in [AK1] and [AK2], we had shown that almost every cocycle in certain one-parameter families has either a positive Lyapunov exponent or admits a sequence of renormalizations converging to a good normal form. The techniques in those works are quite distinct, and our aim is to combine them to show that critical cocycles in fact have zero measure inside a codimension-one subspace. The key difficulty we will face is in establishing an indefiniteness result for the derivative of the Lyapunov exponent, which will enable us to construct appropriate one-parameter families inside the locus where criticality might appear.

[20] A slightly stronger statement follows from our proof: if $(\alpha,v)\in(\mathbb{R}\setminus\mathbb{Q})\times C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$ then $n=n(\alpha,v)$ and $\varepsilon=\varepsilon(\alpha,v,\delta)$ may be chosen so that for every $(\alpha',v')\in(\mathbb{R}\setminus\mathbb{Q})\times C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$ such that $|\alpha-\alpha'|<\varepsilon$ and $\|v-v'\|_\delta<\varepsilon$ and for almost every $w\in\mathcal{P}_0^n(\varepsilon)$ we have $(\alpha',v'+w)\in\mathcal{A}_{\max\{0,k-1\}}$.

In this strategy, Proposition 14 is obtained as a consequence of the following. For all $k \geqslant 1$ let $\mathcal{C}^k$ be the set of all $v \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ such that $L(\alpha, A^{(v)})=0$ and $\omega(\alpha, A^{(v)})=k$. Note that $\bigcup_{j \geqslant k} \mathcal{C}_j$ is closed by upper-semicontinuity of the acceleration and continuity of the Lyapunov exponent.

Proposition 15. *For every $v_0 \in \mathcal{C}^k$, there exist $n \geqslant 1$ and $\varepsilon>0$ such that*

$$\{w \in \mathcal{P}^n(\varepsilon): v_0+w \in \mathcal{C}^k\}$$

has $2n$-dimensional Hausdorff measure zero. Moreover, n and ε can be chosen uniformly over compact subsets of $\mathcal{C}^k$.

Proposition 15 implies Theorem 14 by projecting in the direction of the energy, using that the set of critical points of maximal degree for $H_{\alpha,v}$ is compact. Indeed, suppose that $(\alpha, v) \in \mathcal{A}_k$ with $k \geqslant 1$ (if $k=0$ the result follows from the fact that $\mathcal{A}_0$ is open), and let n and ε be obtained by applying Proposition 15 to the set K of all $v_0=E-v$ with E being a critical point of degree k for v. Shrinking ε if necessary, we may assume it satisfies the extra property that $(\alpha, v+w) \in \mathcal{A}_k$ for every $w \in \mathcal{P}^n_0(\varepsilon)$ (since $\mathcal{A}_k$ is open), and moreover that, if $E \in \mathbb{R}$ is such that $E-v-w \in \mathcal{C}^k$, then there exists E' such that $E'-v \in \mathcal{C}^k$ and $|E-E'|<\frac{1}{4}\varepsilon$ (here we use that $\bigcup_{j \geqslant k} \mathcal{C}^j$ is closed). Choose a finite $\frac{1}{4}\varepsilon$-dense subset $\mathcal{E}$ of the set of critical points of degree k for $H_{\alpha,v}$. The conclusion of Proposition 15 gives that, for almost every $w \in \mathcal{P}^n_0\left(\frac{1}{2}\varepsilon\right)$ and every $E_0 \in \mathcal{E}$, we have $E-v-w \notin \mathcal{C}^k$ provided $|E-E_0|<\frac{1}{2}\varepsilon$ (here we use that the projection from $\mathcal{P}^n$ to $\mathcal{P}^n_0$ takes a set of $2n$-dimensional Hausdorff measure zero to a set of zero Lebesgue measure). Thus $(\alpha, v+w) \in \mathcal{A}_{k-1}$ by the extra property above.

Through the remainder of the paper, $v_0 \in \mathcal{C}^k$ is also fixed.

Fix $\xi'>\xi>0$ such that $v_0 \in C^\omega_{\xi'}(\mathbb{R}/\mathbb{Z}, \mathbb{R})$. The hypothesis $v_0 \in \mathcal{C}^k$ implies that $v_0 \in \Omega_{\xi,k}$. Then there exists a neighborhood $\mathcal{V} \subset C^\omega_\xi(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ of v_0 and $0<\xi_0<\xi$ such that $(\alpha, A^{(v)}_{\xi_0}) \in \mathcal{UH}$ and $\omega(\alpha, A^{(v)}_{\xi_0})=k$ for every $v \in \mathcal{V}$. We fix such ξ_0 and note that $L_{\xi,k}(\alpha, A^{(v)})$, $v \in \mathcal{V}$, is given by $L(\alpha, A^{(v)}_{\xi_0})-2\pi k \xi_0$. For simplicity, we will denote this function by $L_{\xi,k}$. As before, for every $v \in \mathcal{V}$ such that $\omega(\alpha, A^{(v)})=k$ we have $L_{\xi,k}(v)=L(\alpha, A^{(v)})$, and thus $\mathcal{C}^k \cap \mathcal{V} \subset L^{-1}_{\xi,k}(0)$.

Let $U \subset \mathbb{R}^n$ be an open neighborhood of 0 and let $v_\lambda \in \mathcal{V}$, $\lambda \in U$, be an analytic deformation of v_0. For any $\lambda_0 \in U$, let $D_{\lambda_0} v_\lambda: \mathbb{R}^n \to C^\omega_\xi(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ be the derivative

$$D_{\lambda_0} v_\lambda \cdot w = \frac{d}{dt} v_{\lambda_0+tw}\bigg|_{t=0}.$$

The reader should keep in mind the family $v_\lambda=v_0+P_m\lambda$, $\lambda \in \mathbb{R}^{2m+1}$, where

$$P_m: \mathbb{R}^{2m+1} \longrightarrow \mathcal{P}^m$$

is some fixed isomorphism (in this case $\operatorname{Im} D_0 v_\lambda = \mathcal{P}^m$).

We say that $a \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathfrak{sl}(2,\mathbb{R}))$ is *signed* if $\det a(x) > 0$ for every $x \in \mathbb{R}/\mathbb{Z}$. Given $A \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$, we say that $a \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ is *A-signed* if there exists $b \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathfrak{sl}(2,\mathbb{R}))$ such that

$$x \longmapsto A(x)^{-1} b(x+\alpha) A(x) - b(x) + a(x) \tag{35}$$

is signed.

Given $v_0, w \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$, we say that w is v_0*-signed* if

$$\begin{pmatrix} 0 & 0 \\ -w & 0 \end{pmatrix}$$

is $A^{(v_0)}$-signed. Note that the set of (a, A) such that a is A-signed is clearly open, so the set of (v, w) such that w is v-signed is open as well.

Remark 16. It is easy to see that if $\pm w(x) > 0$ for every $x \in \mathbb{R}/\mathbb{Z}$ then w is v-signed (independently of v), just choose

$$b = \begin{pmatrix} 0 & 0 \\ \mp\varepsilon & 0 \end{pmatrix}$$

with sufficiently small $\varepsilon > 0$ in (35).

Remark 17. (Interpretation of signedness) Recall that an analytic one-parameter family $A^\lambda \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ is said to be *monotonic* (in the sense of [AK2]) if for each $x \in \mathbb{R}/\mathbb{Z}$ and each unit vector m in $\mathbb{R}^2$ the derivative, with respect to λ, of the argument of $A^\lambda(x) \cdot m$ is non-zero. It is easy to see that this condition is equivalent to positivity of $\det a^\lambda$, where

$$a^\lambda = (A^\lambda)^{-1} \frac{d}{d\lambda} A^\lambda.$$

Monotonicity is a powerful concept that allows one to efficiently use complexification techniques in the analysis of the parameter space (generalizing Kotani theory).

It turns out that monotonicity is not invariant under coordinate changes. Indeed, let us consider a one-parameter analytic family of coordinate changes $B^\lambda \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ giving rise to the family $\tilde{A}^\lambda(x) = B^\lambda(x+\alpha) A^\lambda(x) B^\lambda(x)^{-1}$. Also define b^λ and $\tilde{a}^\lambda$ analogously to a^λ. Then

$$\tilde{a}^\lambda(x) = B^\lambda(x)(A^\lambda(x)^{-1} b^\lambda(x+\alpha) A^\lambda(x) - b^\lambda(x) + a^\lambda(x)) B^\lambda(x)^{-1}, \tag{36}$$

so that the determinant of $\tilde{a}^\lambda(x)$ is the same as that of

$$A^\lambda(x)^{-1} b^\lambda(x+\alpha) A^\lambda(x) - b^\lambda(x) + a^\lambda(x).$$

Thus a family A^λ can be made monotonic by a coordinate change near some parameter λ_0 if and only if a_0^λ is A_0^λ-signed.

Proposition 18. *Let* v_λ, $\lambda\in U$, *be an analytic family as above such that there exists a* v_0*-signed vector* w *in the image of* $D_0 v_\lambda$ *with* $DL_{\xi,k}(v_0)\cdot w=0$, *but* $DL_{\xi,k}(v_0)$ *does not vanish over* $D_0 v_\lambda$. *Then there exists* $\varepsilon>0$ *such that the set of all* λ *which are* ε*-close to* 0 *and such that* $v_\lambda\in\mathcal{C}^k$ *has* $(n-1)$*-dimensional Hausdorff measure zero.*

Theorem 8 shows that the linear functional $DL_{\xi,k}(v_0)$ has rank 1 (as $v_0\in\mathcal{C}^k$), so Proposition 18 reduces (see below) the proof of Proposition 15, and hence also the proof of the main theorem, to the following indefiniteness estimate for the derivative of the Lyapunov exponent.

Theorem 10. (Indefiniteness of the derivative) *There exists a* v_0*-signed trigonometrical polynomial* w *such that* $DL_{\xi,k}(v_0)\cdot w=0$.

This result is the technical core of this paper, and will be proved in §3.3. In the remainder of this subsection, we will focus on the proof of Proposition 18. But first, let us show in more detail how it combines with Proposition 10 to yield a proof of Proposition 15.

Proof of Proposition 15 (*assuming Propositions* 18 *and* 10). We say that n is *good* for v_0 if there exist $w_0\in\mathcal{P}^n$ such that $DL_{\xi,k}(v_0)\cdot w_0\neq 0$ and a v_0-signed vector $w_s\in\mathcal{P}^n$ such that $DL_{\xi,k}(v_0)\cdot w_s=0$. Using Theorem 8 and Proposition 10, we see that any large n is good for v_0. For such n, Proposition 18 yields directly that for $\varepsilon>0$ small enough, the set of all $w\in\mathcal{P}^n(\varepsilon)$ such that $v_0+w\in\mathcal{C}^k$ has $2n$-dimensional Hausdorff measure zero.

It remains to check the uniformity over any compact set $K\subset\mathcal{C}^k$. Note that if n is good for v_0, then it is also good for any v in a neighborhood of v_0. Indeed $DL_{\xi,k}(v)\cdot w_0$ depends continuously on v, so it remains non-zero, and hence there exists a unique small t such that $DL_{\xi,k}(v)\cdot(w_s+tw_0)=0$. Since signedness is open, we still have that w_s+tw_0 will be v-signed.

Thus we can choose n which is good for all v in a neighborhood $\mathcal{W}$ of K. Take $\varepsilon>0$ small enough so that $v+\mathcal{P}^n(\varepsilon)\subset\mathcal{W}$ for every $v\in\mathcal{K}$. For such v and any $v'\in v+\mathcal{P}^n(\varepsilon)$, if $v'\in\mathcal{C}^k$ then Proposition 18 provides $\varepsilon'>0$ such that $v'+w'\notin\mathcal{C}^k$ for almost every $w'\in\mathcal{P}^n(\varepsilon')$. This implies that $v+w\notin\mathcal{C}^k$ for almost every $w\in\mathcal{P}^n(\varepsilon)$, as desired. □

Proof of Proposition 18. Let us say that (α,A) is L^2-conjugate to rotations if there exists a measurable $B:\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{R})$ such that $B(x+\alpha)A(x)B(x)^{-1}\in\mathrm{SO}(2,\mathbb{R})$ for almost every x and $\int_{\mathbb{R}/\mathbb{Z}}\|B(x)\|^2dx<\infty$. It is clear that if (α,A) is L^2-conjugate to rotations then $L(\alpha,A)=0$.

The following is a convenient restatement of a result of [AK2].

THEOREM 11. *Let $v_\lambda \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ be an analytic family defined for $\lambda \in \mathbb{R}$ near 0 such that*

$$w = \frac{d}{d\lambda} v_\lambda \bigg|_{\lambda=0}$$

is v_0-signed. Then for almost every λ near 0, $(\alpha, A^{(v_\lambda)})$ is L^2-conjugate to rotations.

Proof. Let b be such that (35) is signed, and let $A^\lambda(x) = e^{\lambda b(x+\alpha)} A^{(v_\lambda)}(x) e^{-\lambda b(x)}$. Then $\lambda \mapsto A^\lambda$ is a monotonic family (in the sense of [AK2]), for λ near 0. By the generalized Kotani theory of [AK2] (see Theorem 1.7 therein), for almost every λ near 0, (α, A^λ), and hence $(\alpha, A^{(v_\lambda)})$, is L^2-conjugate to rotations. □

COROLLARY 12. *If 0 is v_0-signed then $A^{(v_0)}$ is L^2-conjugate to rotations.*

Proof. Apply the previous theorem to the constant family $v_\lambda = v_0$. □

Let p_n/q_n be the sequence of continued fraction approximations of α. Let $\beta_n = (-1)^n(q_n\alpha - p_n) > 0$ and $\alpha_n = \beta_n/\beta_{n-1}$. For $A \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$, we say that (α', A') is an nth *renormalization* of (α, A) if $\alpha' = \alpha_n$, $A' \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ and there exist $x_0 \in \mathbb{R}/\mathbb{Z}$ and $N: \mathbb{R} \to \mathrm{SL}(2,\mathbb{R})$ analytic such that

$$N(x+1) A_{(-1)^{n-1} q_{n-1}}(x_0 + \beta_{n-1} x) N(x)^{-1} = \mathrm{id}, \tag{37}$$

$$N(x+\alpha_n) A_{(-1)^n q_n}(x_0 + \beta_{n-1} x) N(x)^{-1} = A'(x). \tag{38}$$

Here $A_{-k}(x) = A_k(x - k\alpha)^{-1}$ for $k \geqslant 1$.([21])

THEOREM 13. ([AK2, Theorem 4.3]) *Let $A \in C^\omega(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$ be homotopic to a constant. If (α, A) is L^2-conjugate to rotations then for every $\varepsilon > 0$ there exists n and $\theta \in \mathbb{R}$ such that (α, A) has an n-th renormalization (α', A') with $\|A' - R_\theta\|_{\varepsilon^{-1}} < \varepsilon$.*

COROLLARY 14. *If (α, A) is homotopic to a constant and L^2-conjugate to rotations then $\omega(\alpha, A) = 0$.*

Proof. Recall that $\beta_{n-1} = 1/(q_n + \alpha_n q_{n-1})$. Let (α', A') be an nth renormalization of (α, A) and let $N: \mathbb{R} \to \mathrm{SL}(2,\mathbb{R})$ be an analytic map satisfying (37) and (38). It follows that

$$A_{k(-1)^n q_n + l(-1)^{n-1} q_{n-1}}(x_0 + \beta_{n-1} x) = N(x + k\alpha' + l)^{-1} A'_k(x) N(x) \tag{39}$$

for $k, l \in \mathbb{Z}$ (naturally we define $A'_k(x) = A'(x + (k-1)\alpha') \ldots A'(x)$ using translations by α' and not by α).

([21]) Heuristically, the nth renormalization is obtained by inducing to $[x, x+\beta_{n-1}]$ the cocycle dynamics and then rescaling the interval to unit length. However, this does not output a one-frequency cocycle, since an appropriate gluing must be made (37). This gluing is not canonical, so the nth renormalization (38) is only defined up to conjugation. See [AK1] and [AK2].

Let $\varepsilon_0>0$ be such that $A'\in C^\omega_{\varepsilon_0}(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{R}))$ and N admits an analytic extension to an open neighborhood of $\mathbb{R}$ containing $Q=[0,2]\times[-\varepsilon_0,\varepsilon_0]$. Let $C_0=\sup_{z\in Q}\|N(z)\|^2$. If k is an arbitrary integer, $l=l(k)$ is the unique integer such that $0\leqslant k\alpha'+l<1$ and $t=t(k)=k(-1)^nq_n+l(-1)^{n-1}q_{n-1}$, then we have

$$C_0^{-1}\leqslant\frac{\|A_t(y+\beta_{n-1}\varepsilon i)\|}{\|A'_k(x+\varepsilon i)\|}\leqslant C_0, \tag{40}$$

where $x,y\in\mathbb{C}/\mathbb{Z}$ are related by $y=x_0+\beta_{n-1}x$ and we assume that $|\mathrm{Im}(x+\varepsilon i)|<\varepsilon_0$. It follows that

$$\begin{aligned}\biggl|\int_{\mathbb{R}/\mathbb{Z}}\log\|A'_k(x+\varepsilon i)\|\,dx-\int_{\mathbb{R}/\mathbb{Z}}\log\|A_{(-1)^nt}(x+\beta_{n-1}\varepsilon i)\|\,dx\biggr|\\ =\biggl|\int_{\mathbb{R}/\mathbb{Z}}\log\|A'_k(x+\varepsilon i)\|\,dx-\int_{\mathbb{R}/\mathbb{Z}}\log\|A_t(x+\beta_{n-1}\varepsilon i)\|\,dx\biggr|\leqslant\log C_0.\end{aligned} \tag{41}$$

Notice that, when k is large, t satisfies

$$(-1)^n\frac{t}{k}=q_n-\frac{l}{k}q_{n-1}=q_n+\alpha_nq_{n-1}+o(1)=\frac{1}{\beta_{n-1}}+o(1).$$

It follows that, for large k,

$$\frac{1}{k}\int_{\mathbb{R}/\mathbb{Z}}\log\|A'_k(x+\varepsilon i)\|\,dx=\frac{1+o(1)}{\beta_{n-1}}\frac{1}{(-1)^nt}\int_{\mathbb{R}/\mathbb{Z}}\log\|A_{(-1)^nt}(x+\beta_{n-1}\varepsilon i)\|\,dx, \tag{42}$$

and taking the limit we get

$$L(\alpha',A'_\varepsilon)=\frac{1}{\beta_{n-1}}L(\alpha,A_{\beta_{n-1}\varepsilon}), \tag{43}$$

from which it follows that $\omega(\alpha,A)=\omega(\alpha',A')$.

If (α,A) is L^2-conjugate to rotations, then by the previous theorem we can take $\|A'-R_\theta\|_1<1$. This easily implies that $L(\alpha',A'_\varepsilon)<\log 2$ for $0<\varepsilon<1$, so

$$\omega(\alpha',A')\leqslant\frac{\log 2}{2\pi}<1$$

by convexity. Hence $\omega(\alpha',A')=0$ by quantization. □

Now, as $DL_{\xi,k}\cdot D_0v_\lambda$ is non-trivial, the implicit function theorem allows us to shrink U and change coordinates near 0 so that $L_{\xi,k}$ becomes a linear function $\tilde{L}(z_1,...,z_n)=z_n$.

The hypothesis implies that there exists $t_0\in\mathbb{R}^n$ such that $w=D_0v_\lambda\cdot t_0$ is v_0-signed and $DL_{\xi,k}(v_0)\cdot w=0$. By Corollaries 12 and 14, we have $t_0\neq 0$, so we may assume that $t_0=(1,0,...,0)$.

Shrinking U further, we may assume that $D_{\lambda_0}v_\lambda\cdot t_0$ is v_{λ_0}-signed at every λ_0 near 0. For every $(z_2,...,z_{n-1})$ and almost every z_1, if $(\alpha,A^{(v_{(z_1,...,z_{n-1},0)})})$ has zero Lyapunov exponent then, by Theorem 11, it is L^2-conjugate to rotations, and hence, by Corollary 14, its acceleration is zero, and thus $v_{(z_1,...,z_{n-1},0)}\notin\mathcal{C}^k$. This concludes the proof of Proposition 18. □

3.2. The critical regime as the boundary of non-uniform hyperbolicity

The indefiniteness estimate (Theorem 10) also has the following consequence.

THEOREM 15. *Let $\alpha\in\mathbb{R}\setminus\mathbb{Q}$ and let $v\in C^{\omega}(\mathbb{R}/\mathbb{Z},\mathbb{R})$. If $E\in\Sigma_{\alpha,v}$ is critical, then there exists a trigonometric polynomial w, and arbitrarily small $t>0$, such that E belongs to the spectrum and is supercritical for $H_{\alpha,v+tw}$.*

Proof. Let $\omega(\alpha,A^{(E-v)})=k>0$. Define as before a function $L_{\xi,k}$ on a neighborhood $\mathcal{V}\subset C^{\omega}_{\xi}(\mathbb{R}/\mathbb{Z},\mathbb{R})$ of $E-v$, for an appropriate choice of $\xi>0$. Choose an $(E-v)$-signed trigonometric polynomial w such that the derivative of $v'\mapsto L_{\xi,k}(E-v')$ at $v'=v$ and in the direction of w is zero. Let v_λ, with $v_0=v$, be an analytic family of trigonometric polynomials which are tangent to w at 0 and satisfy $L_{\xi,k}(E-v_\lambda)=0$. Let $N_{\alpha,v'}\mathbb{R}\to\mathbb{R}$ denote the integrated density of states of $H_{\alpha,v'}$.

By the usual monotonicity argument (see, e.g., [AK2, Lemma 2.4]), and since w is $(E-v)$-signed, $\lambda\mapsto N_{\alpha,v_\lambda}(E)$ is either non-increasing or non-decreasing for λ small. Moreover, since $(\alpha,A^{(E-v)})$ is not uniformly hyperbolic, it cannot be constant near 0.([22])

It follows that there exists a sequence $\lambda_n\to 0$ such that $N_{\alpha,v_{\lambda_n}}(E)\notin\mathbb{Z}\oplus\alpha\mathbb{Z}$. By the gap labeling theorem, this implies that $E\in\Sigma_{\alpha,v_{\lambda_n}}$ and it is accumulated from both sides by points in $\Sigma_{\alpha,v_{\lambda_n}}$.

Let w' be a trigonometric polynomial such that the derivative of $v'\mapsto L_{\xi,k}(E-v')$ at $v'=v$ and in the direction of w' is positive. For every n, there exists a sequence $0<\lambda'_{j,n}<1/j$ such that $N_{\alpha,v_{j,n}}(E)\notin\mathbb{Z}\oplus\alpha\mathbb{Z}$, where $v_{j,n}=v_{\lambda_n}+\lambda'_{j,n}w'$. Taking n and j large then E is supercritical for $H_{\alpha,v_{j,n}}$: On one hand, E belongs to the spectrum (by the gap labeling theorem), and on the other hand, $L_{\xi,k}(E-v_{j,n})>0$ by the choice of w', so by convexity we have $L(\alpha,A^{(E-v_{j,n})})\geqslant L_{\xi,k}(E-v_{j,n})$.

Note that in the generic case $N_{\alpha,v}(E)\notin\mathbb{Z}\oplus\alpha\mathbb{Z}$, the result can be obtained in a much simpler way (using only Theorem 8), since we do not need to assume that w is $(E-v)$-signed in order to find a sequence $0<\lambda_n<1/n$ such that $N_{\alpha,v_{\lambda_n}}(E)\notin\mathbb{Z}\oplus\alpha\mathbb{Z}$. □

3.3. Indefiniteness

Recall the setting of Theorem 10. We will need the expression for the derivative of $L_{\xi,k}:\mathcal{V}\to\mathbb{R}$ at $v\in\mathcal{V}$ that we previously derived. For each such v, there exists a maximal

([22]) Indeed, there exists $\varepsilon\in\{-1,1\}$ and $C>1$ such that, for all λ small, $N_{\alpha,v_\lambda}(E)$ belongs to the closed interval bounded by $N_{\alpha,v}(E+\varepsilon C^{-1}\lambda)$ and $N_{\alpha,v}(E+\varepsilon C\lambda)$. This, or rather the correspondig estimate for the *fibered rotation number* $\varrho=\frac{1}{2}(1-N)$, comes from a comparison of the ϱ-dependence in two monotonic families of cocycles (constructed by suitable coordinate change, see Remark 17). (A related argument appears in the proof of [AK2, Lemma 3.6].) So $\lambda\mapsto N_{\alpha,v_\lambda}(E)$ is non-constant (near $\lambda=0$) if and only if $\lambda\mapsto N_{\alpha,v}(E+\lambda)$ is, and this happens if and only if E is in the spectrum of $H_{\alpha,v}$, that is, $(\alpha,A^{(E-v)})$ is not uniformly hyperbolic.

interval $(\xi_-(v),\xi_+(v))\subset(0,\xi)$ containing ξ_0 such that the cocycle $(\alpha,A^{(v)})$ is uniformly hyperbolic through $\{z:\xi_-(v)<\operatorname{Im} z<\xi_+(v)\}$ (note in particular that $\xi_-(v_0)=0$). Let q_1, q_2 and q_3 be holomorphic functions in the band $\{x:\xi_-<\operatorname{Im} x<\xi_+\}$ given by the formulas for the coefficients of the derivative of the Lyapunov exponent for uniformly hyperbolic cocycles obtained in §2.3.3. Let u and s be the unstable and stable directions and let

$$B=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in \mathrm{SL}(2,\mathbb{C})$$

be such that the first column of B lies in u and the second column of B lies in the s. Then we have $q_1=ad+bc$, $q_2=cd$ and $q_3=-ab$. We let $q(z)=-q_3(z)$. Note that $q(z-\alpha)=q_2(z)$.

The expression for $DL_{\xi,k}(v)$ in a direction $w\in C^\omega_\xi(\mathbb{R}/\mathbb{Z},\mathbb{R})$ is

$$DL_{\xi,k}(v)\cdot w=\operatorname{Re}\int_{\mathbb{R}/\mathbb{Z}} q(x+\varepsilon i)w(x+\varepsilon i)\,dx,\quad \xi_-(v)<\varepsilon<\xi_+(v), \tag{44}$$

see (33).

We say that v is directed if $DL_{\xi,k}(v)\cdot w\neq 0$ for every real-symmetric trigonometric polynomial w with $w(x)>0$ for every $x\in\mathbb{R}/\mathbb{Z}$.

The main step in the proof of Theorem 10 is the following.

Theorem 16. *Assume that v_0 is directed. Then*

(1) *the non-tangential limits of u and s exist almost everywhere*;

(2) $\operatorname{Im} u(x)$ *and* $\operatorname{Im} s(x)$ *are non-zero and have the same sign almost everywhere*;

(3) $\operatorname{Re}[u(x)-s(x)]>0$ *almost everywhere*;

(4) *if $D(x)$ is the open real-symmetric disk with $u(x),s(x)\in\partial\mathbb{D}$, then $0\notin D(x)\cap\mathbb{R}$, but for every $\varepsilon>0$ there exists a positive measure set of x with $D(x)\cap(-\varepsilon,\varepsilon)>0$.*

We delay the proof to the next subsection. We will also need the following result, proved in §3.5.

Theorem 17. *Let $v\in C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$ be non-identically zero. Then there exist a neighborhood $\mathcal{U}$ of $A^{(v)}$ in $C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{R}))$ and analytic functions $\Phi:\mathcal{U}\to C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$ and $\Psi:\mathcal{U}\to C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{R}))$ such that*

(1) $\Psi(\tilde{A})(x+\alpha)\tilde{A}(x)\Psi(\tilde{A})(x)^{-1}=A^{(\Phi(\tilde{A}))}(x)$;

(2) *if $\tilde{A}=A^{(\tilde{v})}$ for some $\tilde{v}$, then $\Phi(\tilde{A})=\tilde{v}$ and $\Psi(\tilde{A})=\mathrm{id}$.*

Proof of Theorem 10. We start with the following simple consequence of Theorem 17.

Lemma 19. *There exist, for t near 0, analytic families*

$$v_t\in\mathcal{V}\quad \textit{and}\quad B_t\in C^\omega(\mathbb{R}/\mathbb{Z},\mathrm{SL}(2,\mathbb{R})),$$

with $B_0=\mathrm{id}$, *such that, for every* $x\in\mathbb{R}/\mathbb{Z}$, $B_t(x+\alpha)A^{(v)}(x)B_t(x)^{-1}=A^{(v_t)}(x)$ *and*

$$\frac{d}{dt}[B_t(x)\cdot 0]>0$$

at $t=0$.

Proof. Recall that $v_0\in C^\omega_{\xi'}(\mathbb{R}/\mathbb{Z},\mathbb{R})$ for some $\xi'>\xi$. Let $b\in C^\omega_\xi(\mathbb{R}/\mathbb{Z},\mathbb{R})$ be a positive analytic funtion such that bv_0 is a trigonometric polynomial. Then there exists a unique trigonometric polynomial a such that $a(x)+a(x+\alpha)=-b(x)v_0(x)$. Set $c(x)=-b(x-\alpha)$ and let

$$\eta=\begin{pmatrix} a & b \\ c & -a\end{pmatrix}.$$

Then, for small s,

$$e^{s\eta}(x+\alpha)A^{(v_0)}(x)e^{-s\eta}(x)=A^{(v_0+s\gamma)}(x)+O(s^2), \tag{45}$$

with $\gamma(x)=(a(x+\alpha)-a(x))v_0(x)+b(x+\alpha)-b(x-\alpha)$. Notice that v_0 is not identically zero, since $\omega(\alpha,A^{(v_0)})\neq 0$. Thus, by Theorem 17, there exist η_s and γ_s with $\|\eta_s\|_\xi=O(s^2)$ and $\|\gamma_s\|_\xi=O(s^2)$ such that

$$e^{\eta_s}(x+\alpha)e^{s\eta}(x+\alpha)A^{(v_0)}(x)e^{-s\eta}(x)e^{-\eta_s}(x)=A^{(v_0+s\gamma+\gamma_s)}(x). \tag{46}$$

Set $B_t=e^{\eta_t}e^{t\eta}$. Then

$$\frac{d}{dt}[B_t(x)\cdot 0]=b(x) \tag{47}$$

at $t=0$. □

LEMMA 20. *Let* v_t *be as in Lemma* 19. *There exists arbitrarily small* $t\in\mathbb{R}$ *such that* v_t *is not directed.*

Proof. We may assume that v_0 is directed, so there are disks $D(x)$ defined for almost every $x\in\mathbb{R}/\mathbb{Z}$ as in Theorem 16. Let B_t be as in Lemma 19, and let u_t and s_t be the unstable and stable directions for v_t. Note that $B_t(z)\cdot u(z)=u_t(z)$ and $B_t(z)\cdot s(z)=s_t(z)$. By Theorem 16, if v_t is directed for every $|t|<\varepsilon$, then for every measurable continuity point x_0 of $x\mapsto D(x)$, $B_t(x_0)\cdot D(x_0)$ must be a disk not containing 0. In particular, we must have $D(x_0)\cap M(x_0)=\varnothing$, where $M(x)$ is the set of all $B_t(x)^{-1}\cdot 0$, for $|t|<\varepsilon$.

Since there exists $\delta>0$ such that $(-\delta,\delta)\subset M(x)$ for every x, this contradicts Theorem 16. □

Let v_t be as in the conclusion of Lemma 20. Since v_t is not directed, there exists a trigonometric polynomial w with $DL_{\xi,k}(v_t)\cdot w=0$ and $\inf_{x\in\mathbb{R}/\mathbb{Z}}w(x)>0$. Define an analytic family $A^\lambda \in C^\omega_\xi(\mathbb{R}/\mathbb{Z}, \mathrm{SL}(2,\mathbb{R}))$, $\lambda\in\mathbb{R}$, by $A^\lambda(x)=B_t(x+\alpha)^{-1}A^{(v_t+\lambda w)}(x)B_t(x)$, so that $A^0=A^{(v_0)}$. Let Φ and Ψ be as in Theorem 17 with $v=v_0$ and $\delta=\xi$, and let $\tilde v_\lambda=\Phi(A^\lambda)$ and $\widetilde B_\lambda=\Psi(A^\lambda)B_t^{-1}$ for $\lambda\in\mathbb{R}$ small, so that $\tilde v_0=v_0$. Let

$$\widetilde w=\frac{d}{d\lambda}\tilde v_\lambda\bigg|_{\lambda=0} \quad \text{and} \quad d=\frac{d}{d\lambda}\widetilde B_\lambda\bigg|_{\lambda=0}\widetilde B_0^{-1}.$$

By construction, we have

$$A^{(\tilde v_\lambda)}(x)=\widetilde B_\lambda(x+\alpha)A^{(v_t+\lambda w)}(x)\widetilde B_\lambda(x)^{-1}, \quad |\mathrm{Im}\, x|<\xi. \tag{48}$$

Differentiating (48) and then multiplying on the left by $A^{(v_0)}(x)^{-1}$, we get

$$\begin{pmatrix} p0 & 0 \\ -\widetilde w(x) & 0\end{pmatrix}=A^{(v_0)}(x)^{-1}d(x+\alpha)A^{(v_0)}(x)-d(x)+B_t^{-1}(x)\begin{pmatrix} 0 & 0 \\ -w & 0\end{pmatrix}B_t(x). \tag{49}$$

From the definition of $L_{\xi,k}$, (48) implies that $L_{\xi,k}(\tilde v_\lambda)=L(v_t+\lambda w)$, and so

$$DL_{\xi,k}(v_0)\cdot\widetilde w=0.$$

Let us now show that

$$\begin{pmatrix} 0 & 0 \\ -\widetilde w & 0\end{pmatrix}$$

is v_0-signed. Since $\inf_{x\in\mathbb{R}/\mathbb{Z}} w(x)>0$, we have that

$$\begin{pmatrix} 0 & 0 \\ -w & 0\end{pmatrix}$$

is v_t-signed (see Remark 16), so there exists $b\in C^\omega(\mathbb{R}/\mathbb{Z}, \mathfrak{sl}(2,\mathbb{R}))$ such that

$$a(x)=A^{(v_t)}(x)^{-1}b(x+\alpha)A^{(v_t)}(x)-b(x)+\begin{pmatrix} 0 & 0 \\ -w & 0\end{pmatrix} \tag{50}$$

is signed, i.e., $\det a(x)>0$. Notice that (50) gives

$$\begin{aligned} B_t(x)^{-1}a(x)B_t(x)=&A^{(v_0)}(x)^{-1}B_t(x+\alpha)^{-1}b(x+\alpha)B_t(x+\alpha)A^{(v_0)}(x)\\ &-B_t(x)^{-1}b(x)B_t(x)+B_t(x)^{-1}\begin{pmatrix} 0 & 0 \\ -w(x) & 0\end{pmatrix}B_t(x).\end{aligned} \tag{51}$$

Let

$$\tilde b=B_t^{-1}bB_t-d \tag{52}$$

and let

$$\tilde a(x)=A^{(v_0)}(x)^{-1}\tilde b(x+\alpha)A^{(v_0)}(x)-\tilde b(x)+\begin{pmatrix} 0 & 0 \\ -\widetilde w & 0\end{pmatrix}. \tag{53}$$

Putting together (49), (51), (52) and (53), we get that $\tilde a=B_t^{-1}aB_t$, and thus $\det\tilde a=\det a$. In particular, since $\tilde a$ is signed, (53) gives that $\widetilde w$ is v_0-signed, as desired.

While $\widetilde w$ is not necessarily a trigonometric polynomial, it can be approximated by a trigonometric polynomial in the kernel of $DL_{\xi,k}(v_0)$, which will be v_0-signed as well (since v_0-signedness is an open condition). □

3.4. When the derivative of the Lyapunov exponent is a measure

In this subsection we prove Theorem 16. We let $v=v_0$ and $A=A^{(v)}$ for simplicity of notation.

The starting observation is that if v is directed then $DL_{\xi,k}(v)$ extends to a functional on $C^0(\mathbb{R}/\mathbb{Z},\mathbb{R})$ with norm $|DL_{\xi,k}(v)\cdot 1|$, which is either non-negative or non-positive on positive functions. By the Riesz representation theorem, it is given by a measure with finite mass μ on $\mathbb{R}/\mathbb{Z}$. By (44), this means that for any $w\in C^\omega_\xi(\mathbb{R}/\mathbb{Z},\mathbb{R})$ we have

$$\operatorname{Re}\int_{\mathbb{R}/\mathbb{Z}} q(x+\varepsilon i)w(x+\varepsilon i)\,dx=\int_{\mathbb{R}/\mathbb{Z}} w(x)\,d\mu(x),\quad 0<\varepsilon<\xi_0. \tag{54}$$

We will assume from now on that μ is non-negative, the other case being analogous.

Our plan is to show that the non-negativity of μ leads to good estimates for q which imply one of two conclusions:

(C1) either u or s extend analytically through $\mathbb{R}/\mathbb{Z}$;

(C2) the conclusion of Theorem 16 holds.

Let us first show that (C1) implies that $\omega(\alpha,A)=0$, which contradicts the standing hypothesis that $v\in\mathcal{C}^k$.

Assume for simplicity that u extends analytically. Then, either $u(x)\in\overline{\mathbb{R}}$ for every $x\in\mathbb{R}/\mathbb{Z}$, or $u(x)\notin\overline{\mathbb{R}}$ for every $x\in\mathbb{R}/\mathbb{Z}$ (since the $\mathrm{SL}(2,\mathbb{R})$ action preserves $\overline{\mathbb{R}}$ and $x\mapsto x+\alpha$ is minimal).

If $u(x)\notin\overline{\mathbb{R}}$ for $x\in\mathbb{R}/\mathbb{Z}$, this extends to $\operatorname{Im} x>0$ small. In this case we can select $a=u$ and $c=1$ when defining $B(x)$, and it follows that $\lambda=u$, so

$$L(\alpha,A_\varepsilon)=\int_{\mathbb{R}/\mathbb{Z}}\log|u(x+\varepsilon i)|\,dx$$

is independent of ε small (the argument of u is always different from $k\pi$, $k\in\mathbb{Z}$), thus $\omega(\alpha,A)=0$.

If $u(x)\in\overline{\mathbb{R}}$, we can use u to define analytic functions

$$A'\colon\mathbb{R}/\mathbb{Z}\longrightarrow\mathrm{SL}(2,\mathbb{R})\quad\text{and}\quad B'\colon\mathbb{R}/\mathbb{Z}\longrightarrow\mathrm{PSL}(2,\mathbb{R})$$

such that $A'(x)$ is upper triangular and $B'(x+\alpha)^{-1}A(x)B'(x)=A'(x)$: take the first column of B' parallel to $\binom{u}{1}$. We have $\omega(\alpha,A')=\omega(\alpha,A)$, and we just need to show that the $\omega(\alpha,A')=0$. But if

$$A'=\begin{pmatrix} a' & b'\\ 0 & d'\end{pmatrix},$$

then $L(\alpha,A'_\varepsilon)=\int_{\mathbb{R}/\mathbb{Z}}\log|a'(x+\varepsilon i)|\,dx$ for $\varepsilon>0$ small. This is independent of ε, since near $\mathbb{R}/\mathbb{Z}$ the argument of a' is always different from $\left(k+\frac{1}{2}\right)\pi$, $k\in\mathbb{Z}$. We conclude that $\omega(\alpha,A')=0$.

The remainder of this subsection is dedicated to showing that one of (C1) or (C2) always holds.

3.4.1. Non-tangential limits and analytic continuation

Recall that for any bounded holomorphic function $f:\mathbb{D}\to\mathbb{C}$, the non-tangential limits $f(z)=\lim_{r\to 1-} f(rz)$ exist for almost every $z\in\partial\mathbb{D}$ (see, e.g., [G]), and the Poisson formula holds, i.e. $f(0)=\int_0^1 f(e^{2\pi i\theta})\,d\theta$. Applying appropriate conformal maps, we see that, if $U\subset\mathbb{C}$ is any real-symmetric domain and $f:U\cap\mathbb{H}\to\mathbb{C}$ is a holomorphic function which is either bounded, or takes values in $\mathbb{H}$, or takes values in $\mathbb{C}\setminus(-\infty,0]$, then the non-tangential limits $f(x)=\lim_{\varepsilon\to 0+} f(x+\varepsilon i)$ also exist for almost every $x\in U\cap\mathbb{R}$.

We will use the following simple version of the Schwarz reflection principle.

PROPOSITION 21. *Let U be a real-symmetric domain and let $f:U\cap\mathbb{H}\to\mathbb{C}$ be holomorphic. Then, the following statements hold*:

(1) *if f takes values in $\mathbb{H}$ and the non-tangential limits at $U\cap\mathbb{R}$ are almost surely imaginary, then f extends analytically to a function on U such that $f(\bar z)=-\overline{f(z)}$*;

(2) *if $f:U\cap\mathbb{H}\to\mathbb{C}\setminus(-\infty,0]$ is holomorphic and its non-tangential limits at $U\cap\mathbb{R}$ are almost surely real, then f extends analytically to a function on U such that $f(\bar z)=\overline{f(z)}$.*

Proof. Assume that $\operatorname{Re} f(x)=0$ (resp. $\operatorname{Im} f(x)=0$ and $\operatorname{Re} f(x)>0$) for almost every $x\in U\cap\mathbb{R}$. Let $\phi:\mathbb{H}\to\mathbb{D}$ (resp. $\phi:\mathbb{C}\setminus(-\infty,0]\to\mathbb{D}$) be a conformal map commuting with the symmetry about the imaginary axis (resp. real axis). Then $\phi\circ f$ is bounded and its non-tangential limits are imaginary (real). Thus the usual Schwarz reflection principle applies.([23]) Since $\phi\circ f$ extends, the same holds for f. ☐

3.4.2. Initial estimates on q

Let us write $q(z)=if(z)+g(z)$, where f is analytic and real-symmetric on $\mathbb{R}/\mathbb{Z}$ and g is a holomorphic function with $\hat g_k=0$ for $k\leqslant -1$. Thus g is defined on $\operatorname{Im} z>0$ and is bounded at ∞.

LEMMA 22. *We have $\operatorname{Re} g(z)\geqslant 0$ for every z such that $\operatorname{Im} z>0$.*

([23]) The Schwarz reflection principle is usually stated assuming continuity at the boundary, the version for bounded holomorphic functions following immediately (as we can consider convolution approximations satisfying the continuity requirement). See also [G, Exercise 12, p. 95].

Proof. (All integrals in this proof are supposed to be over $\mathbb{R}/\mathbb{Z}$.) Let $\phi\colon\mathbb{R}/\mathbb{Z}\to\mathbb{R}$ be a positive C^∞ function with $\hat{\phi}_0=1$, and let

$$g^\phi(z)=\int g(z+x)\phi(x)\,dx.$$

It suffices to show that $\operatorname{Re} g^\phi(z)\geqslant 0$ for every such ϕ. Let

$$h^\phi(x)=\int \phi(y-x)\,d\mu(y),$$

which is a non-negative C^∞ function. Then, for any real-symmetric trigonometric polynomial w and any $\varepsilon>0$, we have

$$\begin{aligned}\int \operatorname{Re}[g^\phi(x+\varepsilon i)w(x+\varepsilon i)]\,dx &= \iint \operatorname{Re}[g(x+y+\varepsilon i)\phi(y)w(x+\varepsilon i)]\,dy\,dx\\ &= \int \phi(y)\int \operatorname{Re}[g(x+y+\varepsilon i)w(x+\varepsilon i)]\,dx\,dy\\ &= \int \phi(y)\int \operatorname{Re}[g(x'+\varepsilon i)w(x'-y+\varepsilon i)]\,dx'\,dy\\ &= \int \phi(y)\int w(x'-y)\,d\mu(x')\,dy\\ &= \iint \phi(y)w(x'-y)\,dy\,d\mu(x')\\ &= \iint \phi(x'-y)w(y)\,dy\,d\mu(x')\\ &= \iint \phi(x'-y)\,d\mu(x')w(y)\,dy\\ &= \int h^\phi(y)w(y)\,dy,\end{aligned} \tag{55}$$

where the first identity uses the definition of g^ϕ, the fourth one is (54) and the last one is the definition of h^ϕ. There exists a bounded holomorphic function H^ϕ on $\operatorname{Im} z>0$ which extends smoothly to $\operatorname{Im} z\geqslant 0$ and satisfies $\operatorname{Re} H^\phi(x)=h^\phi(x)$ (it can be constructed with the help of the Hilbert transform). Obviously $H^\phi(z)>0$ on $\operatorname{Im} z>0$ by the Poisson formula. If w is a real-symmetric trigonometric polynomial, we have

$$\int \operatorname{Re}[H^\phi(x+\varepsilon i)w(x+\varepsilon i)]\,dx=\int h^\phi(x)w(x)\,dx \tag{56}$$

for every $\varepsilon>0$. Since both $\widehat{H}^\phi_k=0$ and $\hat{g}^\phi_k=0$ for every $k\leqslant -1$, this implies that $\widehat{H}^\phi_k=\hat{g}^\phi_k$ for every $k\geqslant 1$ and $\operatorname{Re}\widehat{H}^\phi_0=\operatorname{Re}\hat{g}^\phi_0$. Thus $g^\phi-H^\phi$ is a purely imaginary constant and $\operatorname{Re} g^\phi(z)=\operatorname{Re} H^\phi(z)>0$ when $\operatorname{Im} z>0$. □

Since g takes values in a half-plane, it admits non-tangential limits. This allows us to make conclusions for q as well, and thus for almost every $x\in\mathbb{R}/\mathbb{Z}$ the non-tangential limits $q(x)=\lim_{\varepsilon\to 0} q(x+\varepsilon i)$ exist and satisfy $\operatorname{Re} q(x)\geqslant 0$. Notice that $\operatorname{Re} g(x)\in L^1$ (since $\operatorname{Re} g(z)>0$),[24] and hence $\operatorname{Re} q(x)\in L^1$.

3.4.3. The generic case

By a quick computation, we conclude that limits also exist, almost everywhere, for the unstable and the stable directions. Indeed, from $q(x)=a(x)b(x)$ and

$$q(x-\alpha)=a(x-\alpha)b(x-\alpha)=c(x)d(x),$$

we get

$$q(x)=\frac{u(x)s(x)}{u(x)-s(x)} \quad \text{and} \quad q(x-\alpha)=\frac{1}{u(x)-s(x)}, \tag{57}$$

from which we conclude that

$$1+4q(x)q(x-\alpha)=\left(\frac{u(x)+s(x)}{u(x)-s(x)}\right)^2. \tag{58}$$

Assume the non-tangential limits of q at x and $x-\alpha$ exist and are finite. If $q(x-\alpha)\neq 0$ then

$$u(x)-s(x)=\frac{1}{q(x-\alpha)} \quad \text{and} \quad u(x)s(x)=\frac{q(x)}{q(x-\alpha)}$$

define u and s uniquely up to a choice of sign for $\sqrt{1+4q(x)q(x-\alpha)}$. So the set of non-tangential accumulation values for each of u and s is made of one or two points, and since it must be connected, the non-tangential limits must be well defined. If $q(x-\alpha)=0$, then as z approaches x non-tangentially, either $u(z)$ is close to ∞ and $s(z)$ is close to $q(x)$, or $s(z)$ is close to ∞ and $u(z)$ is close to $-q(x)$. By the same argument as before, the non-tangential limits of u and s also exist in this case. In either case, we also conclude that the existence and finiteness of the non-tangential limits of $q(x)$ at x and $x-\alpha$ imply that $s(x)\neq u(x)$. Also, by the following lemma, u and s must be finite almost everywhere.

LEMMA 23. *Let $w:\mathbb{R}/\mathbb{Z}\to\overline{\mathbb{C}}$ be measurable and satisfy $A(x)\cdot w(x)=w(x+\alpha)$. Then $w(x)\neq\infty$ almost everywhere.*

Proof. Otherwise, there would exist $k,l>0$ and a positive measure set $X\subset\mathbb{R}/\mathbb{Z}$ such that $w(x)=w(x+k\alpha)=w(x+(kl+1)\alpha)=\infty$ for every $x\in X$. It follows from analyticity that $A_k(x)\cdot\infty=\infty$ and $A_{kl+1}(x)\cdot\infty=\infty$ for every x. Thus $A(x)\cdot\infty=\infty$ for every x, which is impossible since $A(x)\cdot\infty=v(x)$. □

[24] Indeed, by pointwise convergence, $\int_{\mathbb{R}/\mathbb{Z}} |\operatorname{Re} g(x)|\,dx$ is at most $\lim_{\varepsilon\to 0}\int_{\mathbb{R}/\mathbb{Z}} |\operatorname{Re} g(x+\varepsilon i)|\,dx$. Since $\operatorname{Re} g(z)>0$, we have that $\int_{\mathbb{R}/\mathbb{Z}} |\operatorname{Re} g(x+\varepsilon i)|\,dx$ is constant and equal to $\operatorname{Re}\hat{g}_0$.

Consider now the following possibilities:

(1) $s(x)=\bar{u}(x)\notin\overline{\mathbb{R}}$ for almost every $x\in\mathbb{R}/\mathbb{Z}$;

(2) $s(x), u(x)\in\overline{\mathbb{R}}$ for almost every $x\in\mathbb{R}/\mathbb{Z}$;

(3) $\bar{s}(x)\neq u(x)$ and $m(x)\in\mathbb{H}$ for almost every $x\in\mathbb{R}/\mathbb{Z}$, and for some choice of $m\in\{u, s, \bar{u}, \bar{s}\}$ (independent of x).

Those possibilities exhaust all cases since $x\mapsto x+\alpha$ is ergodic and $A(x)$ preserves $\mathbb{H}$, for $x\in\mathbb{R}/\mathbb{Z}$. We will now deal with the first two cases (which we call *generic* because the existence of three invariant sections is very constraining for an $\mathrm{SL}(2,\mathbb{R})$ cocycle) and leave the third, and hardest, case for §3.4.4.

In the first case, assuming, say, that $s(x)\in\mathbb{H}$, we have $\operatorname{Re} q(x)=0$ and $\operatorname{Im} q(x)>0$ for almost every $x\in\mathbb{R}/\mathbb{Z}$. Consider a decomposition $q=if+g$, where f is real-symmetric and g is holomorphic in $\mathbb{H}$ and bounded at ∞. We may also assume that $f(x)<0$ for $x\in\mathbb{R}/\mathbb{Z}$. As we saw, $\operatorname{Re} g\geqslant 0$ when $\operatorname{Im} x>0$ and now we also get that the non-tangential limits satisfy $\operatorname{Re} g(x)=0$ and $\operatorname{Im} g(x)>0$ for almost every $x\in\mathbb{R}/\mathbb{Z}$. By Proposition 21, $-ig$ admits an analytic continuation. This implies successively that q, u and s also admit analytic continuations, so we have reached conclusion (C1).

In the second case, $\operatorname{Im} q(x)=0$ for $x\in\mathbb{R}/\mathbb{Z}$. Consider a decomposition $q=f+g$, where f is analytic real-symmetric, with $f(x)<0$ for $x\in\mathbb{R}/\mathbb{Z}$, and g is holomorphic on $\operatorname{Im} x>0$ and bounded at ∞. By comparison with the decomposition considered before, $\operatorname{Re} g>0$ on $0<\operatorname{Im} x<\varepsilon$. Since $\operatorname{Im} q(x)=0$ and $\operatorname{Im} g(x)=0$ for almost every $x\in\mathbb{R}/\mathbb{Z}$, Corollary 21 implies that ig admits an analytic continuation. Hence q, u and s also admit analytic continuations, so we have reached conclusion (C1).

3.4.4. Many sections

We now consider the third case. We will assume that we can take $m=u$, the other possibilities being analogous. Notice that (α, A) admits at least three invariant sections u, s and $\bar{u}$.([25])

LEMMA 24. *We have* $\operatorname{Re} q(x)>0$ *for almost every* x.

Proof. Notice that $\operatorname{Re} q(x)=0$ implies that either $s(x+\alpha)=\infty$ or $u(x+\alpha)-s(x+\alpha)$ is purely imaginary and hence $\operatorname{Im} u\neq|\operatorname{Im} s|$. Let us show that the sets

$$X_\pm=\{x\in\mathbb{R}/\mathbb{Z}:\operatorname{Re} q(x)=0 \text{ and } \pm\operatorname{Im} u(x)>\pm|\operatorname{Im} s(x)|\}$$

([25]) This implies that there exists a measurable function $B\colon\mathbb{R}/\mathbb{Z}\to\mathrm{SL}(2,\mathbb{R})$ such that

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

have zero Lebesgue measure.

If $X_\pm$ has positive measure then there exist $k,l>0$ and a positive measure set of $x\in\mathbb{R}/\mathbb{Z}$ such that $x, x+k\alpha, x+(kl+1)\alpha\in X_\pm$. It follows that $A_k(x+\alpha)\cdot\infty=\infty$ and $A_{kl+1}(x+\alpha)\cdot\infty=\infty$.([26]) Since this happens for a positive measure set of x, this implies that $A_k(x)\cdot\infty=\infty$, $A_{kl+1}(x)\cdot\infty=\infty$, and hence $A(x)\cdot\infty=\infty$, hold for every $x\in\mathbb{R}/\mathbb{Z}$. But $A(x)\cdot\infty=v(x)\neq\infty$, yielding a contradiction. □

For real x, consider the real-symmetric open disk $D(x)$ with u and s on the boundary. If $0\in D$, then $\operatorname{Re}[u(x)-s(x)]>0$ implies that $\operatorname{Re}[u(x+\alpha)-s(x+\alpha)]<0$, which yields a contradiction. So $0\notin D$ for almost every x.

In order to show that (C2) holds, it remains to check that for every $\varepsilon>0$ there exists a positive measure set of $x\in\mathbb{R}/\mathbb{Z}$ such that $D(x)$ intersects $(-\varepsilon,\varepsilon)$.

Assume that this is not the case. Then $D(x)\cap\mathbb{R}\subset[-C,C]$, where $C=1/\varepsilon+\|v\|_0$. We claim that there exists $\varepsilon'>0$ such that $\operatorname{Re} q(x-\alpha)>\varepsilon'$ for almost every $x\in\mathbb{R}/\mathbb{Z}$. There are three cases to consider:

(1) $\operatorname{Im} s(x)=0$. Then $\operatorname{Re} q(x-\alpha)$ is the inverse of the diameter of D, so

$$\operatorname{Re} q(x-\alpha)=\frac{1}{2C}.$$

(2) $\operatorname{Im} s(x)>0$. Then $\operatorname{Re} q(x-\alpha)$ is the inverse of the diameter of the real-symmetric disk through $u(x)-\operatorname{Im} s(x)$ and $s(x)-\operatorname{Im} s(x)$, which is bigger than the diameter of D, so we get

$$\operatorname{Re} q(x-\alpha)>\frac{1}{2C}.$$

(3) $\operatorname{Im} s(x)<0$. Then

$$\operatorname{Re} q(x-\alpha)=\frac{1}{u(x)-\overline{s(x)}}\,\frac{|u(x)-\overline{s(x)}|^2}{|u(x)-s(x)|^2}. \tag{59}$$

We have

$$\frac{1}{u(x)-\overline{s(x)}}>\frac{1}{2C} \tag{60}$$

as in the previous case, so we just have to show that there is a lower bound for

$$\frac{|u(x)-\overline{s(x)}|}{|u(x)-s(x)|}. \tag{61}$$

([26]) If $z_1,z_2\in\mathbb{C}$ and $B\in\mathrm{SL}(2,\mathbb{R})$ are such that $\operatorname{Re} z_1=\operatorname{Re} z_2$, $\operatorname{Re}[B\cdot z_1]=\operatorname{Re}[B\cdot z_2]$, $\pm|\operatorname{Im} z_2|<\pm\operatorname{Im} z_1$ and $\pm|\operatorname{Im} B\cdot z_2|<\pm\operatorname{Im} B\cdot z_1$, then $B\cdot\infty=\infty$.

This is equivalent to showing that there is a lower bound for

$$\frac{|u(x)-\bar{s}(x)|}{2\operatorname{Im} u(x)}, \tag{62}$$

which is equivalent to showing that the hyperbolic distance in $\mathbb{H}$ between $u(x)$ and $\overline{s(x)}$, $d(x)>0$, is bounded from below. Since the hyperbolic metric is invariant by the $\mathrm{SL}(2,\mathbb{R})$ action, $d(x)=d(x+\alpha)$ for almost every $x\in\mathbb{R}/\mathbb{Z}$, and so, by ergodicity, $d(x)$ is constant.

It follows that $\operatorname{Re} q(z)$ is bounded away from 0 for every z with $0<\operatorname{Im} z<\delta$, for some small δ. Thus we can define $t(z)=\sqrt{1+4q(z)q(z-\alpha)}$, such that $\operatorname{Re} t(z)>0$ for every z with $0<\operatorname{Im} z<\delta$.([27]) Thus

$$u(x)=\frac{\pm t(x)+1}{2q(x-\alpha)} \quad \text{and} \quad s(x)=\frac{\pm t(x)-1}{2q(x-\alpha)}. \tag{63}$$

Also, we have

$$t(x)=\pm\frac{u(x)+s(x)}{u(x)-s(x)}.$$

Notice that

$$\operatorname{Re} t=\pm\frac{|u|^2-|s|^2}{|u-s|^2},$$

so $\operatorname{Re} t>0$ (by the choice of t) implies that $\pm|u|>\pm|s|$. Further, $\operatorname{Re}[u(x)-s(x)]>0$ and $\pm|u|>\pm|s|$ imply that $\pm\operatorname{Re} u(x)$ and $\pm\operatorname{Re} s(x)$ are positive.([28])

Thus, for almost every $x\in\mathbb{R}/\mathbb{Z}$, $\pm D(x)$ is contained in the right half-plane.

Let us assume that $D(x)$ is contained in the right half-plane, so that $D(x)\cap\mathbb{R}\subset(\varepsilon',C)$.

Let $\varepsilon'\leqslant z^-(x)<z^+(x)\leqslant C$ be the extremes of $D(x)\cap\mathbb{R}$. Notice that

$$\log\left\|A_n(x)\cdot\begin{pmatrix} z^\pm(x)\\ 1\end{pmatrix}\right\|\geqslant c\sum_{k=0}^{n-1}\log z^\pm(x+k\alpha), \tag{64}$$

where $c>0$ depends only on ε' and C. Since the Lyapunov exponent is 0, we must have

$$\lim_{k\to\infty}\frac{1}{k}\sum_{n=0}^{k-1}\log z^\pm(x+n\alpha)=0, \tag{65}$$

([27]) Notice that the arguments of $q(z)$ and $q(z-\alpha)$ can be taken in $\left(-\frac{1}{2}\pi,\frac{1}{2}\pi\right)$, and hence the argument of $q(z)q(z-\alpha)$ can be taken in $(-\pi,\pi)$, so that $1+4q(z)q(z-\alpha)\notin(-\infty,1]$.

([28]) If $\pm D(x)\subset\{z:\operatorname{Re} z>0\}$, then

$$s(x)\in\partial D(x)\cap\{z:\operatorname{Re} z<u(x)\}\subset\partial D(x)\cap\{z:\pm|z|<\pm u(x)\},$$

so $\pm|u(x)|>\pm|s(x)|$.

so that

$$\int_{\mathbb{R}/\mathbb{Z}} \log z^{\pm}(x)\, dx = 0, \tag{66}$$

which is impossible since $z^+(x) > z^-(x)$ almost everywhere.

The case where $-D(x)$ is contained in the right half-plane is analogous. Thus the proof of Theorem 16 is complete.

3.5. Conjugating $\mathrm{SL}(2,\mathbb{R})$ perturbations to Schrödinger form

Let $d \geqslant 1$ be an integer and $\delta \in \mathbb{R}^d_+$. Let $\Omega_\delta = \{z \in \mathbb{C}^d/\mathbb{Z}^d : |\mathrm{Im}\, z_k| < \delta_k\}$ and $C^\omega_\delta(\mathbb{R}^d/\mathbb{Z}^d, \cdot)$ stand for spaces of analytic functions on $\mathbb{R}^d/\mathbb{Z}^d$ with continuous extensions to $\bar{\Omega}_\delta$ which are holomorphic on Ω_δ.

We will prove the following generalization of Theorem 17 to arbitrary dimensions.

Theorem 18. *Assume that $v \in C^\omega_\delta(\mathbb{R}^d/\mathbb{Z}^d, \mathbb{R})$ is not identically zero. There exists $\varepsilon > 0$ such that, if $A' \in C^\omega_\delta(\mathbb{R}^d/\mathbb{Z}^d, \mathrm{SL}(2,\mathbb{R}))$ satisfies $\|A' - A^{(v)}\|_{C^\omega_\delta} < \varepsilon$, then there exist $v' \in C^\omega_\delta(\mathbb{R}^d/\mathbb{Z}^d, \mathbb{R})$ and $B' \in C^\omega_\delta(\mathbb{R}^d/\mathbb{Z}^d, \mathrm{SL}(2,\mathbb{R}))$, depending analytically on A', such that $B'(x+\alpha)A'(x)B'(x)^{-1} = A^{(v')}(x)$. Moreover, if A' is already of the form $A^{(\tilde{v})}$, then $v' = \tilde{v}$ and $B = \mathrm{id}$.*

A version of this result, for smooth cocycles over more general dynamical systems, was obtained in [ABD]. The proof of [ABD] makes use of partitions of unity to localize perturbations to some small region with disjoint first few iterates, one then tries to define functions in disjoint closed regions of space without worrying about interaction. The only additional care is to select the localizing region away from the critical locus $v(x+\alpha) = 0$, where the relevant equations develop singularities. Our approach is different: We take a disconnected finite cover of the dynamical system to realize the non-interacting condition, and concentrate on the linearized version of the problem, which can be broken up into several subproblems each of which involves a perturbation dominated by $v(x+\alpha)$ in such a way that it compensates the singularity.

Proof. Let $A = A^{(v)}$. Writing $A' = Ae^{s'}$, $B = e^w$ and $v' = v + t'$, we see that the linearized form of the problem is: For $s' \in C^\omega_\delta(\mathbb{R}^d/\mathbb{Z}^d, \mathfrak{sl}(2,\mathbb{R}))$, solve the equation

$$A(x)^{-1} w(x+\alpha) A(x) + s'(x) - w(x) = t'(x) L, \tag{67}$$

where

$$L = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.$$

We will show below how to obtain a solution (w,t') of (67) which is linear in s' and satisfies $\|w\|_{C^\omega_\delta}\leqslant C\|s'\|_{C^\omega_\delta}$ for some $C=C(v)>0$. Moreover, $C(v')$ will be uniformly bounded in a neighborhood of v. This allows one to construct the solution of the non-linear problem by, say, Newton's method.

Let

$$s'=\begin{pmatrix} s'_1 & s'_2 \\ s'_3 & -s'_1 \end{pmatrix}, \quad s=s'+s'_3L \quad \text{and} \quad t=t'+s'_3.$$

Then (67) is equivalent to

$$A(x)^{-1}w(x+\alpha)A(x)+s(x)-w(x)=t(x)L. \tag{68}$$

We will in fact construct a solution (w,t) to (68) which is linear in s and satisfies the required bounds. Notice that when A' is already of the form $A^{(\tilde v)}$, then $s=0$, so $w=0$ and the iterative procedure yields $v'=\tilde v$ and $B=\mathrm{id}$.

Choose $N\geqslant 4$ such that $|\sum_{k=2}^{N-2} v(x+k\alpha)^2|>1$ for every $x\in\Omega_\delta$.([29])

Note that N is constant in a neighborhood of v. Write

$$s_k(x)=\frac{v(x+k\alpha)^2}{\sum_{j=2}^{N-2} v(x+j\alpha)^2}s(x), \quad 2\leqslant k\leqslant N-2. \tag{69}$$

Let us show that there are, for $2\leqslant k\leqslant N-2$, functions $w_{k,l}$, with $l\in\mathbb{Z}_N$, and $t_{k,l}(x)$, with $l\in\{k-1,k,k+1\}$, such that

(1) $w_{k,0}=0$;

(2) $A(x)^{-1}w_{k,1}(x+\alpha)A(x)+s_k(x)-w_{k,0}(x)=0$;

(3) $A(x)^{-1}w_{k,l+1}(x+\alpha)A(x)-w_{k,l}(x)=t_{k,l}(x)L$, $l\in\{k-1,k,k+1\}$;

(4) $A(x)^{-1}w_{k,l+1}(x+\alpha)A(x)-w_{k,l}(x)=0$, $l\notin\{0,k-1,k,k+1\}$.

If we then set $w(x)=\sum_{k=2}^{N-2}\sum_{l=0}^{N-1} w_{k,l}(x)$ and $t(x)=\sum_{k=2}^{N-2}\sum_{l=k-1}^{k+1} t_{k,l}(x)$, we will have $A(x)^{-1}w(x+\alpha)A(x)+s(x)-w(x)=t(x)L$.

Conditions (1), (2) and (4) clearly define all $w_{k,l}$ except $w_{k,k}$ and $w_{k,k+1}$, in particular

$$w_{k,k-1}(x)=-A_{k-1}(x-(k-1)\alpha)s_k(x-(k-1)\alpha)A_{k-1}(x-(k-1)\alpha)^{-1}. \tag{70}$$

Using (69), we see that

$$\|w_{k,k-1}(x)\|\leqslant C|v(x+\alpha)|^2\|s(x-(k-1)\alpha)\|. \tag{71}$$

([29]) By unique ergodicity of $x\mapsto x+\alpha$ on $\mathbb{R}^d/\mathbb{Z}^d$, the Birkhoff averages of $v(z)^2$ converge uniformly to $\int_{\mathbb{R}^d/\mathbb{Z}^d} v(z+x)^2\,dx$, which equals $\int_{\mathbb{R}^d/\mathbb{Z}^d} v(x)^2\,dx>0$ by holomorphicity.

The key equations are thus

$$A(x)^{-1}w_{k,k}(x+\alpha)A(x)-w_{k,k-1}(x)=t_{k,k-1}(x)L, \tag{72}$$

$$A(x)^{-1}w_{k,k+1}(x+\alpha)A(x)-w_{k,k}(x)=t_{k,k}(x)L, \tag{73}$$

$$-w_{k,k+1}(x)=t_{k,k+1}(x)L. \tag{74}$$

From this we get an equation only involving unknown t's,

$$\begin{aligned}-w_{k,k-1}(x)=t_{k,k-1}(x)L+A(x)^{-1}t_{k,k}(x+\alpha)LA(x)\\ +A(x)^{-1}A(x+\alpha)^{-1}t_{k,k+1}(x)LA(x+\alpha)A(x).\end{aligned} \tag{75}$$

Once the t's satisfying (75) are known, one can determine the w's, so from now on we try to solve (75). Rewriting this equation we get

$$-w_{k,k-1}(x)=t_{k,k-1}(x)L+t_{k,k}L_1(x)+t_{k,k+1}L_2(x), \tag{76}$$

where

$$L_1(x)=\begin{pmatrix}-v(x) & 1\\ -v(x)^2 & v(x)\end{pmatrix} \tag{77}$$

and

$$L_2(x)=\begin{pmatrix}v(x+\alpha)-v(x)v(x+\alpha)^2 & v(x+\alpha)^2\\ -(1-v(x)v(x+\alpha))^2 & -v(x+\alpha)+v(x)v(x+\alpha)^2\end{pmatrix}. \tag{78}$$

Thus

$$L_1(x)-v(x)^2L=\begin{pmatrix}-v(x) & 1\\ 0 & v(x)\end{pmatrix} \tag{79}$$

and

$$L_2(x)-v(x+\alpha)^2L_1(x)+(2v(x)v(x+\alpha)-1)L=\begin{pmatrix}v(x+\alpha) & 0\\ 0 & -v(x+\alpha)\end{pmatrix}. \tag{80}$$

We conclude that if $v(x+\alpha)\neq 0$ then L, $L_1(x)$ and $L_2(x)$ span $\mathfrak{sl}(2,\mathbb{C})$, and there exists a unique solution $(t_{k,k-1},t_{k,k},t_{k,k+1})$ of (75), which is bounded by

$$C\frac{\|w_{k,k-1}(x)\|}{|v(x+\alpha)|}.$$

As mentioned before, the singularity that seems to arise when $v(x+\alpha)=0$ was well understood to be one source of difficulties in this problem, but here it emerges from (71) that whenever $v(x+\alpha)\neq 0$, the solutions are bounded by a constant times $|v(x+\alpha)|$. Hence they extend continuously as zero to $\{x:v(x+\alpha)=0\}$, and by holomorphic removability we conclude holomorphicity in Ω_δ. The result follows. □

Appendix A. Some almost Mathieu computations

Throughout this appendix, we let $v(x)=2\cos 2\pi x$.

THEOREM 19. *If* $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, $\lambda>0$, $E\in\mathbb{R}$ *and* $\varepsilon\geqslant 0$, *then*

$$L(\alpha, A_\varepsilon^{(E-\lambda v)}) = \max\{L(\alpha, A^{(E-\lambda v)}), \log\lambda+2\pi\varepsilon\}.$$

Proof. A direct computation shows that if E and λ are fixed, then for every $\delta>0$ there exists $0<\xi<\frac{1}{2}\pi$ such that if ε is large and $w\in\mathbb{C}^2$ makes an angle at most ξ with the horizontal line, then for every $x\in\mathbb{R}/\mathbb{Z}$, $w'=A^{(E-\lambda v)}(x+\varepsilon i)\cdot w$ makes an angle at most $\frac{1}{2}\xi$ with the horizontal line and $\left|\log\|w'\|-(\log\lambda+2\pi\varepsilon)\right|<\delta$.

This implies that

$$L(\alpha, A_\varepsilon^{(E-\lambda v)}) = 2\pi\varepsilon+\log\lambda+o(1) \quad \text{as } \varepsilon\to\infty.$$

By quantization of acceleration, for every ε sufficiently large, $\omega(\alpha, A_\varepsilon^{(E-\lambda v)})=1$ and $L(\alpha, A_\varepsilon^{(E-\lambda v)})=2\pi\varepsilon+\log\lambda$. By real-symmetry, $\omega(\alpha, A_\varepsilon^{(E-\lambda v)})$ is either 0 or 1 for $\varepsilon\geqslant 0$. This implies the given formula for $L(\alpha, A_\varepsilon^{(E-\lambda v)})$. □

For completeness, let us give a contrived rederivation of the Aubry–André formula.

COROLLARY 20. ([BJ1]) *If* $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, $\lambda>0$ *and* $E\in\mathbb{R}$, *then*

$$L(\alpha, A^{(E-\lambda v)}) \geqslant \max\{0, \log\lambda\},$$

with equality if and only if $E\in\Sigma_{\alpha,v}$.

Proof. The complement of the spectrum consists precisely of energies with positive Lyapunov exponent and zero acceleration (as those two properties characterize uniform hyperbolicity for $\mathrm{SL}(2,\mathbb{R})$-valued cocycles by Theorem 6).

The previous theorem gives the inequality, and shows that it is strict if and only if $L(\alpha, A^{(E-\lambda v)})>0$ and $\omega(\alpha, A^{(E-\lambda v)})=0$. □

A.1. Proof of the example theorem

Fix $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, $\lambda>1$ and $w\in C^\omega_\delta(\mathbb{R}/\mathbb{Z},\mathbb{R})$. Let $v_\varepsilon=\lambda v+\varepsilon w$.

LEMMA 25. *If* ε *is sufficiently small and* $E\in\Sigma_{\alpha,v_\varepsilon}$, *then* $\omega(\alpha, A^{(E-v_\varepsilon)})=1$.

Proof. By Theorem 19 and Corollary 20, we have that $L(\alpha, A^{(E-\lambda v)})\geqslant\log\lambda$ and $\omega(\alpha, A^{(E-\lambda v)})\leqslant 1$ for every $E\in\mathbb{R}$.

For ε small we have $\Sigma_{\alpha,v_\varepsilon}\subset[-4\lambda,4\lambda]$. By continuity of the Lyapunov exponent and upper semicontinuity of the acceleration, we get $\omega(\alpha,A^{(E-v_\varepsilon)})\leqslant 1$ and $L(\alpha,A^{(E-v_\varepsilon)})>0$ for every $E\in\Sigma_{\alpha,v_\varepsilon}$.

Since $A^{(E-v_\varepsilon)}$ is real-symmetric, $\omega(\alpha,A^{(E-v_\varepsilon)})\geqslant 0$ as well, and if $\omega(\alpha,A^{(E-v_\varepsilon)})=0$, with $E\in\Sigma_{\alpha,v_\varepsilon}$, then $(\alpha,A^{(E-v_\varepsilon)})$ is regular. This last possibility cannot happen: since the Lyapunov exponent is positive, this would imply uniform hyperbolicity, which cannot happen in the spectrum. We conclude that $\omega(\alpha,A^{(E-v_\varepsilon)})>0$ for $E\in\Sigma_{\alpha,v_\varepsilon}$. By quantization, this forces $\omega(\alpha,A^{(E-v_\varepsilon)})=1$. □

By Proposition 5, $E\mapsto L(\alpha,A^{(E-v_\varepsilon)})$ coincides in the spectrum with the restriction of the analytic function $E\mapsto L_{\delta,1}(\alpha,A^{(E-v_\varepsilon)})$ defined in some neighborhood. This concludes the proof of the example theorem.

Appendix B. Coexistence near critical coupling

Here we show that perturbations of the critical almost Mathieu operator (with potential $v(x)=2\cos 2\pi x$) may exhibit coexistence of subcritical and supercritical energies, and in fact one may have arbitrarily many alternances between subcritical and supercritical regimes. As far as we know, all previous examples of coexistence present only a small number of alternances.

Theorem 21. *Let $\alpha\in\mathbb{R}\setminus\mathbb{Q}$, $n\geqslant 1$ and $\{E_j\}_{j=1}^n$ be n distinct points in $\Sigma_{\alpha,v}$. Then, for any $\delta>0$, there exists a trigonometric polynomial $w\in C^\omega(\mathbb{R}/\mathbb{Z},\mathbb{R})$ such that for every $\varkappa\neq 0$ sufficiently small, and every $1\leqslant j\leqslant n$, there is $E_j^\varkappa\in\Sigma_{\alpha,v+\varkappa w}\cap(E_j-\delta,E_j+\delta)$ such that $E_j^\varkappa$ is subcritical if $(-1)^j\varkappa>0$ and $E_j^\varkappa$ is supercritical if $(-1)^j\varkappa<0$.*

Proof. For $H_{\alpha,v}$, all energies in the spectrum are critical, with zero Lyapunov exponent and acceleration 1 (see Appendix A). For $E\in\mathbb{C}\setminus\Sigma_{\alpha,v}$, we have that $(\alpha,A^{(E-v)})$ is uniformly hyperbolic (this is general) with zero acceleration (this is obvious for real energies and can be analytically continued to complex energies).

As for $(\alpha,A_\varepsilon^{(E-v)})$, we have that it is uniformly hyperbolic with zero acceleration for $0<\varepsilon<L(E)/2\pi$ and uniformly hyperbolic with acceleration 1 for $\varepsilon>L(E)/2\pi$ (here $L(E)=L(\alpha,A^{(E-v)})$). This follows from the asymptotic estimate $L(\alpha,A_\varepsilon^{(E-v)})=2\pi\varepsilon$ for $\varepsilon\gg 1$ (see the proof of Theorem 19). In particular, for $E\in\mathbb{C}\setminus\Sigma_{\alpha,v}$, $(\alpha,A_\varepsilon^{(E-v)})$ is not uniformly hyperbolic for $\varepsilon=L(E)/2\pi$.

Let U be the set of all E such that $L(\alpha,A^{(E-v)})<1$. This is an open neighborhood of $\Sigma_{\alpha,v}$. Following §3.3 define, for $E\in U$, a holomorphic function q^E on $\operatorname{Im} x>1/2\pi$ by

$q^E(x)=a^E(x)b^E(x)$, where

$$B^E=\begin{pmatrix} a^E & b^E \\ c^E & d^E \end{pmatrix}\in \mathrm{SL}(2,\mathbb{C})$$

has columns parallel to the unstable and the stable directions of $(\alpha, A^{(E-v)})$. Notice that q^E is holomorphic with respect to (E,x) and, for each $E\in U$, q^E admits holomorphic extensions up to $\operatorname{Im} x>L(E)/2\pi$. Thus, when $E\in\Sigma_{\alpha,v}$, q^E is defined in the entire upper half-plane $\mathbb{H}$.

Fix some $1\leqslant\xi_0<2$ and let $\mathcal{V}\subset C^\omega_2(\mathbb{R}/\mathbb{Z},\mathbb{R})$ be an open neighborhood of all $E-v$, $E\in U$, such that for every $v'\in\mathcal{V}$ the cocycle $(\alpha, A^{(v')}_{\xi_0})$ is uniformly hyperbolic with acceleration 1. Define $L_{2,1}\colon\mathcal{V}\to\mathbb{R}$ by $L_{2,1}(v')=L(\alpha, A^{(v')}_{\xi_0})-2\pi\xi_0$.

For $E\in U$ and $w\in C^\omega_2(\mathbb{R}/\mathbb{Z},\mathbb{R})$, the derivative of $t\mapsto L_{2,1}(E-v-tw)$ at $t=0$ is given by

$$-\int_{\mathbb{R}/\mathbb{Z}}\operatorname{Re} q^E(x+\varepsilon i)w(x+\varepsilon i)\,dx, \tag{81}$$

where ε can be chosen arbitrarily with $L(E)/2\pi<\varepsilon<2$ (see §3.3). Let Φ^E be the (bounded) linear functional on $C^\omega_2(\mathbb{R}/\mathbb{Z},\mathbb{R})$ taking w to (81).

We claim that, for every finite subset $\mathcal{E}\subset\Sigma_{\alpha,v}$ and any $E_*\in\Sigma_{\alpha,v}$, there exists $E'\in\Sigma_{\alpha,v}$ arbitrarily close to E_* such that $\Phi^{E'}$ is not a linear combination of the $\{\Phi^E\}_{E\in\mathcal{E}}$. Once this has been done, one can inductively obtain points $E'_j\in\Sigma_{\alpha,v}\cap\big(E_j-\frac{1}{2}\delta, E_j+\frac{1}{2}\delta\big)$ and a trigonometric polynomial $w\in C^\omega(\mathbb{R}/\mathbb{Z},\mathbb{R})$ such that $(-1)^j\Phi^{E'_j}\cdot w<0$. Now choose $0<r<\frac{1}{2}\delta$ such that $E'\in U$ for every $E'\in K_j=[E'_j-r, E'_j+r]$ and moreover $(-1)^j\Phi^{E'}\cdot w<0$. Then, for $\varkappa\neq0$ small and every $E'\in K_j$, we have $\varkappa(-1)^jL_{2,1}(E'-v-\varkappa w)<0$. Notice that if $\varkappa\neq0$ is small, then for every $E'\in\Sigma_{\alpha,v+\varkappa w}$ we have

(1) if $L_{2,1}(E'-v-\varkappa w)>0$ then E' is supercritical for $H_{\alpha,v+\varkappa w}$;

(2) if $L_{2,1}(E'-v-\varkappa w)<0$ then E' is subcritical for $H_{\alpha,v+\varkappa w}$.

Indeed, in the first case, we just use that $L\geqslant L_{2,1}$ (by convexity), and in the second case we notice that we must have $L\geqslant0>L_{2,1}$, so $\omega(\alpha, A^{(E'-v-\varkappa w)})<1$, hence by quantization $\omega(\alpha, A^{(E'-v-\varkappa w)})=0$ and $(\alpha, A^{(E'-v-\varkappa w)})$ is regular. The result then follows since, for every $\varkappa$ small, $\Sigma_{\alpha,v+\varkappa w}$ intersects each of the $\operatorname{int} K_j$ (as $\varkappa\mapsto\Sigma_{\alpha,v+\varkappa w}$ is continuous in the Hausdorff topology).

To conclude, let us prove the claim. Note that, by Theorem 8, $\Phi^{E_*}\neq0$, which in particular implies the claim when $\mathcal{E}$ is empty. We may assume that the Φ^E, $E\in\mathcal{E}$, are linearly independent. If the claim does not hold, then for every $E'\in\Sigma_{\alpha,v}$ close to E_*, $\Phi^{E'}$ is a linear combination of Φ^E, for $E\in\mathcal{E}$. Thus, we can write (in a unique way) $\Phi^{E'}=\sum_{E\in\mathcal{E}}c_E(E')\Phi^E$. Note that the coefficients $c_E(E')$, originally defined for E' near E_* in $\Sigma_{\alpha,v}$, coincide with restrictions of real-analytic functions defined in a small open interval I_* around E_*, which we still denote by $c_E(E')$.(30) Let $D \subset U$ be a small disk around E_* with $D \cap \mathbb{R} \subset I_*$ and such that $E' \mapsto c_E(E')$ extends holomorphically to D.

For $E' \in D$, we set $\gamma^{E'} = q^{E'} - \sum_{E \in \mathcal{E}} c_E(E') q^E$. Then, for $E' \in \Sigma_{\alpha,v} \cap D$, $\gamma^{E'}$ is a holomorphic function defined on the upper half-plane $\mathbb{H}$ such that

$$\int_{\mathbb{R}/\mathbb{Z}} \operatorname{Re} \gamma(x+\varepsilon i) w(x+\varepsilon i)\, dx = 0$$

for every $w \in C_2^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ and any $0 < \varepsilon < 2$. This implies that, for $E' \in \Sigma_{\alpha,v} \cap \mathbb{D}$, $\gamma^{E'}$ extends to an entire function, which is purely imaginary on $\mathbb{R}$ (see §2.4). But for every $E' \in D$, $\gamma^{E'}$ defines a holomorphic function on $\{z : \operatorname{Im} z > 1/2\pi\}$ which depends holomorphically on E'. Since $\Sigma_{\alpha,v} \cap D$ has positive logarithmic capacity (see [S, Theorem 7.2]), $\gamma^{E'}$ must define an entire function for every $E' \in D$ (just use Hartogs' theorem). It follows that $q^{E'} = \gamma^{E'} + \sum_{E \in \mathcal{E}} c_E(E') q^E$ defines a holomorphic function on $\mathbb{H}$ for every $E' \in D$. By a similar argument as in §3.4.3, we can, for every $E' \in D$, analytically continue the unstable and stable directions of $(\alpha, A^{(E'-v)})$ defined for $\operatorname{Im} x > L(E')/2\pi$ to holomorphic functions $u^E, s^E : \mathbb{H} \to \mathbb{PC}^2$ which satisfy $A^{(E'-v)}(x) \cdot u^{E'}(x) = u^{E'}(x+\alpha)$, $A^{(E'-v)}(x) \cdot s^{E'}(x) = s^{E'}(x+\alpha)$ and $u^{E'}(x) \neq s^{E'}(x)$ for every $x \in \mathbb{H}$.

Since $L(\alpha, A_\varepsilon^{(E'-v)}) > 0$ for every E' and any $\varepsilon > 0$, this implies that $(\alpha, A_\varepsilon^{(E'-v)})$ is uniformly hyperbolic for any $\varepsilon > 0$ and every $E' \in D$. But this cannot happen when $E' \in D \setminus \Sigma_{\alpha,v}$, as $(\alpha, A_\varepsilon^{(E'-v)})$ is not uniformly hyperbolic when $2\pi\varepsilon = L(E')$. This gives a contradiction and proves the claim. □

Remark 26. For perturbations of the almost Mathieu operator, the acceleration is bounded by 1, which implies that the number of alternances between the subcritical, critical and supercritical regimes is always finite. Indeed, for any $v' \in C_\xi^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$, and near any critical energy $E_0 \in \Sigma_{\alpha,v'}$ with acceleration 1, we can define an analytic function $L_{\xi,1}$ as before which has the property that energies $E \in \Sigma_{\alpha,v'}$ near E_0 are supercritical, critical, or subcritical according to whether $L_{\xi,1} > 0$, $L_{\xi,1} = 0$ or $L_{\xi,1} < 0$.

Remark 27. For $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $w \in C_\xi^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ and $\varkappa$ small, one may investigate the transition from subcriticality to supercriticality within the one-parameter family of operators $H_{\alpha,\lambda(v+\varkappa w)}$, $\lambda > 0$. It is convenient to look simultaneously at all $\Sigma_{\alpha,\lambda(v+\varkappa w)}$ in the (E, λ)-plane. Our work implies that there is a (possibly disconnected) nearly horizontal analytic curve $L_{\xi,1}^{-1}(0)$,(31) close to $\Sigma_{\alpha,v} \times \{1\}$, which separates the subcritical energies (below it) and the supercritical energies (above it). From the point of view of this paper, the study of this family is straightforward, since transversality can be checked by the

(30) Fix a $\#\mathcal{E}$-dimensional subspace $F \subset C_2^\omega(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ such that the $\Phi^E|_F$, $E \in \mathcal{E}$, are linearly independent, and define $c_E(E')$ in a neighborhood of E_* so to have $\Phi^{E'}|_F = \sum_{E \in \mathcal{E}} c_E(E') \Phi^E|_F$.

(31) Here we use the explicit computation $L_{\xi,1}(E - \lambda v) = \log \lambda$ for λ near 1 and E near $\Sigma_{\alpha,v}$.

direct computation of $L_{\xi,1}$ in the almost Mathieu case. In particular, since the critical curve is nearly horizontal, it defines a premonotonic family of cocycles, so the arguments in §3.1 show that the intersection of this curve with the spectra has zero linear measure.

References

[A1] Avila, A., The absolutely continuous spectrum of the almost Mathieu operator. Preprint, 2008. arXiv:0810.2965 [math.DS].

[A2] — Almost reducibility and absolute continuity I. Preprint, 2010. arXiv:1006.0704 [math.DS].

[A3] — KAM, Lyapunov exponents and the spectral dichotomy for one-frequency schrödinger operators. In preparation.

[ABD] Avila, A., Bochi, J. & Damanik, D., Cantor spectrum for Schrödinger operators with potentials arising from generalized skew-shifts. *Duke Math. J.*, 146 (2009), 253–280.

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

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

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

[AJS] Avila, A., Jitomirskaya, S. & Sadel, C., Complex one-frequency cocycles. *J. Eur. Math. Soc.* (*JEMS*), 16 (2014), 1915–1935.

[AK1] Avila, A. & Krikorian, R., Reducibility or nonuniform hyperbolicity for quasiperiodic Schrödinger cocycles. *Ann. of Math.*, 164 (2006), 911–940.

[AK2] — Monotonic cocycles. To appear in *Invent. Math.*

[AS] Avron, J. & Simon, B., Almost periodic Schrödinger operators. II. The integrated density of states. *Duke Math. J.*, 50 (1983), 369–391.

[BL] Benyamini, Y. & Lindenstrauss, J., *Geometric Nonlinear Functional Analysis*. Vol. 1. American Mathematical Society Colloquium Publications, 48. Amer. Math. Soc., Providence, RI, 2000.

[Bj1] Bjerklöv, K., Positive Lyapunov exponent and minimality for a class of one-dimensional quasi-periodic Schrödinger equations. *Ergodic Theory Dynam. Systems*, 25 (2005), 1015–1045.

[Bj2] — Explicit examples of arbitrarily large analytic ergodic potentials with zero Lyapunov exponent. *Geom. Funct. Anal.*, 16 (2006), 1183–1200.

[B] Bourgain, J., *Green's Function Estimates for Lattice Schrödinger Operators and Applications*. Annals of Mathematics Studies, 158. Princeton Univ. Press, Princeton, NJ, 2005.

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

[BJ1] Bourgain, J. & Jitomirskaya, S., Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential. *J. Stat. Phys.*, 108 (2002), 1203–1218.

[BJ2] — Absolutely continuous spectrum for 1D quasiperiodic operators. *Invent. Math.*, 148 (2002), 453–463.

[C] Christensen, J. P. R., On sets of Haar measure zero in abelian Polish groups, in *Proceedings of the International Symposium on Partial Differential Equations and the Geometry of Normed Linear Spaces* (Jerusalem, 1972). *Israel J. Math.*, 13 (1972), 255–260.

[E] Eliasson, L. H., Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation. *Comm. Math. Phys.*, 146 (1992), 447–482.

[G] Garnett, J. B., *Bounded Analytic Functions*. Graduate Texts in Mathematics, 236. Springer, New York, 2007.

[GS1] Goldstein, M. & Schlag, W., Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. *Ann. of Math.*, 154 (2001), 155–203.

[GS2] — Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues. *Geom. Funct. Anal.*, 18 (2008), 755–869.

[GS3] — On resonances and the formation of gaps in the spectrum of quasi-periodic Schrödinger equations. *Ann. of Math.*, 173 (2011), 337–475.

[H] Herman, M. R., Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnol′d et de Moser sur le tore de dimension 2. *Comment. Math. Helv.*, 58 (1983), 453–502.

[HPS] Hirsch, M. W., Pugh, C. C. & Shub, M., *Invariant Manifolds*. Lecture Notes in Mathematics, 583. Springer, Berlin–Heidelberg, 1977.

[HSY] Hunt, B. R., Sauer, T. & Yorke, J. A., Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc.*, 27 (1992), 217–238.

[J] Jitomirskaya, S., Metal-insulator transition for the almost Mathieu operator. *Ann. of Math.*, 150 (1999), 1159–1175.

[JKS] Jitomirskaya, S., Koslover, D. A. & Schulteis, M. S., Continuity of the Lyapunov exponent for analytic quasiperiodic cocycles. *Ergodic Theory Dynam. Systems*, 29 (2009), 1881–1905.

[S] Simon, B., Equilibrium measures and capacities in spectral theory. *Inverse Probl. Imaging*, 1 (2007), 713–772.

[SS] Sorets, E. & Spencer, T., Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials. *Comm. Math. Phys.*, 142 (1991), 543–566.

[Y] Young, L.-S., Lyapunov exponents for some quasi-periodic cocycles. *Ergodic Theory Dynam. Systems*, 17 (1997), 483–504.

Artur Avila
CNRS, IMJ-PRG, UMR 7586
Université Paris Diderot
Sorbonne Paris Cité
Sorbonnes Universités
UPMC Université Paris 06
FR-75013 Paris
France

and

Instituto Nacional de Matemática Pura e Aplicada
Estrada Dona Castorina 110
Rio de Janeiro 22460-320
Brasil
artur@math.univ-paris-diderot.fr

Received April 8, 2013
Received in revised form June 15, 2015