Asymptotic laws for a class of quasi-periodic Schrödinger cocycles at the lowest energy of the spectrum

Thomas Ohlson Timoudas

September 17, 2018

Abstract

Let \((\omega, A_E)\) be a quasi-periodic Schrödinger cocycle, where \(\omega\) is a Diophantine irrational. The potential is assumed to be \(C^2\) with a unique non-degenerate minimum, and the coupling constant is assumed to be large.

We show that, as the energy approaches the lowest energy of the spectrum from below, the distance between the Oseledets-directions, in projective coordinates, is asymptotically linear. Moreover, we show that the \(C^2\)-norm of the Oseledets-directions, in projective coordinates, grows asymptotically (almost) like the inverse of the square root of the distance.

Both of these results confirm numerical observations.

1 Introduction

Consider a cocycle \(A : \mathbb{T} \to \text{SL}(2, \mathbb{R})\) over an irrational circle rotation, given by
\[
(\theta, x) \mapsto (\theta + \omega, A(\theta)x),
\]
where \(\omega\) is irrational. In this paper we shall consider the family of cocycles
\[
A_E(\theta) = \begin{pmatrix} 0 & 1 \\ -1 & \lambda v(\theta) - E \end{pmatrix},
\]
where \(E \in \mathbb{R}\) is a parameter (the energy), \(\lambda > 0\) the coupling constant, and \(v : \mathbb{T} \to \mathbb{R}\). The resulting system is called a quasi-periodic Schrödinger cocycle, due to its relation to the Schrödinger equation. For more information about this connection, we refer to [Dam17]. Set
\[
A^n(\theta) = \begin{cases} 
A(\theta + (n-1)\omega) \cdots A(\theta) & n \geq 1, \\
\text{Id} & n = 0, \\
A(\theta - n\omega)^{-1} \cdots A(\theta - \omega)^{-1} & n \leq -1.
\end{cases}
\]
For every \(E\) we have an important quantity \(L(E)\), the (top) Lyapunov exponent. For Lebesgue-a.e. \(\theta \in \mathbb{T}\), it holds that
\[
L(E) = \lim_{n \to \infty} \frac{1}{n} \log \|A^n_E(\theta)\| \geq 0.
\]
We say that a cocycle \(A\) is uniformly hyperbolic if there are two continuous functions \(W^u, W^s : \mathbb{T} \to \text{Gr}(1, \mathbb{R}^2)\) spanning the whole space \((W^u(\theta) \oplus W^s(\theta) = \mathbb{R}^2)\), that are invariant
\[
A(\theta)W^s(\theta) = W^s(\theta + \omega), \quad \text{and} \quad A(\theta)W^u(\theta) = W^u(\theta + \omega),
\]
(1.2)
and satisfy for some $c > 0, 0 < r < 1$ that
\[
\|A^n(\theta)v\| \leq cr^n\|v\|, \text{ for } v \in W^s(\theta), \text{ and}
\]
\[
\|A^{-n}(\theta)v\| \leq cr^n\|v\|, \text{ for } v \in W^u(\theta),
\] (1.3)
for every $n \geq 0$, and $\theta \in \mathbb{T}$. We call $W^u$ and $W^s$ the unstable and stable subspaces, respectively. Since they are continuous and span the whole of $\mathbb{R}^2$, it is clear that the minimum angle between the spaces is bounded away from 0:
\[
\min_{\theta \in \mathbb{T}} \angle(W^s(\theta), W^u(\theta)) > 0.
\]
In summary, if $L(E) > 0$ and we have such a continuous splitting, the cocycle is uniformly hyperbolic. In the case that $L(E) > 0$ but there is not such continuous splitting, the cocycle is called non-uniformly hyperbolic, the splitting is only measurable, and
\[
\inf_{\theta \in \mathbb{T}} \angle(W^s(\theta), W^u(\theta)) = 0.
\]
In this case, the constant $c$ in (1.3), will depend non-uniformly on $\theta$.

Naturally, we ask ourselves how a system can bifurcate from uniformly hyperbolic behaviour, to non-uniformly hyperbolic. In [HdlIL06], they numerically studied how the minimum distance and Lyapunov exponent behaves at the bifurcation point, but for a different class of systems. Their findings were that
\[
\min_{\theta \in \mathbb{T}} \angle(W^s_t(\theta), W^s_t(\theta)) \sim t - t_0, \text{ and}
\]
\[
L(t) - L(t_0) \sim (t - t_0)^\alpha, \text{ for some } \alpha < 0,
\]
where $t$ is a parameter and the bifurcation happens at the critical parameter $t_0$. That is, the angle between the directions was observed to behave asymptotically linearly in the parameter, and the Lyapunov exponent according to some power law in the parameter.

Recently, the linear behaviour of the angle was verified for a certain class of systems, in [BS08]. In a different setting, the distance between two invariant tori was shown to behave asymptotically linearly at the point of collision, in [OT17].

More generally, we may ask how the directions of these subspaces (the curves given by their graphs) merge at the point of collision. In [Her83, 4.14], there is a discussion about this process. One of the problems given there, about minimal sets, was answered positively in the paper [Bjc07], for a class of Schrödinger cocycles. That result was later generalized to a larger class of systems (without linear structure) in [FGJ17]. We believe that the results in this paper should also be possible to generalize in the same way.

2 Our results

In this paper, we will assume that $v : \mathbb{T} \to \mathbb{R}$ is a $C^2$ function, having a unique non-degenerate minimum. We will consider the system
\[
A_E(\theta) = \begin{pmatrix}
0 & 1 \\
-1 & \lambda^2 v(\theta) - E
\end{pmatrix},
\] (2.1)
with the coupling constant \( \lambda \) in (1.1) replaced by \( \lambda^2 \). Since we consider only positive coupling constants, this is no restriction. This system is exactly the one considered in [Bje07], where, using methods similar to the ones in [You97], it was shown that \( L(E) > 0 \) uniformly for \( E \in (-\infty, E_0] \), where \( E_0 \) is the lowest energy of the spectrum, provided \( \lambda \) is large enough.

From now on, we will use projective coordinates \((1, r)\), and let \( r \) represent the direction \((1, r)\), and \( \infty \) the direction \((0, 1)\). Then \( A_E \) gives us the projective cocycle

\[
\Phi_E(\theta, r) = (\theta + \omega, \lambda^2 v(\theta) - E - 1/r).
\]

Note that, given \( r \), it is possible to recover the expansion rate of the original system \( A_E \), since

\[
A_E(\theta) \left( \begin{array}{c} 1 \\ r \end{array} \right) = r \left( \begin{array}{c} 1 \\ \Phi_E(\theta, r) \end{array} \right).
\]

Given invariant subspaces \( W_u^E \) and \( W_s^E \) of the cocycle \( A_E \) as above, we obtain directions that are invariant under \( \Phi_E \). That is, (1.2) gives us functions \( r_u^E : T \to \hat{\mathbb{R}} \) and \( r_s^E : T \to \hat{\mathbb{R}} \), where the projective line \( \hat{\mathbb{R}} \) is simply the real line together with a point at infinity, that satisfy the invariance relations

\[
(\theta + \omega, r_u^E(\theta + \omega)) = \Phi_E(\theta, r_u^E(\theta)), \quad \text{and}

(\theta + \omega, r_s^E(\theta + \omega)) = \Phi_E(\theta, r_s^E(\theta)).
\]

When \( E < E_0 \), where \( E_0 \) is the lowest energy of the spectrum of the corresponding Schrödinger operator, the graphs will satisfy \( \frac{1}{C} < r_s^E < r_u^E < C \) for some \( C \) uniformly in \( \theta \) and \( E < E_0 \). This is explained and shown in [Her83, 4.8–4.14].

![Invariant cone and development of peaks](image)

(a) Invariant cone \((0, \pi/2)\) (shaded gray) when \( E < E_0 \). Lower curve (red) is the stable direction, and the upper curve (blue) the unstable one.

(b) Development of peaks as \( E \) approaches \( E_0 \). Lower curve (red) is the stable direction, the upper one (blue) the unstable one.

Figure 1: For the simulations we used the almost-Mathieu potential \( \cos(2\pi \theta) \) with \( \omega = (\sqrt{5} - 1)/4 \) and \( \lambda^2 = 30 \).

In fig. 1, the invariant directions have been approximated through simulation, and we can clearly see that they lie within some positive cone in \((0, \pi/2)\). Due to numerical reasons, we were not able to reliably depict the development of successive peaks (the wrinkling process). For illustrations of that process at a more advanced stage, we refer to [BS08, FH15, OT17], and their references. We wish to study this process.

In order to quantify what we mean by wrinkling, we have chosen to focus on the \( C^1 \)-norm of the curves \( r_u^E(\theta) \) and \( r_s^E(\theta) \). Since the curves are in some invariant set \([\frac{1}{C}, C]\) (where \( C \) is positive), when
$E < E_0$, the norm is finite. We remark that it is in general not finite, and other coordinates may be more appropriate for treating general energies. We show that the second derivatives blow up according to the asymptotic law

$$\frac{1}{C_1} |E - E_0|^{-1/2 - \varepsilon} \leq \|r'_E\|_{C^1} \leq C_1 |E - E_0|^{-1/2},$$

where $C_1 > 0$ and $\varepsilon$ goes to 0 as $E \to E_0$. The same type of asymptotics holds for $r''_E$. Our method also gives local information about the $C^1$-norm. Higher derivatives could be studied using the same method, but it is not clear to us how one might achieve this without avoiding long computations.

The results and methods are similar to the ones in the paper [OT17], by the present author, where the system was given by a quasi-periodically forced logistic map. In that paper, there are two invariant graphs, one repelling and one attracting. There, the attracting graph $\psi_t$ satisfied the asymptotics

$$\frac{1}{C_1} (t - t_0)^{-1/2} \leq \|\psi_t\|_{C^1} \leq C_1 (t - t_0)^{-1/2},$$

at some critical parameter $t_0$, and the repelling one was just 0 at every point. That repelling graph was in an expanding region at every point, whereas in the present model, the repelling graph $r'_E$ cycles between expanding and contracting regions. This cycling is exactly why there is a loss of exponent, and we suspect that it can not be removed. However, we do remark that, for a large measure of parameters, there is no $\varepsilon$ in the lower bound.

Such norms have also been studied numerically, and found to satisfy similar power laws. For instance, in [FH15], they numerically observe a similar asymptotic, but for the blow-up of the $C_2$-norm.

In general, the Lyapunov exponent is not continuous (see [WY13], for an example where the potential is perturbed). At least for analytic potentials, it has been shown to be continuous in the parameter $E$ (see [BJ02]). For such potentials, the Lyapunov exponent is known to be at least Hölder continuous in the parameter $E$ (see [Bou00, GS01]).

In a subsequent paper, written jointly with Jordi-Lluís Figueras, we will show an asymptotic law for the Lyapunov exponent in the same setting as the one considered here, as $E \to E_0$.

We are now ready to state our main results. Denote by $E_0$ be the lowest energy in the spectrum, and let $\psi^u_E : T \to \hat{\mathbb{R}}$ and $\psi^s_E : T \to \hat{\mathbb{R}}$ be the unstable and stable projective directions, respectively, where $\hat{\mathbb{R}}$ is the real line with a point at infinity. We only consider irrationals $\omega$ satisfying the Diophantine condition

$$\inf_{p \in \mathbb{Z}} |n\omega - p| > \frac{\kappa}{|n|^\tau}, \text{ for every } n \in \mathbb{Z} \setminus \{0\}, \quad (DC)_{\kappa, \tau}$$

for some constants $\kappa > 0$ and $\tau \geq 1$. This condition allows us to obtain lower bounds on return times.

**Theorem 2.1** (Main result). Suppose that $\omega$ satisfies $(DC)_{\kappa, \tau}$, and the potential $v : T \to \mathbb{R}$ is $C^2$ and has a unique minimum. Then there is a $\lambda_0(\omega) > 0$ such that, if $\lambda > \lambda_0$, the minimum distance (in projective coordinates) between $\psi^u_E$ and $\psi^s_E$, is attained in a unique point $\theta_c(E)$ depending only on $E$, and is asymptotically linear:

$$\delta(E) = \min_{\theta \in T} |\psi^u_E(\theta) - \psi^s_E(\theta)| = C_1 \cdot (E_0 - E) + o(E_0 - E), \quad (2.2)$$

as $E \to E_0$, where $C_1 > 0$ is independent of $E$.  

4
Furthermore, there is a positive $\varepsilon = \varepsilon(E)$ satisfying $\lim_{E \to E_0} \varepsilon = 0$, and a $C_2 > 0$ independent of $E$, such that

$$\frac{1}{C_2} \cdot \left( \frac{1}{\sqrt{d(E)}} \right)^{1-\varepsilon} \leq \| \psi^E \|_{C^1} \leq C_2 \cdot \frac{1}{\sqrt{d(E)}},$$

and the same inequality is true if we replace $\| \psi^E \|_{C^1}$ with $\| \psi^E_s \|_{C^1}$.

Using the first statement about $d(E)$, the second one reduces to the inequality

$$\frac{1}{C} \cdot \left( \frac{1}{\sqrt{E_0 - E}} \right)^{1-\varepsilon} \leq \| \psi^E \|_{C^1} \leq C \cdot \frac{1}{\sqrt{E_0 - E}},$$

where the constant $C > 0$ and independent of $E$. We obtain a similar inequality for $\| \psi^E_s \|_{C^1}$.

Remark. For a relatively large set of $E$ close to $E_0$ (in the sense of Lebesgue measure), we can in fact get rid of this $\varepsilon$. That is, up to uniform constants, the asymptotics behaves like the square root for most energies. By increasing $\lambda$, the relative measure of such energies can be made arbitrarily close to full.

However, it also seems like the $\varepsilon$ can not be removed. That is, for some positive measure of energies (going to 0 as $\lambda$ increases), the $\varepsilon$ can not be removed!

We stress that the methods in this paper do not rely on the linear structure of the model, and should be possible to generalize to other systems. However, the asymptotics obtained in this paper may not be universal, but depend on resonances and certain properties of the forcing map. We will shed some light on this dependence in the the next section, where we discuss the mechanisms behind the process. We are confident that the methods contained in this paper can be extended to cover the spectral gaps as well; however, this may need some further work to obtain appropriate estimates for the spectral gaps. The reason we have chosen to study only the lowest energy is because the required estimates have already been established in [Bje07].

3 Outline of the paper

The model we consider has already been studied in [Bje07], and in order to avoid redoing a lot of work, we will simply summarize the main statements about the inductive construction used in that paper (see appendix A). We introduce the notation we use, as well as some basic assumptions and results, in section 4. There is a sketch of the proof, as well as a toy model to illustrate why we might expect the result to hold, in section 5.

In appendices B to D, we develop formulae, and collect some statements about growth estimates that we will later use together with the results in appendix A. These are all used in section 7 to prove that the list of assumptions, that are stated at the beginning of section 6, hold for our model.

From these assumptions, we prove theorem 2.1 in section 6. The assumptions have nothing to do with the linear structure of the system, and similar formulae can be developed for other systems. Therefore, the method should work for more general systems.
4 Notation, assumptions and basics

4.1 Diophantine irrationals

We recall that an irrational \( \omega \) is said to be Diophantine if

\[
\inf_{p \in \mathbb{Z}} |n\omega - p| > \frac{\kappa}{|n|^\tau}, \quad \text{for every } n \in \mathbb{Z}\setminus\{0\},
\]

(\( DC \))\( \kappa, \tau \)

where \( \kappa > 0 \) and \( \tau \geq 1 \). Diophantine irrationals are desirable in these types of problems because they have very good return properties.

**Lemma 4.1.** Let \( I \) be an interval in \( \mathbb{T} \) of length \( \varepsilon > 0 \). Then

\[ I \cap \bigcup_{0 < |n| \leq N} (I + n\omega) = \emptyset, \]

where \( N = \left\lfloor \left( \frac{\varepsilon}{\kappa} \right)^{1/\tau} \right\rfloor \) (\( \lfloor x \rfloor \) denotes the integer part of \( x \)).

That is, the first return time from an interval \( I \) to itself is always greater than some fixed constant times \( |I|^{-1/\tau} \). For a proof of this fact, see for instance [OT17, Lemma 3.1].

4.2 Basic notation

As is customary, we use the notation \((\theta_k, r_k) = \Phi_E(\theta_0, r_0)\). We also use \( \pi_1, \pi_2 \) to denote the projections onto the first and second coordinates, respectively:

\[
\begin{align*}
\pi_1(\theta_k, r_k) &= \theta_k, \\
\pi_2(\theta_k, r_k) &= r_k.
\end{align*}
\]

The skew-product structure ensures that points that start in the same fibre will always be in the same fibre. Therefore, given a \( \theta_0 \in \mathbb{T} \), and \( r_0, s_0, z_0 \) points in the same fibre, we refer to

\[(\theta_k, r_k), (\theta_k, s_k) \text{ and } (\theta_k, z_k),\]

simply as \( r_k, s_k \) and \( z_k \), respectively. The map \( \Phi_E \) induces the fibre-wise map

\[
r_{k+1} = \lambda^k v(\theta_k) - E - 1/r_k.
\]

(4.1)

We immediately get the relation

\[
r_{k+1} - s_{k+1} = \frac{r_k - s_k}{r_k s_k}.
\]

(4.2)

Since we consider only the invariant set \([\frac{1}{C}, C]\), where the invariant curves \( r_k^E \) and \( r_k^E \) are when \( E < E_0 \), then orientation is preserved: if \( s_0, r_0 \in B \), then \( s_0 \leq r_0 \) implies that \( s_1 \leq r_1 \). From now on, we will assume that \( s_0 \leq r_0 \), but let \( z_0 \) be an arbitrary point of reference, in no particular relation to either \( r_0 \) or \( s_0 \). Let us introduce the notation

\[
\begin{align*}
d(\theta_i) &= r_i - s_i, \\
D_{j,k}(r_0, s_0) &= \frac{1}{r_j s_j \cdots r_k s_k}, \quad \text{and} \\
\Pi_{j,k}(r_0, s_0) &= \prod_{i=j}^{k} \frac{r_i}{s_i}.
\end{align*}
\]
where $j \leq k$ are integers. If $j = k$, we will simply write $D_j(r_0, s_0)$ and $\Pi_j(r_0, s_0)$. Using induction, (4.2) gives us the relation

$$r_{k+1} - s_{k+1} = D_{j,k}(r_0, s_0) \cdot (r_j - s_j),$$

for every $j \leq k$, and so $D_{j,k}$ is simply the factor by which distance is changed between the $j$-th and the $(k+1)$-th step. We may relate these factors for different points:

$$D_{j,k}(r_0, z_0) = \prod_{i=j}^{k} \frac{s_i}{r_i} = D_{j,k}(r_0, s_0) \Pi_{j,k}(s_0, z_0).$$

(4.3)

Thus, $\Pi_{j,k}$ can be considered a sort of distortion factor for comparing distance growth between different points.

### 4.3 Assumptions and specific notation used in the construction

By shifting $E$ and $\theta$ linearly, we may assume that $v(\theta)$ has a unique non-degenerate global minimum equal to 0, at the point $\theta = 0$. Using Taylor expansion, we can see that if $\lambda > 0$ is sufficiently large, the set

$$\{\theta : v(\theta) \leq 10/\lambda\}$$

is contained in an interval of length $c_0/\sqrt{\lambda}$, centered at 0, for some constant $c_0$ depending only on $\nu$. Set

$$I_0 = \{\theta : |\theta| \leq c_0/(2\sqrt{\lambda})\},$$

and

$$M_0 = [\lambda^{1/(4\tau)}].$$

(4.4)

(4.5)

Then $I_0$ contains $\{\theta : v(\theta) \leq 10/\lambda\}$, which can be thought of as the interval where the system experiences rotation. In light of lemma 4.1, we see that the return time from $I_0$ to itself is bounded from below by the constant

$$N_0 = \left(\frac{\kappa \sqrt{\lambda}}{c_0}\right)^{1/\tau} \sim \lambda^{1/(2\tau)},$$

(4.6)

where $\kappa$ and $\tau$ are the constants appearing in the Diophantine condition $(DC)_{\kappa, \tau}$ and depend only on $\omega$. Therefore, $M_0 \sim \sqrt{N_0}$, if $\lambda$ is large enough. Later on, we will construct infinite sequences

$$I_0 \supset I_1 \supset \cdots,$$

$$M_0 < M_1 < \cdots,$$

and

$$N_0 < N_1 < \cdots.$$  

As above, for each $k > 0$, $N_k$ will be a lower bound for the return time from $I_k$ to itself, and $M_k \sim \sqrt{N_k}$ when $\lambda$ is very large. Now, we turn to the invariant sets for our fibres. Set

$$B = [\lambda^{-2}, \lambda^{2}],$$

$$B_\nu = [\lambda, \lambda^{2}],$$

and

$$B^\prime = [\lambda^{-2}, \lambda^{-1}].$$
The set $B$ will be invariant for the set of energies $E$ that we will consider. The system is contracting in the region $T \times B^u$ (the candidate for our first approximation of the unstable direction). Similarly, the system expands in $T \times B^s$ (the candidate for our stable direction).

The set of energies we consider is

$$\mathcal{E}_{-1} = [-1, 1].$$

It can be easily shown that the our cocycle is uniformly hyperbolic for $E \in (-\infty, -1)$. The interval $E_{-1}$ serves as our initial guess as to where $E_0$ is located, and in fact contains it. We will later on construct an infinite sequence of energy intervals

$$\mathcal{E}_{-1} \supset \mathcal{E}_0 \supset \cdots,$$

"zooming" in on the lowest energy $E_0$. If we write $\mathcal{E}_n = [E_n^-, E_n^+]$, then for every $n \geq 0$, we set

$$\mathcal{E}^{ld}_n = [E_n^-, E_{n+1}^-) \subset \mathcal{E}_n \backslash \mathcal{E}_{n+1}. \quad (4.7)$$

We will use the induction scheme in Appendix A to control the dynamics for the energies $E \in \mathcal{E}^{ld}_n$. We remark that $\bigcup_{n=-1}^{\infty} \mathcal{E}^{ld}_n = (-1, E_0)$. In particular, this means that the dynamics is uniformly hyperbolic for $E \in \mathcal{E}^{ld}_n$. The following result says that, $B^u$ ($B^s$) is forwards (backwards) invariant, as long as we stay sufficiently far away from the minimum of the potential $v$ (i.e. outside of the interval $I_0$). Therefore, any interesting effects on the dynamics will be a consequence of getting close to the minimum of $v$.

**Lemma 4.2.** Suppose that $E \in \mathcal{E}_{-1} = [-1, 1]$, and that $z_0 \in B$. Then

- $z_0 \notin B^s$ and $\theta_0 \notin I_0$, imply that $z_1 \in B^u$, and
- $z_0 \notin B^u$ and $\theta_0 \notin I_0 + \omega$, imply that $z_{-1} \in B^s$.

As the energy gets closer to $E_0$, the appropriate scale $n$ will increase, since we will need more time to recover from the worse growth estimates (where the loss of uniformity happens), therefore requiring longer return times before tackling the "bad returns" to $I_n$. Moreover, we set

$$\mathcal{Z}^u_n = \bigcup_{i=0}^{n} \bigcup_{m=1}^{M_i} (I_i + m\omega), \quad (4.8)$$

$$\mathcal{Z}^s_n = \bigcup_{i=0}^{n} \bigcup_{m=0}^{M_i} (I_i - m\omega), \quad (4.9)$$

$$\Theta_n = T \backslash (\mathcal{Z}^u_n \cup \mathcal{Z}^s_n). \quad (4.10)$$

The sets $\mathcal{Z}^u_n$ and $\mathcal{Z}^s_n$ should be thought of as the "immediate vicinity" of $I_0$, where at each scale the immediate vicinity is considered greater in terms of iterates. These sets are where we "lose information" about the invariant directions, and $\Theta_n$ is where we have almost perfect information about them, at scale $n$. Note that, since each $M_i \sim \sqrt{N}$, the vast majority of iterates spend time in $\Theta_n$. This is the basis of the construction.

In order to locate the invariant directions, we have to make an initial guess. They will be, for the two respective directions, the boxes

$$B^u_n = (I_n - M_n\omega) \times B^u = \{ (\theta, r) | \theta \in I_n - M_n\omega, r \in B^u \},$$

$$B^s_n = (I_n + M_n\omega) \times B^s = \{ (\theta, r) | \theta \in I_n + M_n\omega, r \in B^s \}. \quad (4.11)$$
Iterating these boxes will help us construct the invariant curves. To do so, we wish to look at the intersection of the forward iterates of the first box

\[ A_n^+ = \Phi^{M_n+1}(B_{n}^+) = \{ (\theta, r) | \theta \in I_n + \omega, \phi_{n}^{+} (\theta, E) \leq r \leq \phi_{n}^{+} (\theta, E) \}, \]

with the backward iterates of the second one

\[ A_n^- = \Phi^{-M_n+1}(B_{n}^-) = \{ (\theta, r) | \theta \in I_n + \omega, \phi_{n}^{-} (\theta, E) \leq r \leq \phi_{n}^{-} (\theta, E) \}. \]

If they don’t intersect, scale \( I_n \) will be sufficient to establish uniform estimates from these initial guesses. In fact, if \( E \in [-1, E_0) \), where \( E_0 \) is the lowest energy of the spectrum, then there will be an \( n \) such that \( A_n^+ \) and \( A_n^- \) do not intersect, and in fact \( \phi_{n}^{+} < \phi_{n}^{-} \).

5 Ideas and sketch of the proof

5.1 Illustrating the idea behind norm growth through a model example

Suppose that we have two functions \( \phi_0, \psi_0 : [-a, a] \to \mathbb{R} \) that are quadratically separated:

\[ \psi_0(\theta) - \phi_0(\theta) = d + \theta^2, \]

where \( d > 0 \) is some constant. Suppose that we generate a new functions \( \psi_1 \) over \([-a, a] \), given by

\[ \psi_1(\theta) - \phi_1(\theta) = r(\psi_0(\theta) - \phi_0(\theta)) = rd + r\theta^2, \]

for some \( r > 1 \). That is, \( \psi_1 \) is obtained simply by separating \( \psi_0 \) and \( \phi_0 \) by a factor \( r \). We similarly obtain functions \( \psi_2, \ldots, \psi_n \) by

\[ \psi_k(\theta) - \phi_k(\theta) = r(\psi_{k-1}(\theta) - \phi_{k-1}(\theta)) = r^k(\psi_0(\theta) - \phi_0(\theta)). \] (5.1)

For any given \( \theta \in [-a, a] \), we record the first \( \sigma = \sigma(\theta) \geq 0 \) such that

\[ \psi_\sigma(\theta) - \phi_0(\theta) \geq \delta > 0. \]

We say that \( \delta \) is the distance at which the curves (the graphs of the functions) become separated/decorrelated. We immediately see that

\[ \frac{\delta}{d + \theta^2} \leq r^\sigma \leq \frac{r\delta}{d + \theta^2}. \]

When we differentiate the relation (5.1) with respect to \( \theta \), we obtain

\[ \frac{\partial}{\partial \theta}(\psi_k(\theta) - \phi_0(\theta)) = r^k \frac{\partial}{\partial \theta}(\psi_0(\theta) - \phi_0(\theta)) = 2r^k \theta. \]

Therefore, when \( k = \sigma \), we have

\[ 2\delta \frac{\theta}{d + \theta^2} \leq \frac{\partial}{\partial \theta}(\psi_\sigma(\theta) - \phi_0(\theta)) \leq 2r\delta \frac{\theta}{d + \theta^2}, \]

which has a maximum when \( \theta = \sqrt{d} \). Therefore, if \( a \geq \sqrt{d} \), the maximum is realized, and we would have

\[ \frac{\delta}{\sqrt{d}} \leq \max_{\theta \in [-a, a]} \max_{0 \leq k \leq \sigma(\theta)} \frac{\partial}{\partial \theta}(\psi_k(\theta) - \phi_0(\theta)) \leq r\delta \frac{1}{\sqrt{d}}. \]

This model example captures the essential ideas of the construction. It provides a toy model of the local behaviour of the invariant functions in the present model (the one considered in this paper). The main difficulties in our present model are:
• A lack of uniform growth estimates. Namely, the factor \( r \) depends on \( \theta \), and a typical orbit will spend cycles between periods of expansion, and periods of contraction, before becoming separated/decorrelated.

• The initial graph is not perfectly quadratic, but close to one. Moreover, it is not obvious how large the interval is, where it satisfies some given quadratic condition. That is, it is not obvious that we can choose \( a \geq \sqrt{d} \).

The first point may lead to a loss of uniform constants in the above inequality. In fact, this is something we should expect for a small set of exceptional energies. The second point is crucial in obtaining anything close to the exponent \( \frac{1}{2} \). However, it turns out that such intervals are even much longer than what is needed, but it does remains an important part of the proof.

5.2 Sketch of the construction and proof

![Figure 2](image-url)

(a) Zoom-in on the initial peak, which is approximately quadratic.

(b) Zoom-in (same scale as in the left figure) on the peak after a few iterations. Notice that a lot of points have escaped \( B^u \), and the peak is much sharper.

Figure 2: For the simulations we used the almost-Mathieu potential \( \cos(2\pi \theta) \) with \( \omega = (\sqrt{5} - 1)/4 \) and \( \lambda^2 = 30 \).

In \( B^u \) (blue in fig. 2) the system is strongly contracting (by a factor \( \leq \lambda^{-1} \)), and in \( B^s \) (red in fig. 2) the system is strongly expading (by a factor \( \geq \lambda \)). Since we are looking at the projective dynamics, the unstable region is indeed contracting, and the stable one expanding. Recall that, for \( E < E_0 \), we have a stable direction \( \psi^s_E : \mathbb{T} \to B \), and an unstable direction \( \psi^u_E : \mathbb{T} \to B \).

The interval \( I_0 \) is the only place where non-negligible rotation takes place. It is therefore the only place where an invariant direction can change from expanding to expanding. That is, if we have an invariant function \( \psi \), then \( \psi(\theta) \in B^u \) implies \( \psi(\theta + \omega) \in B^u \), unless \( \theta \in I_0 \). Similarly, \( \psi(\theta) \in B^s \) implies \( \psi(\theta - \omega) \in B^s \), unless \( \theta \in I_0 + \omega \). This is what lemma 4.2 says.

Note that the rotation taking place in \( I_0 \) is reflected in the invariant directions over \( I_0 + \omega \), since

\[
\psi(\theta + \omega) = \Phi(\theta, \psi(\theta)).
\]

That is, the direction is lagging one step behind whatever the transformation is doing.
When the system is uniformly hyperbolic, the curve $\psi_E^+ \sigma$ will spend most of its time in $B^u$, and $\psi_E^-$ will spend most of the time in $B^s$. As $E$ gets closer to $E_0$, the curves will approach one another, and the curves will spend progressively less time in their respective regions.

Specifically, as the directions get closer to each other, whatever expansion/contraction one direction experiences, the other one does too. This causes a complicated cycling between expansion and contraction, in our case ultimately leading to non-uniform hyperbolicity.

Using an induction procedure, we identify an interval $I = I(E) \subset I_0 + \omega$, where the minimum of their difference is minimised. It can then be shown that the distance is asymptotically linear in $E$. That is, if we let $\delta(E)$ denote the minimum distance between the curves for the parameter $E$, we have

$$\partial_E \delta(E) = \text{const} \cdot (E_0 - E) + o(E_0 - E),$$

for some positive constant independent of $E$. Moreover, the difference between the curves has an approximately quadratic shape over $I$, that is

$$d(\theta) = \psi_E^+(\theta) - \psi_E^-(\theta) \sim \delta(E) + (\theta - \theta_c)^2$$

for some $\theta_c = \theta_c(E) \in I$ (where the minimum is attained), and every $\theta \in I$.

For every $\theta \in I$, we define stopping times $\sigma^\pm = \sigma^\pm(\theta, E)$, that measure how long the two directions stay close going forwards. Specifically, $\sigma^+$ is the time at which they become separated, going forwards, and $\sigma^-$ is defined similarly, but going backwards. We show that the second derivative has the biggest blow up in the set of $\theta$’s between such stopping times. In fig. 2, we see how the difference between the curves becomes sharper as we iterate the interval $I$ forwards.

How do we show the bounds for the $\mathcal{C}^1$-norms of the curves? The crucial step is relating the growth of the distance to the growth of the derivatives. Indeed, if the difference between the curves is very close to 0, then the expansion the two curves will experience should be very similar.

In the next section, we show that the factor which determines the growth of their difference, is essentially the same as the one controlling the growth of the derivative, with the factors losing only an exponent $\epsilon$ between them. This tight coupling between the factors holds up to the stopping times defined above. Once the stopping time has been reached, the local information of one curve no longer gives any reliable local information about the other, and the procedure stops.

We will use the notation introduced in section 4.2. Heuristically, in order to obtain the stopping times $\sigma^+(\theta_0)$, that is the first time when the curves have been separated by the distance $\text{const}$ starting from $\theta_0$, we can solve the equation

$$d(\theta_{\sigma^+}) = d(\theta_0)D_{0,\sigma^+-1}(\psi_E^+(\theta_0), \psi_E^-(\theta_0)) \approx \text{const}.$$  

The distance factor can then be expressed as

$$D_{0,\sigma^+-1}(\psi_E^+(\theta), \psi_E^-(\theta)) \sim \frac{1}{d(\theta)}.$$  

The expression (C.4) gives us, as long as the remainder term is reasonably small, that

$$\partial_\theta d(\theta_{\sigma^+}) \approx d(\theta_{\sigma^+})\frac{\partial_\theta d(\theta_0)}{d(\theta_0)}\Pi_{0,\sigma^+-1}(\psi_E^+(\theta_0), \psi_E^+(\theta_0)),$$

where $\Pi_{0,\sigma^+-1}$ is some (small) distortion factor. As long as $|I| \gtrsim \sqrt{\delta(E)}$, we may choose a $\theta_0$ that makes $d(\theta_0) \sim \delta(E)$, and $\partial_\theta d(\theta_0) \sim \sqrt{\delta(E)}$. In that case, we see that

$$\frac{\partial_\theta d(\theta_0)}{d(\theta_0)} \sim \frac{1}{\sqrt{\delta(E)}},$$
which in turn shows that
\[ \partial_\theta d(\theta_{\sigma^+}) \sim \frac{1}{\sqrt{d(E)}}. \]

Since \( k \) was chosen such that \( \Delta(\theta_k) \sim 1 \), we find that the only obstacle remaining is controlling the distortion factor \( \Pi_{0,k-1} \). It turns out that it is close to 1, except for some exceptional energies, which causes the loss of exponent in the lower bound of the norm.

In fact, these exceptional energies are precisely the ones where the stopping times align with the cycling of expanding/contracting behaviour. That is, the stopping time occurs shortly before the next return to bad points in \( I_0 \), namely the sets \( I_n \), where a larger \( n \) means the set is worse.

6 Proof of Main Theorem

In order to split up the proof into smaller parts, we will show how the conclusions in theorem 2.1 follow from a list of assumptions. In the next section, we will prove that all of those assumptions hold for Schrödinger cocycles satisfying the assumptions of theorem 2.1.

6.1 List of assumptions

Here is a list of the assumptions we will base the proof on.

(A1) For every \( E \in [-1,E_0) \) there are two \( C^2 \) invariant functions (curves) \( \psi_E^s < \psi_E^u : \mathbb{T} \rightarrow B \).

(A2) There are strictly positive constants \( C_0 \) and \( C_1 \), independent of \( E \), and an increasing function \( l(E) \nearrow \infty \) as \( E \nearrow E_0 \), such that for every \( E \in [-1,E_0) \), there is an interval \( I = I(E) \) satisfying:

(a) The minimum distance between the curves is linearly asymptotic
\[ \delta(E) = \min_{\theta \in \mathbb{T}} (\psi_E^u - \psi_E^s)(\theta) = \min_{\theta \in I(E)} (\psi_E^u - \psi_E^s)(\theta) = C_0 \cdot (E_0 - E) + o(E_0 - E) \tag{6.1} \]
as \( E \nearrow E_0 \).

(b) There is some \( \theta_c = \theta_c(E) \in I \), such that
\[ \delta(E) + \frac{1}{C_1} \cdot (\theta - \theta_c)^2 \leq (\psi_E^u - \psi_E^s)(\theta) \leq \delta(E) + C_1 \cdot (\theta - \theta_c)^2, \tag{6.2} \]
\[ |\partial_\theta \psi_E^u(\theta)| < C_1, \text{ and } |\partial_\theta \psi_E^s(\theta)| \leq C_1 \tag{6.3} \]
for every \( \theta \in I \).

(c) The length of the interval satisfies the lower bound
\[ |I| \geq l(E) \cdot \sqrt{\delta(E)}. \tag{6.4} \]

(A3) For every \( E \in [-1,E_0) \) and \( \theta \in I \) there are stopping times \( \sigma \pm = \sigma(\theta,E) \geq \hat{\sigma} \pm = \hat{\sigma}(\theta,E) > 0 \), and a positive function \( \eta(E) \searrow 0 \) (as \( E \nearrow E_0 \)), satisfying
\[ \sigma^+ - \hat{\sigma}^+ \leq \eta(E)\sigma_E^+, \]
\[ \sigma^- - \hat{\sigma}^- \leq \eta(E)\sigma_E^-, \text{ and:} \]
(a) If we set $\sigma^\pm_E = \max_{\theta \in I} \sigma^\pm(\theta, E)$, then
\[
\left( \bigcup_{m=1}^{15\sigma^+_E} I + m\omega \right) \cap \left( \bigcup_{m=1}^{15\sigma^-_E} I - m\omega \right) = \emptyset.
\] (6.5)

(b) For every $\theta \in I$, and every $0 \leq k \leq \sigma^+(\theta)$,
\[
D_{0,k}(\psi^E_+ (\theta), \psi^E_+ (\theta)) \geq \lambda^{(k+1)/2}.
\] (6.6)

For every $\theta \in I$, and every $0 \leq j \leq \hat{\sigma}^+$,
\[
D_{j,\hat{\sigma}^+}(\psi^E_+ (\theta), \psi^E_+ (\theta)) \geq \lambda^{(\hat{\sigma}^+-j+1)/2}.
\] (6.7)

For every $\theta \in I$, $0 \leq j \leq \sigma^+$ and every $j + \eta(E)\sigma^+_E \leq k \leq N^+(\theta)$, where $N^+(\theta) > 0$ is the smallest integer such that $\theta + N^+\omega \in I$, we have
\[
D_{j,k}(\psi^E_+ (\theta), \psi^E_+ (\theta)) \geq \lambda^{(k-j+1)/2}.
\] (6.8)

(c) For every $0 \leq k \leq \sigma^-$, and every $\theta \in I$,
\[
D_{-k,0}(\psi^E_+ (\theta), \psi^E_+ (\theta)) \leq \lambda^{-(k+1)/2}.
\] (6.9)

For every $\theta \in I$, and every $0 \leq j \leq \hat{\sigma}^-$,
\[
D_{-\hat{\sigma}^-,j}(\psi^E_+ (\theta), \psi^E_+ (\theta)) \geq \lambda^{-(-\hat{\sigma}^-+j+1)/2}.
\] (6.10)

For every $\theta \in I$, $0 \leq j \leq \sigma^-$ and every $j + \eta(E)\sigma^-_E \leq k \leq N^-\theta$, where $N^-\theta > 0$ is the smallest integer such that $\theta - N^-\omega \in I$, we have
\[
D_{-k,j}(\psi^E_+ (\theta), \psi^E_+ (\theta)) \geq \lambda^{-(k-j+1)/2}.
\] (6.11)

We will now briefly discuss each of the assumptions, and how they can be interpreted.

The first one, (A1), is saying that we have two distinct invariant families of directions, the directions of the most expansion ($\psi^E_+ \theta$), and the most contraction ($\psi^E_-$). The estimates in (6.8) and (6.11) give bounds of their respective expansion/contraction.

In the next one, (A2), the interval $I$ will be the interval where these directions are the closest to each other. In the present model, the smallest distance is asymptotically linear, and behaves quadratically at the interval $I$. The quadratic condition ensures that the directions are not too close, too frequently. This is a consequence of the minimum being non-degenerate, and holds generally for the type of model we consider in this paper.

The assumptions in (A3) are essentially growth estimates for Lyapunov exponents, and help us measure how much uniformity is lost at each parameter. Essentially, as long as they are close to each other, they separate exponentially fast (both forwards and backwards).

The stopping times $\sigma^\pm = \sigma^\pm(\theta, E)$, where $\theta \in I$, are simply the largest times such that
\[
(\psi^E_+ - \psi^E_\theta)(\theta + k\omega) < \lambda^{-3},
\]
for every $-\sigma^- \leq k \leq \sigma^+$. They are defined in section 7.2, together with the intervals $I(E)$. 

13
6.2 Proof of the main result

Note that the assumption in (6.1) is in fact the statement (2.2) in theorem 2.1, which follows from proposition 7.7. Therefore, we only need to focus on the statement (2.3), which follows from proposition 6.1. In the next section, we will prove that the above assumptions hold for our model, and so the main results indeed follows if we can prove it from our list of assumptions.

In what follows, let both \( E \in (E_{-1}, E_0) \) and \( \theta \in I = I(E) \) be fixed. In order to ease notation, we set
\[
\begin{align*}
    s_0 &= \psi_E^s(\theta_0), \\
    r_0 &= \psi_E^r(\theta_0).
\end{align*}
\]

(6.12)

Because of (.41), the dynamics is always confined to \( B = [\lambda^{-2}, \lambda^2] \), and everything in appendix C will hold for the sequences \( r_k \) and \( s_k \). Recall that \( \sigma^+ = \sigma^+(\theta_0) \) depends on \( \theta_0 \), and consider the relation in (C.4):
\[
\partial_\theta (r_{\sigma^+ + 1} - s_{\sigma^+ + 1}) = (r_{\sigma^+ + 1} - s_{\sigma^+ + 1}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} \Pi_{0, \sigma^+} (s_0, r_0) + R_{0, \sigma^+} (r_0, s_0).
\]

We will split the proof into three parts. The first part deals with the first term
\[
(r_{\sigma^+ + 1} - s_{\sigma^+ + 1}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} \Pi_{0, \sigma^+} (s_0, r_0).
\]

This term is, as we shall see, the dominant term. The second part deals with the remainder term \( R_{0, \sigma^+} (r_0, s_0) \), which will be shown to be negligible in comparison to the first one. The last part of the proof deals with showing that the maximum of the norm of \( \psi^E_E \) is essentially attained at \( \theta_{\sigma^+ + 1} \) for some appropriate initial point \( \theta_0 \in I \).

6.2.1 Treating the dominant term

Consider the term
\[
(r_{\sigma^+ + 1} - s_{\sigma^+ + 1}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} \Pi_{0, \sigma^+} (s_0, r_0).
\]

The last factor \( \Pi_{0, \sigma^+} (s_0, r_0) \) can be dealt with through the inequality in (C.6), leading us to investigate the sum
\[
\sum_{j=1}^{\sigma^+} \frac{1}{D_{j, \sigma^+} (r_0, s_0)}.
\]

That is, an upper bound for that sum leads to a lower bound for the factor \( \Pi_{0, \sigma^+} (s_0, r_0) \). The problem here is that \( D_{j, \sigma^+} (r_0, s_0) \) may behave badly (not uniformly exponentially) for \( j \) close to \( \sigma^+ \). Therefore, we split the sum into
\[
\sum_{j=0}^{\sigma^+} \frac{1}{D_{j, \sigma^+} (r_0, s_0)} = \sum_{j=0}^{\tau} \frac{1}{D_{j, \sigma^+} (r_0, s_0)} + \sum_{j=\tau+1}^{\sigma^+} \frac{1}{D_{j, \sigma^+} (r_0, s_0)},
\]

where \( \tau = \hat{\sigma}^+ \), and therefore \( |\sigma^+ - \tau| \leq \eta(E)\hat{\sigma}^+ \). For \( 0 \leq j \leq \tau \), we have the inequality
\[
D_{j, \tau} (r_0, s_0) \geq \lambda^{(\tau - j + 1)/2},
\]

14
by (6.7). Since $D_{j, \sigma^+} \geq 1$ for every $0 \leq j \leq \sigma^+$ (otherwise the distance at step $j$ would be greater than at step $\sigma^+$, contradicting the definition of the stopping time), we may estimate

$$\sum_{j=\tau+1}^{\sigma^+} \frac{1}{D_{j, \sigma^+}(r_0, s_0)} \leq \sigma^+ - \tau = \eta(E)\sigma^+_E.$$  

Combining these estimates, we end up with the upper bound

$$\sum_{j=0}^{\sigma^+} \frac{1}{D_{j, \sigma^+}(r_0, s_0)} \leq \frac{1}{1 - \lambda^{-1/2} + \eta(E)\sigma^+_E}.$$  

The inequality in (C.6) immediately implies that

$$\exp\left(-\frac{\lambda^4}{1 - \lambda^{-1/2}} - \lambda^4\eta(E)\sigma^+_E\right) \leq \Pi_{0, \sigma^+}(s_0, r_0) \leq 1.$$  

Unfortunately, the term $\delta(E)\sigma^+$ prevents a uniform lower bound. However, we may still estimate how much we lose. Since the dynamics takes place in $B = \left[\lambda^{-2}, \lambda^2\right]$, we have the upper bound

$$r_{\sigma^++1} - s_{\sigma^++1} \leq \lambda^2.$$  

Therefore, (6.6) gives us the inequality

$$\lambda^{(\sigma^++1)/2} \leq D_{0, \sigma^+}(r_0, s_0) \leq \frac{\lambda^2}{r_0 - s_0}.$$  

This means that

$$\sigma^+ \leq 3 + 2 \log_\lambda \frac{1}{r_0 - s_0}, \quad \text{(6.13)}$$  

and in particular that $\sigma^+_E \lesssim \log_\lambda \frac{1}{\delta(E)}$. Therefore

$$\exp(-\lambda^4\eta(E)\sigma^+_E) \geq \text{const} \cdot \delta(E)^{2\lambda^4\eta(E)},$$  

where the constant is independent of $E$, and uniformly bounded away from 0. Since $\eta(E) \searrow 0$ as $E \nearrow E_0$, there is a positive constant const and a positive $\varepsilon = \varepsilon(E)$ that goes to 0 as $E \nearrow E_0$ such that

$$\text{const} \cdot \delta(E)^\varepsilon \leq \Pi_{0, \sigma^+}(s_0, r_0) \leq 1.$$  

We now turn our attention to the factor

$$\partial_\theta(r_0 - s_0) \quad \text{in (6.14).}$$  

Recall that we set $r_0 = \psi^+_E(\theta)$ and $s_0 = \psi^+_E(\theta)$ in (6.12). Using the inequalities in (6.2), we have

$$r_0 - s_0 = \delta(E) + C_1(\theta)(\theta - \theta_0)^2,$$

and

$$\partial_\theta(r_0 - s_0) = \tilde{C}_1(\theta)(\theta - \theta_0).$$  

15
where $\frac{1}{C_1} \leq C_1(\theta) \leq C_1$ and $\frac{2}{C_1} \leq \tilde{C}_1(\theta) \leq 2C_1$. Now, there is a unique $\beta > 0$ such that

$$C_1(\theta)(\theta - \theta_0)^2 = \delta(E)^{2\beta}.$$ 

This means that

$$\tilde{C}_1(\theta)(\theta - \theta_0) = \tilde{C}_1(\theta)\delta(E)^{\beta},$$

where $2C_1^{-1/2} \leq \tilde{C}_1(\theta) \leq 2C_1^{3/2}$. We end up with the new expressions

$$r_0 - s_0 = \delta(E) + \delta(E)^{2\beta},$$

and

$$\partial_\theta (r_0 - s_0) = \tilde{C}_1 \cdot \delta(E)^{\beta}.$$ 

Plugging these expressions into (6.14), we end up with

$$\frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} = \tilde{C}_1(\theta) \cdot \frac{1}{\delta(E)^{1-\beta} + \delta(E)^{\beta}},$$

which attains its maximum at $\beta = \frac{1}{2}$. If we can show that some $\theta$ satisfies that $\beta = \frac{1}{2}$, this maximum is indeed attained. Since $|l(E)| \geq l(E) \cdot \sqrt{\delta(E)}$, where $L(E) \not\to \infty$ as $E \not\to E_0$, by the assumption in (6.4), it is clear that some $\theta$ has $\beta = \frac{1}{2}$, provided that $E$ is sufficiently close to $E_0$.

Again, since the dynamics is constrained to $B = [\lambda^{-2}, \lambda^2]$, we have the trivial bound $0 \leq r_{\sigma^+} - s_{\sigma^+} \leq \lambda^2$. All this together gives us the inequalities

$$\text{const} \cdot \frac{1}{\sqrt{\delta(E)}} \delta(E)^{\varepsilon} \leq \max_{\theta_0 \in I} (r_{\sigma^+} - s_{\sigma^+}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} \Pi_{0,\sigma^+}(s_0, r_0) \leq \text{const} \cdot \frac{1}{\sqrt{\delta(E)}},$$

where $\varepsilon = \varepsilon(E)$ is positive and $\lim_{E \not\to E_0} \varepsilon = 0$. Since $\Pi_{0,k} \leq 1$, we always have the upper bound

$$(r_{k+1} - s_{k+1}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} \Pi_{0,k}(s_0, r_0) \leq (r_{k+1} - s_{k+1}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0}.$$

By definition of $\sigma^+$, it is also the case that for every $k \leq \sigma^+$ we have the inequality $r_{k+1} - s_{k+1} \leq r_{\sigma^+ - 1} - s_{\sigma^+ + 1}$. Therefore, we immediately get the bounds

$$\text{const} \cdot \frac{1}{\sqrt{\delta(E)}} d(E)^{\varepsilon} \leq \max_{\theta_0 \in I, 0 \leq k \leq \sigma^+} (r_{k+1} - s_{k+1}) \frac{\partial_\theta (r_0 - s_0)}{r_0 - s_0} \Pi_{0,k}(s_0, r_0) \leq \text{const} \cdot \frac{1}{\sqrt{\delta(E)}},$$

where $\lim_{E \not\to E_0} \varepsilon = 0$.

### 6.3 Treating the remainder term

By (C.5), we have

$$|R_{0,\sigma^+}(r_0, s_0)| \leq 2\lambda^3 \cdot \sigma^+ \cdot \max_{0 \leq j \leq \sigma^+} |\partial_\theta (s_j)|.$$ 

As we saw in (6.13),

$$\sigma^+ \leq 3 + 2 \log_\lambda \frac{1}{r_0 - s_0} \leq 3 + 2 \log_\lambda \frac{1}{\delta(E)}.$$
Therefore, there is an \( \varepsilon = \varepsilon(E) \lesssim 0 \) (as \( E \not\nearrow E_0 \)), different from the previous \( \varepsilon \), such that

\[
\sigma^+ \leq \delta(E)^{-\varepsilon}.
\]

The factors \( \max_{0 \leq j \leq \sigma^+} |\partial_\theta(s_j)| \) can be dealt with by considering the expression in (D.4).

\[
\partial_\theta s_{-k} = (\partial_\theta s_0)s_k^2 \cdots s_{k-1} - \lambda^2 \sum_{j=1}^{k} v'(\theta_{j}) s_{j}^2 \cdots s_{k-1} =
\]

\[
\frac{\partial_\theta s_0}{D_{-k-1}(s_0,s_0)} - \lambda^2 \sum_{j=1}^{k+1} \frac{v'(\theta_{j})}{D_{-k-j}(s_0,s_0)}.
\]

Let \( \theta_0 \in I \). Since we wish to estimate \( |\partial_\theta s_j| \) for \( 0 \leq j \leq \sigma^+ \), we consider any \( k > 0 \) satisfying that \( \theta_{-k} \in \bigcup_{m=0}^{\sigma^+} I + m\omega \). Since the iterates of \( I \) cover the circle, every \( \theta \) in the union is in fact \( \theta_{-k} \) for some \( \theta_0 \in I \) and some \( k > 0 \). That is, every \( s_j \) we consider is simply the backward iterate of some \( \tilde{s}_0 \in I \). Then (6.5) gives us that \( k \geq 14 \sigma^+_E \gg \eta(E)\sigma^+_E \). Therefore (6.8) applies, and we obtain

\[
D_{-k-1}(s_0,s_0) \geq \lambda^{k/2}.
\]

As before, we divide the sum into two parts, one behaving like a geometric sum (when \( k - j \geq \eta(E)\sigma^+_E \)), and another when \( k - j < \eta(E)\sigma^+_E \). The part behaving like a geometric sum gives a contribution that is uniformly bounded. The interesting part is therefore \( k - j < \eta(E)\sigma^+ \), and it can be bounded using the trivial estimate

\[
D_{-k-j}(s_0,s_0) \geq \lambda^{-4\eta(E)\sigma^+}.
\]

Since \( \sigma^+ \sim \log \frac{1}{\delta(E)} \), and \( \eta \searrow 0 \), there is an \( \varepsilon \searrow 0 \) (as \( E \not\nearrow E_0 \)) such that

\[
\left| \sum_{j=k-\eta(E)\sigma^+-1}^{k+1} \frac{v'(\theta_{j})}{D_{-k-j}(s_0,s_0)} \right| = \eta(E)\sigma^+ \lambda^{-4\eta(E)\sigma^+} \leq \delta(E)^{-\varepsilon},
\]

and therefore the whole sum behaves like

\[
\sum_{j=1}^{k+1} \frac{v'(\theta_{j})}{D_{-k-j}(s_0,s_0)} \leq \text{const} + \delta(E)^{-\varepsilon} \sim \delta(E)^{-\varepsilon}.
\]

Since \( \partial_\theta s_0 \) is uniformly bounded on \( I \), w.r.t. \( E \) (see (6.3)), it follows that \( \partial_\theta s_{-k} \) satisfies the bound

\[
|\partial_\theta s_{-k}| \leq \delta(E)^{-\varepsilon},
\]

where \( \varepsilon \) is positive, distinct from the other \( \varepsilon \) above, and \( \varepsilon \searrow 0 \) as \( E \not\nearrow E_0 \). Therefore, we have

\[
\max |\partial_\theta \psi' (\theta)| \leq \delta(E)^{-\varepsilon}, \quad (6.16)
\]

where the maximum is taken over the set \( \theta \in \bigcup_{m=1}^{\sigma^+} I + m\omega \). That is,

\[
|R_{0,\sigma^+}(r_0,s_0)| \leq \delta(E)^{-\varepsilon}.
\]
6.4 Locating the global maximum

Putting everything together in the previous subsections, we obtain the inequality

\[
\text{const} \cdot \left( \frac{1}{\sqrt{\delta(E)}} \right)^{1-\epsilon} \leq \max_{\theta_0 \in \mathcal{I}, 0 \leq k \leq \sigma^+ + 1} \partial_\theta(r_k - s_k) \leq \text{const} \cdot \frac{1}{\sqrt{\delta(E)}},
\]

where the constant is uniformly bounded away from 0, and \( \epsilon \searrow 0 \) as \( E \nearrow E_0 \). By the estimate in (6.16), it follows that \( \partial_\theta r_k \) is the dominant term in the maximum, and therefore

\[
\text{const} \cdot \left( \frac{1}{\sqrt{\delta(E)}} \right)^{1-\epsilon} \leq \max_{\theta_0 \in \mathcal{I}, 0 \leq k \leq \sigma^+ + 1} \partial_\theta r_k \leq \text{const} \cdot \frac{1}{\sqrt{\delta(E)}}.
\]

By a simple argument, we will show that this is in fact (essentially) the maximum. Since \( r_k^2 \leq r_k s_k = D_k(r_0, s_0) \), we have for every \( k \geq \sigma^+ + 1 \) that

\[
|\partial_\theta r_{k+1}| \leq |\partial_\theta r_{\sigma^+ + 1}| D_{\sigma^+ + 1, k}(r_0, s_0) + \lambda^2 |v'(\theta_k)| + \lambda^2 \sum_{j=\sigma^+ + 1}^{k-1} |v'(\theta_j)| D_{j+1, k}(r_0, s_0).
\]

Since \( r_{\sigma^+ + 1} - s_{\sigma^+ + 1} \geq \lambda^{-3} \), and we always have \( r_k - s_k \leq \lambda^2 \), it follows that \( D_{\sigma^+ + 1, k}(r_0, s_0) \leq \lambda^5 \). That is, for every \( k \geq \sigma^+ + 1 \), we have

\[
|\partial_\theta r_{k+1}| \leq \text{const} + \lambda^5 |\partial_\theta r_{\sigma^+ + 1}|.
\]

Since the forward iterates of \( I \) cover the circle, this means that we get the following result.

**Proposition 6.1.** There is a positive \( \epsilon = \epsilon(E) \) satisfying \( \lim_{E \nearrow E_0} \epsilon = 0 \), and a \( C_2 > 0 \) independent of \( E \), such that

\[
\frac{1}{C_2} \cdot \left( \frac{1}{\sqrt{\delta(E)}} \right)^{1-\epsilon} \leq \max |\partial_\theta \psi_{E}^{\pm}| \leq C_2 \cdot \frac{1}{\sqrt{\delta(E)}}.
\]

Using the exact same arguments, but iterating the other direction, we can also prove that

\[
\frac{1}{C_2} \cdot \left( \frac{1}{\sqrt{\delta(E)}} \right)^{1-\epsilon} \leq \max |\partial_\theta \psi_{E}^{-}| \leq C_2 \cdot \frac{1}{\sqrt{\delta(E)}},
\]

where the constant is uniformly bounded away from 0, and \( \epsilon \searrow 0 \) as \( E \nearrow E_0 \).

This concludes the proof of the second part of theorem 2.1.

7 Proof of assumptions

In this section we will derive (A1) to (A3) from lemma A.1. We will use all the notation from that section. Most of the statements in this section assume that \( \lambda \) is large enough. In this section, we will therefore assume that \( \lambda \) is large enough (depending only on \( \omega \) and the potential \( \nu \)) for lemma A.1, and all the statements contained within this section, to hold. We observe the following:

\[
M_0 \rightarrow \infty \quad \text{and} \quad \frac{M_0}{N_0} = o(1),
\]

18
as $\lambda \to \infty$ (see (4.6), and the line after it). The sequence $M_i$ therefore grows super-exponentially fast if $\lambda$ is large, since

$$M_i \approx \lambda^{M_{i-1}/(4\tau)}$$

(7.1)

for every $i \geq 1$. Moreover, the return bounds in lemma 4.1 imply that

$$M_i \approx \sqrt{N_i}$$

(7.2)

for every $i \geq 0$. For every $n \geq -1$, set

$$E_U_n = [E^{-n}, E^{-n} + 1) \subset E_n \setminus E_{n+1},$$

(7.3)

where we use the notation $E_n = [E^{-n}, E^n]$. It is worth noting that $\bigcup_{n \geq -1} E_U_n = [-1, E_0)$, where $E_0$ is the lowest energy of the spectrum. That is, given an $E \in [-1, E_0)$, there is a fixed $n \geq -1$ such that $E \in E_U_n$.

We now state a stronger condition that will be satisfied in these energy intervals. The condition is essentially an extension of $(C_1)_n$ in appendix A to iterates past $I_n$.

**Condition $\mathcal{UH}_n$**

Condition $(C1)_m$ and $(C2)_m$ for every $m \leq n$, together with the following conditions:

1. Suppose that $(\theta_0, r_0) \in \Theta_n \times B^u$, then for every integer $k$

$$r_k \in B,$$

$$r_k \notin B^u \implies \theta_k \in \mathbb{Z}_n^u = \bigcup_{i=0}^{n} \bigcup_{m=1}^{M_i} (I_i + m\omega).$$

(7.4)

2. Suppose that $(\theta_0, r_0) \in \Theta_n \times B^s$, then for every integer $k$

$$r_{-k} \in B,$$

$$r_{-k} \notin B^s \implies \theta_{-k} \in \mathbb{Z}_n^s = \bigcup_{i=0}^{n} \bigcup_{m=0}^{M_i} (I_i - m\omega).$$

(7.5)

Later on, we shall show that this condition is satisfied for every $E \in \mathcal{E}_{n-1}^U$, and $n \geq 0$.

### 7.1 Proving $(A1)$

Here we prove that, for every $n \geq 1$, Condition $\mathcal{UH}_n$ is satisfied for every $E \in \mathcal{E}_{n-1}^U$. We will show how this implies the existence of two invariant functions (curves) $\psi^u, \psi^s : \mathbb{T} \to B$, for every $E \in \mathcal{E} = \bigcup_{n=0}^{\infty} \mathcal{E}_n^U$.

The following result is crucial to the whole construction. It allows us to analyse the dynamics for all times, and establish uniform hyperbolicity. This result is implicit in the construction used in lemma A.1, but not explicitly stated in that paper.

**Lemma 7.1.** Suppose that $n \geq 1$, and that $E \in \mathcal{E}_{n-1}^U$. Then Condition $\mathcal{UH}_n$ is satisfied, and

$$A_{n-1}^u \cap A_n^s \neq \emptyset,$$

$$A_n^u \cap A_n^s = \emptyset,$$

$$A_{n-1}^u \cap A_n^s = \emptyset.$$
Proof. Since $E \in \mathcal{E}_{n-1}^{\leq} \subset \mathcal{E}_{n-1} \subset \cdots \subset \mathcal{E}_{-1}$, lemma A.1 implies the conditions (C1)$_m$ and (C2)$_m$ are satisfied for every $m \leq n$. Using the same methods as above (and below), one can show that

$$A_{m-1}^u \cap A_{m-1}^s \neq \emptyset, \text{ and}$$

$$A_m^u \cap A_m^s = \emptyset.$$ 

The sets $A_m^u$ and $A_m^s$ were constructed precisely to satisfy this, when $E \in \mathcal{E}_{n-1}^{\leq}$, but the statement of this fact is buried in the proof of [Bje07, Lemma 5.3]. Since the proof of this is technical, and would add nothing new, we have chosen to exclude it.

In order to check the rest of $(\mathcal{UH})_n$, suppose that $(\theta_0, r_0) \in \Theta_n \times B^u$, and let

$$0 < T_0 < \cdots < T_k < \cdots$$

be the return times to $I_n$ for $\theta_0$. It is clear from (C1)$_n$ that, for every $0 \leq k \leq T_0$,

$$r_k \in B,$$

$$r_k \notin B^u \implies \theta_k \in \Xi_n^u = \bigcup_{i=0}^n \bigcup_{m=1}^M (I_i + m\omega).$$

Since $\theta_0 \in \Theta_n = \bigcup_{i=0}^n \bigcup_{m=1}^M (I_i + m\omega)$, it follows that $T_0 > M_n$. Therefore, there is a time $0 \leq t < T_0$ such that $\theta_t \in I_n - M_n \omega$. By (C2)$_n$, $\theta_t \in \Theta_{n-1}$, and therefore (C1)$_n$ ensures that $r_t \in B^u$, since $t < T_0$ (the first return to $I_n$). That is, $(\theta, r_t) \in B^u$, and ultimately, $(\theta_{T_0+1}, r_{T_0+1}) \in A_n^u$. Since $A_n^u \cap A_n^s = \emptyset$, and $\Phi_{M_n-1}(A_n^u) = B_n^u$, it follows that

$$\Phi_{M_n-1}(\theta_{T_0+1}, r_{T_0+1}) = (\theta_{T_0+M_n}, r_{T_0+M_n}) \in (I_n + M_n \omega) \times (B \setminus B^u).$$

By (C2)$_n$, $\theta_{T_0+M_n} \in \Theta_{n-1}$, and by lemma 4.2, $r_{T_0+M_n+1} \in B^u$. Note that for every $0 \leq k \leq T_0 + M_n$,

$$\theta_k \in \bigcup_{m=1}^{M_n} (I_n + m\omega) \subset \Xi_n^u.$$ 

Now, (C2)$_n$ implies that $\theta_{T_0+M_n+1} \in \Theta_{n-1}$, and therefore $(\theta_{T_0+M_n+1}, r_{T_0+M_n+1}) \in \Theta_{n-1} \times B^u$. By induction, we show that for arbitrary $l > 0$, and every $0 \leq k \leq T_l$,

$$r_k \notin B^u \implies \theta_k \in \Xi_n^u = \bigcup_{i=0}^n \bigcup_{m=1}^M (I_i + m\omega).$$

Condition $(\mathcal{UH})_n$ now follows. 

\[\square\]

Lemma 7.2. Suppose that $E \in \mathcal{E}_{n-1}^{\leq}$. Then there are two invariant $C^2$ functions $\psi^u, \psi^s : \mathbb{T} \to B$ satisfying for every $\theta \in \mathbb{T}$

$$\psi^s(\theta) < \psi^u(\theta),$$

such that $\psi^u$ is uniformly attracting, and $\psi^s$ is uniformly repelling (in a neighbourhood). Furthermore

$$\psi^u(\theta) \notin B^u \implies \theta \in \Xi_n^u = \bigcup_{i=0}^n \bigcup_{m=1}^M (I_i + m\omega) \quad (7.6)$$

$$\psi^s(\theta) \notin B^s \implies \theta \in \Xi_n^s = \bigcup_{i=0}^n \bigcup_{m=0}^M (I_i - m\omega). \quad (7.7)$$

20
Proof. Consider our set \( \Theta_n = \mathbb{T} \setminus \bigcup_{m=-M}^{M} (I_m + m\omega) \), for which it holds that \( \bigcup_{k=0}^{2M_n+1} (\Theta_n + k\omega) = \mathbb{T} \).

In particular, \((\mathcal{UH})_n\) is satisfied (by lemma 7.1), which implies that the set

\[
\Lambda = \bigcup_{k=0}^{2M_n+1} \Phi^k(\Theta_n \times B^u)
\]

is invariant. Suppose that we have the two initial conditions \((\theta_0, r_0), (\theta_0, r_0') \in \Lambda\). Then \(\theta_k \in \Theta_n \Rightarrow r_k, r_k' \in B^u\). Set \(\Sigma_k = \bigcup_{m=-M}^{M} (I_k - m\omega)\), and \(t = 100M_n \ll N_n\). Then \(\Sigma_k\) is an \((N_k - 2M_k - 1, 2M_k + 1)\)-system, and lemma B.1 implies that

\[
\frac{|\{0 \leq j < t : \theta_j \notin \Theta_n\}|}{t} \leq \sum_{k=0}^{n} \frac{2M_k + 1}{100M_n} + \frac{2M_k + 1}{N_k - 2M_k - 1} \leq \frac{3}{100} \sum_{k=0}^{n} \frac{M_k}{M_n} \leq \frac{1}{10}.
\]

For \(t = 100M_n\), (B.5) gives us

\[
\frac{1}{r_0 \cdots r_{t-1}} \leq \lambda^{-t/2},
\]

\[
\frac{1}{r_0' \cdots r_{t-1}'} \leq \lambda^{-t/2}.
\]

Since \(r_i - r_i' = \frac{m - r_0}{r_0 \cdots r_{t-1}}\), this means that

\[
r_i - r_i' = \frac{r_0 - r_0'}{r_0 r_0' \cdots r_{t-1} r_{t-1}'} \leq (r_0 - r_0') \lambda^{-t},
\]

and so \(\Phi = \Phi^{100M_n}\) is a fibre contraction on \(\Lambda\). This gives us a \(C^2\)-family of attracting invariant curves \(\psi^u_{E_n} : \mathcal{C}_{n-1} \times \mathbb{T} \rightarrow B\) (see for instance [Sta97, Theorems 2.1 and 3.1]). We do the same thing, but for \(\Phi^{-1}\), to obtain our repelling curves \(\psi_{E_n} : \mathcal{C}_{n-1} \times \mathbb{T} \rightarrow B\). By construction, they satisfy the conditions in \((\mathcal{UH})_n\). \(\square\)

### 7.2 The interval \(I(E)\) and the stopping times \(\sigma^\pm\)

The obvious way of constructing these intervals would be to let \(I(E) = I_n + \omega\), if \(E \in \mathcal{C}_{n-1}\). However, our method performs badly close to the endpoints of \(\mathcal{C}_{n-1}\). The reason is that, the time taken for \(\psi^u(\theta_k)\) to stabilise in \(B^u\), if \(\theta_0 \in I_n\), is very similar to the time taken to stabilise if \(\theta_0 \in I_{n-1}\). Since points starting in \(I_n\) could potentially enter \(I_{n-1}\) before they stabilise in \(B^u\), according to \((\mathcal{UH})_n\) and \((\mathcal{C})_n\), this appears to create a double-resonance.

This resonance will never occur, but this is not obvious the way the conditions are formulated. We circumvent this by being flexible with our scales; if we are close to the lowest energy \(E_{n-1}\) of \(\mathcal{C}_{n-1}\), we simply slide the scale to use the previous one, that is \(I_{n-1}\), rather than the one given to us by the induction statement, that is \(I_n\). In order to determine when we can slide the scales, we introduce some stopping times:

Suppose that \(E \in \mathcal{C}_{n-1}\), and \(\theta_0 \in I_0 + \omega\). Let \(\sigma^+ = \sigma^+(\theta, E) \geq 0\) be the smallest positive integer satisfying

\[
|\psi_{E_n}^u(\theta_0 + (\sigma^+ + 1)\omega)| \geq \lambda^{-3}
\]

\[
|\psi_{E_n}^u(\theta_0 + j\omega)| < \lambda^{-3}
\]
for every $0 \leq j \leq \sigma^+$. Similarly, let $\sigma^- = \sigma^- (\theta, E) \geq 0$ be the smallest positive integer satisfying

$$\left| (\psi_E^r - \psi_E^l) (\theta_0 - (\sigma^- + 1) \omega) \right| \geq \lambda^{-3}$$

$$\left| (\psi_E^r - \psi_E^l) (\theta_0 - j \omega) \right| < \lambda^{-3}$$

for every $0 \leq j \leq \sigma^-$. By Lemma 7.1, $E \in \mathcal{L}'_{n-1}$ implies $\langle \mathcal{U} \mathcal{H} \rangle_n$, which implies that the stopping times are well-defined. Indeed, $\Theta_n$ is non-empty, and by $\langle \mathcal{U} \mathcal{H} \rangle_n$ we have $\theta \in \Theta_n \implies \psi^u_E (\theta) \in B^u, \psi^l_E (\theta) \in B^l$, and therefore $|\psi^u_E (\theta) - \psi^l_E (\theta)| \geq \lambda - \lambda^{-1}$. Set

$$\sigma^+_n = \sigma^+ (I_n, E) = \max_{\theta \in I_n + \omega} \sigma^+ (\theta, E)$$

$$\sigma^-_n = \sigma^- (I_n, E) = \max_{\theta \in I_n + \omega} \sigma^- (\theta, E),$$

If, for any $0 < k \leq n$, we have

$$\frac{1}{30} N_{k-1} \leq \max \{ \sigma^+_n, \sigma^-_n \} < \frac{1}{30} N_k, \quad (7.9)$$

then we set

$$I (E) = I_k + \omega. \quad (7.10)$$

**Remark.** The $k$ above goes to infinity as $n$ goes to infinity, that is as $E \not\Rightarrow E_0$.

By $\langle \mathcal{U} \mathcal{H} \rangle_n$ and $(C2)_n$, it follows that $\psi^u_E (\theta, M_k) \in B^u$, and $\psi^l_E (\theta, M_k) \in B^l$, which immediately implies that $\sigma^+ \leq M_n < \frac{1}{30} N_k$. That is, for the parameters $E \in \mathcal{L}'_{n-1}$, (7.9) is satisfied for some $0 < k \leq n$.

Since the return time from $I_k$ to itself is at least $N_k$, this immediately gives us

$$\left( \bigcup_{m=1}^{15 \sigma^+} I + m \omega \right) \cap \left( \bigcup_{m=1}^{15 \sigma^-} I - m \omega \right) = \emptyset,$$

which is the assumption in (6.5).

Now, suppose that $k$ is such that $I = I(E) = I_k + \omega$, and let $\theta_0 \in I + \omega$. Set $\sigma^+ = \sigma^+ (\theta_0, E)$, $r_i = \psi^u (\theta_i)$ and $s_i = \psi^l (\theta_i)$. Since $\langle \mathcal{U} \mathcal{H} \rangle_n$ is satisfied, (7.4) implies that

$$r_j \not\in B^u \implies \theta_j \in \Xi_n^u = \bigcup_{i=0}^{n} \bigcup_{m=1}^{M_i} (I_i + m \omega).$$

However, we might be dealing with the situation where $k < n$, in which case

$$\bigcup_{i=k+1}^{n} \bigcup_{m=1}^{M_i} (I_i + m \omega)$$

could be replaced by something even better, since in that case $\sigma^+_n < N_k \ll M_{k+1}$. That is, $r_j$ might stabilise in $B^u$ much earlier than predicted by $\langle \mathcal{U} \mathcal{H} \rangle_n$. If we set

$$\Sigma_j^u = \bigcup_{m=1}^{M_i} (I_i + m \omega) \quad \text{for } 0 \leq i < k,$$

and

$$\Sigma_k^u = \bigcup_{m=1}^{10 M_{k-1}} I_k + m \omega,$$
\[ \Sigma_i^t = \bigcup_{m=0}^{M_i} (I_i - m\omega) \] for \( 0 \leq i < k \), and 
\[ \Sigma_k^t = \bigcup_{m=0}^{\sigma_n^{-10M_{k-1}}} I_k - m\omega, \]
then we have the following result.

**Lemma 7.3.** Suppose that \( I = I(E) = I_k + \omega \), then

\[ r_j \notin B^u \implies \theta_j \in \bigcup_{i=0}^{k} \Sigma_i^t, \]  \hspace{1cm} (7.11)

and

\[ s_j \notin B^s \implies \theta_j \in \bigcup_{i=0}^{k} \Sigma_i^t, \]  \hspace{1cm} (7.12)

**Proof.** Since the forward iterates of \( I = I_k + \omega \) cover the circle, and we start with \( \theta_0 \in I \), it suffices to show that it holds for every \( 0 \leq j \leq N(\theta_0) \), where \( N(\theta_0) \) is the first return of \( \theta_0 \) to \( I \). That is, every iterate of \( \theta_0 \) can be identified with the iterate \( \tilde{\theta}_j \), where \( \_0 \in I \), and \( 0 \leq j \leq N(\theta_0) \), which implies the claim for arbitrary iterates \( r_j \).

By lemma B.5, there is a 0 \( j \leq 10M_{k-1} \) (depending on \( \theta_0 \)), satisfying that

\[ (\tilde{\theta}_{\sigma^+(\theta_0)+j}, r_{\sigma^+(\theta_0)+j}) \in \Theta_{k-1} \times B^u. \]

Condition \((\mathcal{UH})_n\) implies \((C1)_k\), which further implies that

\[ r_j \notin B^u \implies \theta_j \in \Sigma_{k-1}^t = \bigcup_{i=0}^{k-1} (I_i + p\omega), \]

for every \( \sigma^+(\theta_0) + 10M_{k-1} \leq j \leq N \), where \( N \) is the first return to \( I_k \). The only iterates we haven’t covered are \( 0 \leq j < \sigma^+(\theta_0) + 10M_{k-1} \), which are in \( \Sigma_k^t \), and thus (7.11) follows.

The other statement is proved in the exact same way, but iterating backwards. \( \square \)

### 7.3 Proving \((A3)\)

Suppose that \( E \in \mathcal{E}_{n}^{L,1} \), for some \( n \geq 0 \), and that \( I(E) = I_k + \omega \) where \( k \leq n \). By lemma 7.1, \((\mathcal{UH})_n\) is satisfied. We begin with proving (6.6). The proof of (6.9) is completely analogous, but iterating the other direction.

Let \( \theta_0 \in I \), and set \( r_0 = \psi_E^t(\theta_0), s_0 = \psi_E^s(\theta_0) \) and

\[ \Sigma_i^t = \bigcup_{m=0}^{M_i} (I_i - m\omega) \text{ for } 0 \leq i < k, \text{ and } \Sigma_k^t = \bigcup_{m=0}^{\sigma_n^{-10M_{k-1}}} I_k - m\omega, \]

For every \( 0 \leq j < k \), set \( a_j = N_j - M_j - 2 \), \( r_k = N_j - M_j - 1 \), and \( l_j = M_j + 1 \). Set \( a_k = N_k - (\sigma_n^{-1} + 10M_{k-1} + 1) \geq \frac{28}{30} N_k \) (recall (7.9) and that \( M_{k-1} \ll N_k \)), \( r_k = N_k - (\sigma_n^{-1} + 10M_{k-1} + 1) \geq \frac{28}{30} N_k \), and \( l_k = \sigma_n^{-1} + 10M_{k-1} + 1 \leq \frac{2}{30} N_k \). Then lemma B.1 applies to \( \Sigma_j^t \), and \( 0 \leq j \leq k \), giving us

\[ \frac{|\{0 \leq i < t : \theta_i \in \bigcup_{j=0}^{k} \Sigma_j^t\}|}{t} \leq \sum_{j=0}^{k-1} \frac{M_j + 1}{N_j - 1} + \frac{2}{28} \leq \frac{1}{2}, \]  \hspace{1cm} (7.13)
if $\lambda$ is sufficiently large. Since (7.12) is satisfied, lemma B.4 implies that, for every $0 \leq i \leq \sigma^+$, we have

$$D_{0,i}(r_0,s_0) \geq \lambda^{(i+1)/2}. \quad (7.14)$$

Analogously, one can show that

$$D_{-i,0}(r_0,s_0) \leq \lambda^{-(i+1)/2}, \quad (7.15)$$

for every $0 \leq i \leq \sigma^-$. We now turn to (6.7) and (6.10). Again, their proofs are nearly identical, and we will only write down the proof of the first one. Let $0 \leq \hat{\sigma} \leq \sigma^+$ be the largest such that

$$\theta_{\hat{\sigma}} \not\in \bigcup_{j=0}^{k-1} \bigcup_{j=0}^{k} \Sigma^j + m\omega,$$

where $\hat{M}_j = 20 \cdot 2^j M_j$. Since, the ratio between $M_j$ and $N_j$ grows super-exponentially fast (see (7.1) and (7.2)), it follows that $20 \cdot 2^j M_j < N_j$, for every $j > 0$, provided that $\hat{\lambda}$ is sufficiently large. Since $N_j$ is a lower bound of the return time from $I_j$ to itself, there has to be such a $\hat{\sigma}$. That is, if we set

$$r_j = N_j - (M_j + 1) - \hat{M}_j,$$

and $l_j = M_k + 1 + \hat{M}_k$, then lemma B.1 gives us that

$$\frac{|\{0 \leq i < t : \theta_i \in \bigcup_{j=0}^{k} \Sigma_j\}|}{t} \leq \sum_{j=0}^{k} \frac{M_j + 1 + \hat{M}_j}{t} \frac{M_j + 1 + \hat{M}_j}{N_j} \leq \frac{1}{2},$$

provided $t \geq 3\hat{M}_j$, and $\lambda$ is large enough. This means that $\sigma^+ - \hat{\sigma} \leq 3\hat{M}_j = 60 \cdot 2^{k-1} M_{k-1}$. Since $\sigma^+_n \geq \frac{1}{20} N_{k-1}$, by (7.9), it follows that

$$\frac{\sigma^+ - \hat{\sigma}}{\sigma^+_n} \leq \frac{60 \cdot 2^{k-1} M_{k-1}}{\frac{1}{20} N_{k-1}} \xrightarrow{k \to \infty} 0,$$

or that $\sigma^+ - \hat{\sigma} = \eta(E) \sigma^+_n$, where $\eta(E) \searrow 0$ as $E \to E_0$.

Since $\theta_{\hat{\sigma}}$ starts far away (iterating backwards) from the sets $\Sigma^j$, we can get good estimates going backwards. That is, if we set $r_j = N_j - (M_j + 1), l_j = M_j + 1$, and $a_j = \hat{M}_j = 20 \cdot 2^j M_j$, then lemma B.1 gives us for every $0 \leq \tau \leq \hat{\sigma}$ that

$$\frac{|\{j : \tau < j \leq \hat{\sigma}, \theta_j \in \bigcup_{i=0}^{k-1} \Sigma_i\}|}{\hat{\sigma} - \tau} \leq \sum_{i=0}^{m-1} \frac{M_i + 1}{20 \cdot 2^i M_i} \leq \frac{1}{2},$$

Again, using lemma B.4, we obtain for every $0 \leq \tau \leq \hat{\sigma}$ the inequality

$$D_{\tau,\hat{\sigma}} \geq \lambda^\left(\frac{\hat{\sigma} - \tau + 1}{2}\right).$$
This concludes the proof of (6.7).

In the same way, one can show that there is a \( \tilde{\sigma} \) and \( \eta(E) \) \( \to \) 0 (as \( E \to E_0 \)), such that for every \( 0 \leq \tau \leq \tilde{\sigma} \)

\[
D_{-\tilde{\sigma},-\tau} \geq \lambda^{(\tilde{\sigma}-\tau+1)/2},
\]

where again \( \sigma^- - \tilde{\sigma} = \eta(E)\sigma_n^- \).

In order to prove (6.8), we note that

\[
s_j \notin B^\varepsilon \implies \theta_j \in \bigcup_{i=0}^k \Sigma_i^l,
\]

by (7.12). Let \( \theta \in I, 0 \leq \tau \leq \sigma^+ \) and \( t \geq \tau \). We wish to bound \( D_{\tau,t}(\psi_E^l(\theta), \psi_E^l(\theta)) \). Similarly to before, find the smallest \( \tau \leq \tilde{\sigma} \) such that

\[
\theta_{\tilde{\sigma}} \notin \bigcup_{j=0}^{k-1} \Sigma_j^l - m \omega.
\]

The rest of the proof is simply showing that \( \tilde{\sigma} - \tau = \eta(E)\sigma_n^- \) for some \( \eta(E) \) that goes to 0 as \( E \to E_0 \), and that we have

\[
D_{\tilde{\sigma},t}(\psi_E^l(\theta), \psi_E^l(\theta)) \geq \lambda^{(t-\tilde{\sigma}+1)/2},
\]

whenever \( \tilde{\sigma} \leq t \). The proof proceeds in a manner analogous to the proof above. One can prove (6.11) in a similar way.

### 7.4 Proving (A2)

The assumption (6.1), and also (2.2) in theorem 2.1, follow from proposition 7.7. The assumption (6.3) follows from (7.19). The assumption (6.2) requires a little bit more care, but follows from the interval \( I \) having a global minimum, by lemma 7.5, and the uniform bounds on the second derivative in (7.16). As for the length of the interval, the assumption (6.4), it is shown in lemma 7.8.

We remark that, as has already been said, as \( E \) gets closer to \( E_0 \), the \( k \) that satisfies \( I(E) = I_k + \omega \) goes to infinity. That is, for the asymptotic statements, we can simply choose \( k \) as big as we want.

The first result here is about the derivative above the critical interval \( I = I(E) \).

**Lemma 7.4.** Suppose that \( n \geq 1 \) and \( E \in \epsilon_n^{l+1} \). Then there is a positive constant \( C, \) independent of \( n \) and \( E, \) such that

\[
\frac{1}{C} \lambda^2 \leq |\partial^2 \psi_E^l(\theta)| \leq C \lambda^2,
\]

\[
-1 - \frac{4}{\lambda^2} \leq |\partial E(\psi_E^l(\theta))| \leq -1 + \frac{4}{\lambda^2}.
\]

\[
-\frac{32}{\lambda^2} \leq |\partial^2 E(\psi_E^l(\theta))| \leq \frac{32}{\lambda^2},
\]

where \( I = I(E), \) provided \( \lambda \) is sufficiently large. Moreover, there is a positive constant \( C_1, \) independent of \( n \) and \( E, \) such that

\[
|\partial \psi_E^l(\theta)| \leq C_1, \text{ and }
|\partial \psi_E^l(\theta)| \leq C_1,
\]

for every \( \theta \in I. \)
Proof. Suppose that \( k \) is such that \( I = I(E) = I_k + \omega \), and let \( \theta_0 \in I + \omega \). Set \( \sigma^+ = \sigma^+(\theta_0) \) and \( r_i = \psi^\mu(\theta_i) \). By lemma 7.3, we have

\[
\forall j \notin B^\mu \implies \theta_j \in \bigcup_{i=0}^{k} \Sigma_j^i,
\]

where

\[
\Sigma_j^i = \bigcup_{m=1}^{M_j} (I_j + m\omega) \text{ for } 0 \leq j < k, \text{ and } \Sigma_k^i = \bigcup_{m=1}^{\sigma_k^+ + 10M_k-1} I_k + m\omega.
\]

Since we will iterate backwards from \( \theta_{-1} \), we set \( a_j = N_j - M_j, r_j = N_j - M_j \), and \( l_j = M_j \), when \( 0 \leq j < k \), and \( a_k = N_k - (\sigma_k^+ + M_k-1) + 1 \geq \frac{28}{30}N_k, r_k = N_k - (\sigma_k^+ + M_k-1) \geq \frac{28}{30}N_k \), and \( l_k = \sigma_k^+ + M_k-1 \leq \frac{2}{3}N_k \). Therefore lemma B.1 gives us the estimates

\[
|\{1 \leq i \leq t : \theta_{-i} \in \bigcup_{j=0}^{k} \Sigma_j^i\}| \leq \frac{k-1}{t} M_j \leq \frac{k-1}{28} + \frac{2}{28} \leq \frac{1}{2},
\]

for every \( t > 0 \), provided \( \lambda \) is large enough. Then lemma B.3 gives us the estimate

\[
r_{-j}^a \cdots r_{-1}^a \geq \lambda^{\frac{1}{2}}. \]

Moreover, lemma 7.3 implies that \( r_0 \in B^\mu = [\lambda^{-2}, \lambda^2] \), since \( I_k \cap \bigcup_{i=0}^{k} \Sigma_i^i = \emptyset \). Therefore

\[
r_{-i}^a \cdots r_{-1}^a r_0^a \geq \lambda^{\frac{1}{2} + a_0}.
\]

Since \( \psi^\mu_E \) is \( C^2 \) in both \( \theta \) and \( E \), there is a constant \( C_E > 0 \) depending only on \( E \) such that \( \| \psi^\mu_E \|_{C^2(\theta, E)} \leq C_E \). If we choose \( t > 0 \) large enough (depending only on \( E \)), we can ensure that

\[
|\frac{\partial^r_j r_{-j}}{r_{-i}^a \cdots r_{-1}^a r_0^a}| \text{ and } |\frac{\partial^l_j r_{-j}}{r_{-i}^a \cdots r_{-1}^a r_0^a}|
\]

are as small as we wish, for \( i = 1, 2 \). Therefore, lemma D.1 applies with \( c_1 = c_2 = 1 \). We thus obtain the inequalities

\[
|\partial_\theta^r \psi^\mu(\theta_1) + 1| \leq \frac{2}{\lambda^2}
\]

\[
|\partial_\theta^2 \psi^\mu(\theta_1)| \leq \frac{16}{\lambda^2}
\]

\[
|\partial_\theta^r \psi^\mu(\theta_1) - \lambda^2 v^r(\theta_N)| \leq 2\| v \|_{C^1}
\]

\[
|\partial_\theta^2 \psi^\mu(\theta_1) - \lambda^2 v''(\theta_N)| \leq 16\lambda \| v \|_{C^1}^2 + 8\lambda \| v \|_{C^2}^2 \leq 50\lambda \| v \|_{C^2},
\]

provided that \( \lambda \) is sufficiently large. Iterating backwards, the same can be done for \( \psi^\nu \), letting \( \theta_0 \in I \). Upon applying lemma D.2, for the same \( c_1 = c_2 = 1 \), we obtain the inequalities

\[
|\partial_\theta^r \psi^\nu(\theta_0)| \leq \frac{2}{\lambda^2}
\]

\[
|\partial_\theta^2 \psi^\nu(\theta_0)| \leq \frac{16}{\lambda^2}
\]

\[
|\partial_\theta^r \psi^\nu(\theta_0)| \leq 2\| v \|_{C^1}
\]

\[
|\partial_\theta^2 \psi^\nu(\theta_0)| \leq \frac{16\lambda^2 \| v \|_{C^1}^2}{\lambda^2} + 8\lambda \| v \|_{C^1} + 2\| v \|_{C^2} \leq 50\lambda \| v \|_{C^2}.
\]
The inequalities (7.17) and (7.18) follow immediately. For any \( \theta \in I \), we also have
\[
|\partial^2_\theta \psi''(\theta) - \partial^2_\theta \psi'(\theta) - \lambda^2 \psi''(\theta)| \leq 100\lambda \|v\|_{C^2}.
\]
Since \( \psi''(\theta) > 0 \) in \( I_0 \), we see that the difference \( \partial^2_\theta (\psi''_E - \psi''_E) \geq \text{const} \cdot \lambda^2 \) on \( I \), provided that \( \lambda \) is large enough.

**Lemma 7.5.** Suppose that \( E \in \mathcal{E}_{n-1}^{cl} \). Then the minimum of
\[
\min_{\theta \in I} |\psi^\mu_E(\theta) - \psi^\nu_E(\theta)|
\]
is globally unique, and attained in \( I = I(E) \). That is, there is a unique \( \theta_c = \theta_c(E) \in I \) such that
\[
|\psi^\mu_E(\theta_c) - \psi^\nu_E(\theta_c)| < \min_{\theta \in \mathcal{T} \setminus \{\theta_c\}|} |\psi^\mu_E(\theta) - \psi^\nu_E(\theta)| < \min_{\theta \in \mathcal{T} \setminus I} |\psi^\mu_E(\theta) - \psi^\nu_E(\theta)|.
\]

**Proof.** Since \( E \in \mathcal{E}_{n-1}^{cl} \), lemma 7.1 implies that \((\mathcal{U}_H)_n\) is satisfied. Suppose that \( 0 \leq k \neq n \) is such that \( I(E) = I_k + \infty \), and let \( \theta_0 \in I \). Set \( \sigma^\pm = \sigma^\pm(\theta_0,E), r_i = \psi^\mu_E(\theta_i) \) and \( s_i = \psi^\nu_E(\theta_i) \). Then (7.14) gives us for every \( 0 \leq i \leq \sigma^+ \), that
\[
D_{0,i-1} \geq \lambda^{i/2}
\]
and (7.15) gives us for every \( 0 \leq i \leq \sigma^- \), that
\[
D_{-i+1,0} \geq \lambda^{i/2}.
\]
The bounds in (7.9) imply that either
\[
\max_{\theta \in I(E)} D_{0,\sigma^+-1}(\psi^\mu_E(\theta), \psi^\nu_E(\theta)) \geq \lambda^{\frac{1}{2}N_k-1},
\]
or
\[
\min_{\theta \in I(E)} D_{-\sigma^-,0}(\psi^\mu_E(\theta), \psi^\nu_E(\theta)) \leq \lambda^{-\frac{1}{2}N_k-1}.
\]
Since, \( r_i - s_i = D_{0,i-1}(r_0 - s_0) \) and \( r_i - s_i = D_{-i+1,0}(r_0 - s_0) \), for \( i \geq 1 \), this shows that
\[
r_0 - s_0 < \min_{-\sigma^- \leq \sigma^+ \neq 0} r_i - s_i.
\]
Moreover, since \( \mathbb{T} \times [\lambda^{-2}, \lambda^2] \) is invariant, it shows that
\[
\min_{\theta \in \mathcal{T}} \psi^\mu_E(\theta) - \psi^\nu_E(\theta) \leq \lambda^{2-\frac{1}{2}N_k-1}. \tag{7.20}
\]
Using lemmas B.5 and B.6, we get \( 0 \leq j^\pm \leq 10M_{k-1} \), such that \( \theta_{\pm(\sigma^+ + j^+)} \in \Theta_{k-1}, r_{\sigma^+ + j^+} \in B^\mu \) and \( s_{-\sigma^- - j^-} \in B^\nu \). Using the trivial bound \( D_{\sigma^+, \sigma^+ + 10M_{k-1}-1} \geq \lambda^{-40M_{k-1}} \), this means that
\[
r_i - s_i \geq (r_{\sigma^+} - s_{\sigma^+}) \cdot \lambda^{-40M_{k-1}} \geq \lambda^{-40M_{k-1}-3}, \tag{7.21}
\]
for \( \sigma^+ < i \leq \sigma^+ + 10M_{k-1} \), and similarly for \( -(\sigma^- + 10M_{k-1}) \leq i < -\sigma^- \). This is much larger than the lower bound in (7.20). This proves the statement for \( \theta \in I + m\omega \), and every \( -\sigma^- - 10M_{k-1} \leq m \leq \sigma^+ + 10M_{k-1} \). For the remaining \( \theta \), we use (C1)\(_k\) (which is guaranteed by \((\mathcal{U}_H)_n\)), to get
\[
r_i \notin B^\mu \iff \theta_i \in \mathcal{Z}^\mu_{k-1} = \bigcup_{j=0}^{M_j} (I_{j} + m\omega),
\]
for some
for every $\sigma^+ + j^+ \leq i \leq N^+$, where $N^+$ is the first return to $I$ iterating forward, and
\[ s_{-i} \not\in B^i \implies \theta_{-i} \in \mathbb{Z}_k^+ = \bigcup_{j=0}^{k-1} (I_j - m\omega), \]
for every $\sigma^- + j^- \leq i \leq N^-$, where $N^-$ is the first return to $I$ iterating backward. Setting $r_j = N_j - M_j - 1$ and $l_k = M_j + 1$ for every $0 \leq j \leq k - 1$, both sets $\mathbb{Z}_k^+$ and $\mathbb{Z}_k^-$ satisfy the conditions in lemma B.1, giving us
\[ |\{0 \leq i < t : \theta_i \in \bigcup_{j=0}^{k-1} \Sigma_j\}| \leq \sum_{j=0}^{k-1} \frac{M_j + 1}{t} + \frac{M_j + 1}{N_j}, \]
where $\Sigma_j$ is either $\bigcup_{m=1}^{M_j} (I_j + m\omega)$ or $\bigcup_{m=0}^{M_j} (I_j - m\omega)$. If we choose $t = 2M_{k-1}$, and $\lambda$ is sufficiently large, then
\[ |\{0 \leq i < t : \theta_i \in \bigcup_{j=0}^{k-1} \Sigma_j\}| \leq \frac{1}{2} \lambda. \]
showing that there are $\theta_i$’s spaced at most $2M_{k-1}$ steps apart, satisfying $\theta_i \not\in (\mathbb{Z}_k^+ \cup \mathbb{Z}_k^-)$, and therefore that $r_i \in B^i, s_i \in B^i$, meaning that
\[ r_i - s_i \geq \lambda - \lambda^{-1} \geq \frac{1}{2} \lambda. \]
Again, using the trivial bounds on the distance increase, we see that difference has to be $\geq \frac{1}{2} \lambda^{1 - 8M_{k-1}}$, between such $\theta_i$’s. This proves that the minimum is attained in
\[ \bigcup_{m=-\sigma^- - 10M_{k-1}}^{\sigma^+ + 10M_{k-1}} I + m\omega, \]
which together with the bounds in (7.20) shows that the difference is minimised in $I$. This minimum has to be unique, because of the non-degeneracy condition provided by the bounds in (7.16).

For an upcoming paper, we need a result that is hidden in the proof of the above lemma. Specifically, we have the following result.

**Lemma 7.6.** For $E < E_0$ sufficiently close to $E_0$, we have
\[ d(\theta) \gg \sqrt{d(\theta_e)} \] (7.22)
for every $\theta \not\in \{\theta \in I + m\omega : -\sigma^- (\theta, E) \leq m \leq \sigma^+ (\theta, E)\}$.

**Proof.** The bounds in (7.20) imply the bounds
\[ d(\theta_e) \leq \lambda - 40M_{k-1}. \]
In the rest of the above proof, we show that, outside of $\{\theta \in I + m\omega : -\sigma^- (\theta, E) \leq m \leq \sigma^+ (\theta, E)\}$, the difference is at least $\lambda^{-40M_{k-1} - 3}$, using (7.21), and the bound $\frac{1}{2} \lambda^{1 - 8M_{k-1}}$ given at the end of the proof. Since $M_{k-1} \ll N_{k-1}$, we see that $d(\theta) \gg \sqrt{d(\theta_e)}$ outside the set $\{\theta \in I + m\omega : -\sigma^- (\theta, E) \leq m \leq \sigma^+ (\theta, E)\}$. □
The next result shows that the minimum difference is asymptotically linear as $E \nearrow E_0$.

**Proposition 7.7.** Let $\theta_\ast = \theta_\ast(E)$ be the point where that minimises the difference between the two curves $\psi^\pm_E$ and $\psi^\pm_{E'}$. Then the difference at $\theta_\ast$ satisfies

$$d(\theta_\ast) = \min_{\theta \in I} |\psi^\pm_E(\theta) - \psi^\pm_{E'}(\theta)| = \text{const} \cdot (E_0 - E) + o(E_0 - E),$$

(7.23)
as $E \nearrow E_0$, where the constant satisfies $-1 - \frac{4}{\lambda^2} \leq \text{const} \leq -1 + \frac{4}{\lambda^2}$.

**Proof.** For any $E < E_0$ sufficiently close to $E_0$, $(\mathcal{U}(\lambda))_n$ is satisfied for some $n \geq 0$. By lemma 7.5, it is sufficient to consider only $\theta \in I$. Set $\delta(E) = d(\theta_\ast(E))$, and extend it continuously up to $E_0$, where the value is 0. Taylor expansion of $\delta$ gives

$$\delta(E) = \delta(E) - \delta(E_0) = \partial_E \delta(E)(E - E_0) - \partial_E^2 \delta(E)(E - E_0)^2,$$

where $E < \tilde{E} < E_0$. By the estimate in (7.18), the second derivative is uniformly bounded, and the inequality in (7.17) gives us the desired bounds of the constant.

**Lemma 7.8.** For any $E < E_0$ sufficiently close to $E_0$, the length of the interval $I(E)$ satisfies

$$|I(E)| \geq C \sqrt{d(E)},$$

where $C > 0$ can be made arbitrarily large as $E \nearrow E_0$.

**Proof.** Suppose that $I_E = I_k + \omega$ for some $0 \leq k \leq n$. Then, by (7.9), we must have

$$\frac{1}{15} N_{k-1} < \max \{\sigma^+_n, \sigma^-_n\} \leq \frac{1}{15} N_k.$$

By (A.6), we have $|I_k| = c_0/\lambda^{M_{k-1}/2}$. Suppose that the maximum is attained for $\sigma^+_n$, and let $\theta_0 \in I$ be such that $\sigma^+_n(\theta_0, E) = \sigma^+_n$. Then

$$d(\theta_0) D_{0, \sigma^+_n}(\theta_0) = d(\theta_{\sigma^+_n}) \leq \lambda^2,$$

and using the estimate (7.14) for $k = \sigma^+_n$ gives us the inequality

$$\sigma^+_n \leq \text{const} + \max_{\theta \in I_k} \log \sqrt{\lambda} \left( \frac{1}{d(\theta_0)} \right) \leq \text{const} + \log \sqrt{\lambda} \left( \frac{1}{\delta(E)} \right),$$

where $\delta(E)$ is the smallest distance $\delta(E) = \min_{\theta \in I} d(\theta)$. Therefore

$$2M_{k-1} \leq \frac{1}{15} N_{k-1} \leq \text{const} + \log \sqrt{\lambda} \left( \frac{1}{\delta(E)} \right).$$

As $E$ gets closer to $E_0$, the distance approaches 0, and it follows that

$$|I_m| = c_0/\lambda^{M_{k-1}/2} \geq \text{const} \cdot \sqrt{\delta(E)},$$

where the constant can be made arbitrarily large as $n \to \infty$, and therefore as $E \nearrow E_0$. \qed
A Summary of inductive construction

In this section we will summarize the results in [Bje07] that we will use. In particular, the results hold for sufficiently large $\lambda$, and $E \in [-1, E_0)$, where $E_0$ is the lowest energy of the spectrum. Recall the definitions and notation we introduced in section 4.3. We recall briefly that

$$A_n^\mu = \{(\theta, r) | \theta \in I_n + \omega, \phi_{n, \mu}^-(\theta, E) \leq r \leq \phi_{n, \mu}^+(\theta, E)\},$$
$$A_n^\nu = \{(\theta, r) | \theta \in I_n + \omega, \phi_{n, \nu}^-(\theta, E) \leq r \leq \phi_{n, \nu}^+(\theta, E)\}.$$

Moreover, we had the sets

$$\Xi_n^\mu = \bigcup_{i=0}^{n-1} \bigcup_{m=1}^{M_i} (I_i + m\omega),$$
$$\Xi_n^\nu = \bigcup_{i=0}^{n-1} \bigcup_{m=1}^{M_i} (I_i - m\omega),$$
$$\Theta_n = \mathbb{T} \setminus (\Xi_n^\mu \cup \Xi_n^\nu).$$

The following conditions appear in the statement of the result:

**Condition (C1)$_n$**

If $(\theta_0, r_0) \in \Theta_{n-1} \times B^\mu$, and $N \geq 0$ is the smallest positive integer such that $\theta_N \in I_n$, then for every integer $0 \leq k \leq N$

$$r_k \in B, \quad r_k \not\in B^\mu \implies \theta_k \in \Xi_{n-1}^\mu = \bigcup_{i=0}^{n-1} \bigcup_{m=1}^{M_i} (I_i + m\omega).$$

If $(\theta_0, r_0) \in \Theta_{n-1} \times B^\nu$, and $N \geq 0$ is the smallest positive integer such that $\theta_{-N} \in I_n + \omega$, then for every integer $0 \leq k \leq N$

$$r_{-k} \in B, \quad r_{-k} \not\in B^\nu \implies \theta_{-k} \in \Xi_{n-1}^\nu = \bigcup_{i=0}^{n-1} \bigcup_{m=1}^{M_i} (I_i - m\omega).$$

**Condition (C2)$_n$**

For $i = 0, 1$

$$I_n \pm (M_n + i)\omega \subset \Theta_{n-1}.$$  \hspace{1cm} (A.4)

Note that in the below statement, (C2)$_n$ appears in a different place in the original article. Since the base dynamics is independent of $E$, we see that (C2)$_n$ does indeed depend only on $I_n$ and $M_n$. Therefore, the contents of the result remain unchanged.

**Lemma A.1** ([Bje07, Lemma 5.3]). Assume that $\lambda$ is sufficiently large. Then there is an infinite sequence of integers $M_0 < \cdots < M_k < \cdots$, and infinite sequences of closed non-empty intervals $I_0 \supset \cdots \supset I_k \supset \cdots$ and $\mathcal{E}_1 \supset \cdots \supset \mathcal{E}_k \supset \cdots$, satisfying (C2)$_n$ and

$$\lambda^{M_{j-1}/(4\tau)} \leq M_j \leq 2\lambda^{M_{j-1}/(4\tau)}$$ \hspace{1cm} (A.5)

$$|I_j| = c_0/\lambda^{M_{j-1}/2}$$ \hspace{1cm} (A.6)

$$\mathcal{E}_j \subset \text{Interior}(\mathcal{E}_{j-1})$$ \hspace{1cm} (A.7)
for every $j \geq 1$. The condition (C1)$_n$ above is satisfied for every $E \in \mathcal{E}_{n-1}$ and $n \geq 0$. Finally, for every $E \in \mathcal{E}_n$ and $\theta \in (I_n + \omega) \setminus (\frac{1}{2} I_n + \omega)$

$$\phi_n^{u-}(\theta) > \phi_n^{s+}(\theta),$$

(A.8)

and if we write $\mathcal{E}_n = [E_n^-, E_n^+]$, then for $E = E_n^-$, there is a unique $\theta^* \in \frac{1}{2} I_n + \omega$ such that

$$\phi_n^{u-}(\theta^*) = \phi_n^{s+}(\theta^*).$$

(A.9)

Recall that, for every $n \geq 0$, we have set

$$\mathcal{E}^l_n = [E_n^-, E_{n+1}^-] \subset \mathcal{E}_n \setminus \mathcal{E}_{n+1},$$

(A.10)

where we use the notation $\mathcal{E}_n = [E_n^-, E_n^+]$. The following simple observation is buried in the proof of lemma A.1, and is not crucial to the argument. We include it, simply to reassure the readers, that the sets $\mathcal{E}^l_n$ are non-empty.

**Lemma A.2.** For every $n \geq -1$, the set $\mathcal{E}^l_n$ is non-empty.

**Proof.** Let $E = E_n^- \in \mathcal{E}_n$, and consider the sets $B_{n+1}^u$ and $B_{n+1}^s$ given in (4.11). Iterating $B_{n+1}^u$ forward by $M_{n+1} - M_n - 1$ steps, the result lies over $I_{n+1} - (M_n + 1) \omega$. Similarly, iterating $B_{n+1}^u$ backwards by $M_{n+1} - M_n - 1$ steps, the result lies over $I_{n+1} + (M_n + 1) \omega$. Since $E = E_n^- \in \mathcal{E}_n$, the conditions (C1)$_{n+1}$ and (C2)$_n$ are satisfied. Therefore (C1)$_{n+1}$ implies that

$$\Phi_{E}^{M_{n+1} - M_n} (B_{n+1}^u) \subset (I_n - (M_n + 1) \omega) \times B^s,$$

and (C2)$_n$ implies that both $I_n + (M_n + 1) \omega$ and $I_n - (M_n + 1) \omega$ have empty intersection with $I_0$. Applying lemma 4.2, we obtain

$$\Phi_{E}^{M_{n+1} - M_n} (B_{n+1}^u) \subset (I_n - M_n \omega) \times (\lambda, \lambda^2] \subset B^u_n,$$

and $\Phi_{E}^{-(M_{n+1} - M_n)} (B_{n+1}^s) \subset (I_n + M_n \omega) \times [\lambda^{-2}, \lambda^{-1}) \subset B^s_n$.

Recall that $\phi_n^{u-} = \Phi_{E}^{M_{n+1}} (I_n - M_n \omega) \times \{\lambda\}$, the lower boundary of $\Phi_{E}^{M_{n+1}} (B^u_n)$, and $\phi_n^{s+} = \Phi_{E}^{M_{n+1}} (I_n - M_n \omega) \times \{\lambda^{-1}\}$, the upper boundary of $\Phi_{E}^{-(M_{n+1})} (B^s_n)$. Since we have restricted to parameters that preserve orientation, the lower (upper) boundary of $B^u_n$ ($B^s_n$) are indeed the forward (backward) iterates of the endpoints $\lambda$ and $\lambda^{-1}$.

Since the intervals $(\lambda, \lambda^2]$ and $[\lambda^{-2}, \lambda^{-1})$ don’t include the endpoints, this means that $\phi_n^{s+} < \phi_n^{s+}$, and similarly $\phi_n^{u-} < \phi_n^{u-}$. Since $E \in \mathcal{E}_n$, (A.8) and (A.9) imply that $\phi_n^{u-} \leq \phi_n^{u-}$ in $I_{n+1} + \omega$. It follows that $A_{n+1}^u$ and $A_{n+1}^s$ do not intersect for $E = E_n^-$. Since $E = E_n^- \leq E_{n+1}^-$ and they do intersect for $E_{n+1}^-$ (again using (A.9), but for $n+1$), it follows that $E_n^- < E_{n+1}^-$. Therefore $\mathcal{E}^l_n$ is non-empty. \qed

**B Abstract growth estimates**

This section is divided into two parts. The first part is independent of the model at hand, and simply gives bounds on the relative time spent in certain collections of interval systems. The second part gives growth estimates for the expansion, given the previous estimates applied to interval systems satisfying some conditions. In the end, these will all be applied to the collection of interval systems $\mathfrak{U}_h$ (recall their definition in (4.8) and (4.9)).
B.1 Relative time spent in interval systems

Suppose that we are given a $\Sigma \subset T$. We call $r > 0$ the minimal return time if $\theta_i, \theta_{i+j} \in \Sigma$, but $\theta_{i+s} \notin \Sigma$ for some $0 < s < j$, forces $j > r$. That is, once $\theta_i$ leaves $\Sigma$, then it won’t return to $\Sigma$ for at least $r$ iterates.

Similarly, we call $l > 0$ the maximal confinement time if $\theta_i, \ldots, \theta_{i+j} \in \Sigma$ forces $j \leq l$. That is, a point can stay in $\Sigma$ for at most $l$ successive iterates.

If $\theta_0 \in \Sigma$, then we say that it has accumulation time $a \geq 0$, with respect to $\Sigma$, if $\theta_i \not\in \Sigma$ for $0 \leq i < a$. That is, $\theta_0$ enters $\Sigma$ after $a$ iterations, but not before that.

We will refer to $\Sigma$ as an $(r,l)$-system, and to $(\Sigma, \theta_0)$ as an $(r,l,a)$-system.

In the same way, we define reversed $(r,l,a)$-systems, having an accumulation criterion, but iterating backwards. The return and confinement conditions are the same, but instead, we say that the system $(\Sigma, \theta_0)$ has reversed accumulation time $a \geq 0$, if $\theta_i \not\in \Sigma$ for $0 \leq l < i$.

Lemma B.1. Let $n \geq 0$ be an integer and $\theta_0 \in T$. Suppose that for every $0 \leq k \leq n$ we are given a $\Sigma_k$ such that $(\Sigma_k, \theta_0)$ is an $(r_k, l_k, a_k)$-system. Then for every $0 < t$ we have the upper bounds

$$\frac{|\{0 \leq j < t : \theta_j \in \bigcup_{k=0}^n \Sigma_k\}|}{t} \leq \sum_{k=0}^n \frac{l_k}{r_k + l_k}, \tag{B.1}$$

and

$$\frac{|\{0 \leq j < t : \theta_j \in \bigcup_{k=0}^n \Sigma_k\}|}{t} \leq \sum_{k=0}^n \frac{l_k}{m_k + l_k}, \tag{B.2}$$

where $m_k = \min\{a_k, r_k\}$.

Remark. It is clear that any $(r_k, l_k)$-system $\Sigma_k$ makes an $(r_k, l_k, a_k)$-system by simply adding an arbitrary $\theta_0 \in T$. Therefore the first inequality in the above result can be applied directly to systems without a reference point $\theta_0$. That is given a collection of $(r_k, l_k)$-systems $\Sigma_k$ for $0 \leq k \leq 0$, we have the inequality

$$\frac{|\{0 \leq j < t : \theta_j \in \bigcup_{k=0}^n \Sigma_k\}|}{t} \leq \sum_{k=0}^n \frac{l_k}{r_k + l_k},$$

for any choice of $\theta_0 \in T$ and $t > 0$.

Proof. If $t < a_k$, then

$$\{0 \leq j < t : \theta_j \in \Sigma_k\} = \emptyset.$$

Therefore suppose that $t \geq a_k$, and partition the interval $[a_k, t]$ into smaller intervals $[t_i, t_{i+1})$, where $a_k = t_0 < \cdots < t_{p_k} \leq t$ are the times such that $\theta_j \in \Sigma_k$ for $t_i \leq j < s_i < t_{i+1}$, and $\theta_j \not\in \Sigma_k$ for $s_i \leq j < t_{i+1}$. Then for every $0 \leq i \leq p_k$,

$$\pi_i = \frac{|\{t_i \leq j < t_{i+1} : \theta_j \in \Sigma_k\}|}{t} \leq \frac{l_k}{t}.$$
Since \( t_{i+1} - t_i \geq r_k + \pi_i t \) and \( t - t_{nk} \geq \pi_{nk} t \), we get the inequality
\[
t - a_k \geq p_k r_k + \sum_{i=0}^{p_k} \pi_i t \geq p_k r_k + (p_k + 1) l_k.
\]

Now, consider the sum
\[
\sum_{i=0}^{p_k} \pi_i = \frac{|\{0 \leq j < t : \theta_j \in \bigcup_{k=0}^{n} \Sigma_k\}|}{t}
\]
We will treat this sum in two different ways. The first one is rewriting
\[
\sum_{i=0}^{p_k} \frac{\pi_i t}{t} \leq \frac{l_k}{a_k + p_k r_k + (p_k + 1) l_k} \leq \frac{l_k + l_k}{r_k}.
\]
The second way proceeds by writing \( m_k = \min\{a_k, r_k\} \) and using the bounds
\[
\sum_{i=0}^{p_k} \frac{\pi_i t}{t} \leq \frac{\sum_{i=0}^{p_k} \pi_i t}{a_k + p_k r_k + \sum_{i=0}^{p_k} \pi_i t} \leq \frac{(p_k + 1) l_k}{(p_k + 1) m_k + (p_k + 1) l_k}.
\]
Doing the same for every \( 0 \leq k \leq n \), and adding them together, we end up with the two inequalities above.

In the same way, one can prove the following.

**Lemma B.2.** Let \( n \geq 0 \) be an integer and \( \theta_0 \in \mathbb{T} \). Suppose that for every \( 0 \leq k \leq n \) we are given a \( \Sigma_k \) such that \( (\Sigma_k, \theta_0) \) is a reversed \( (r_k, l_k, a_k) \)-system. Then for every \( 0 < t \) we have the upper bounds
\[
|\{0 \leq j < t : \theta_j \in \bigcup_{k=0}^{n} \Sigma_k\}| \leq \sum_{k=0}^{n} \frac{l_k}{r_k + l_k}, \tag{B.3}
\]
and
\[
|\{0 \leq j < t : \theta_j \in \bigcup_{k=0}^{n} \Sigma_k\}| \leq \sum_{k=0}^{n} \frac{l_k}{m_k + l_k}, \tag{B.4}
\]
where \( m_k = \min\{a_k, r_k\} \).

### B.2 Growth estimates

In this section, we will assume that our starting points \( (\theta_0, r_0) \) and \( (\theta_0, s_0) \) satisfy that \( r_k \in B = [\lambda^{-2}, \lambda^2] \) for every \( k \in \mathbb{Z} \). That is, we assume that the set \( \mathbb{T} \times B \) is invariant. We recall the other notation in section 4.3, namely \( B' = [\lambda^{-2}, \lambda^{-1}] \) and \( B'' = [\lambda, \lambda^2] \). Moreover, we will assume that \( \lambda \) is sufficiently large for the statements in this section to hold. It will be clear in the proofs where we assume that \( \lambda \) is large.
Lemma B.3. Suppose that we are given a set \( \Sigma \subset T \), a point \( (\theta_0, r_0) \), and a \( t > 0 \) such that
\[
\left| \{ 0 \leq j < t : \theta_j \in \Sigma \} \right| \leq \rho,
\]
for some \( 0 \leq \rho \leq 1 \). If \( \theta_j \not\in \Sigma \implies r_j \in B^u \), for every \( j \in [0, t) \), then
\[
r_0^{\alpha_0} \cdots r_{t-1}^{\alpha_{t-1}} \geq \lambda^{t(1-5\rho)} \tag{B.5}
\]
for every choice of \( \alpha_0, \ldots, \alpha_{t-1} \in [1, 2] \).

Similarly, if \( \theta_j \not\in \Sigma \implies r_j \in B^s \), for every \( j \in [0, t) \), then
\[
r_0^{\alpha_0} \cdots r_{t-1}^{\alpha_{t-1}} \leq \lambda^{-t(1-5\rho)} \tag{B.6}
\]
for every choice of \( \alpha_0, \ldots, \alpha_{t-1} \in [1, 2] \).

Proof.

If \( r_j \in B^u = [\lambda, \lambda^2] \), then for every \( \alpha_j \in [1, 2] \),
\[
\lambda^{\alpha_j} \leq r_j^{\alpha_j} \leq \lambda^{2\alpha_j}.
\]
Moreover, if \( r_j \in B \setminus B^u = [\lambda^{-2}, \lambda] \), then
\[
\lambda^{-2\alpha_j} \leq r_j^{\alpha_j} \leq \lambda^{\alpha_j}.
\]

Therefore
\[
r_0^{\alpha_0} \cdots r_{t-1}^{\alpha_{t-1}} \geq \lambda^{t(1-\rho)} \lambda^{-4\rho} = \lambda^{t(1-5\rho)}.
\]
The proof of the second statement is analogous, noting that if \( r_j \in (\lambda^{-1}, \lambda^2] \), then
\[
\lambda^{-2} \leq r_j^{\alpha_j} \leq \lambda^4,
\]
and that
\[
\lambda^{-4} \leq r_j^{\alpha_j} \leq \lambda^{-1},
\]
if \( r_j \in [\lambda^{-2}, \lambda^{-1}] \).

\( \square \)

Lemma B.4. Suppose that we are given a set \( \Sigma \), a \( \theta_0 \in \mathbb{T} \), and a \( t > 0 \) such that
\[
\left| \{ 0 \leq j < t : \theta_j \in \Sigma \} \right| \leq \rho,
\]
for some \( 0 \leq \rho \leq 1 \). If \( 0 \leq r_j - s_j < \lambda^{-3} \) and \( \theta_j \not\in \Sigma \implies s_j \in B^s \), for every \( j \in [0, t) \), then
\[
r_0 s_0 \cdots r_{t-1} s_{t-1} \leq \lambda^{t(5-\rho)} \tag{B.7}
\]
Similarly, if \( 0 \leq r_j - s_j < \lambda^{-3} \) and \( \theta_j \not\in \Sigma \implies r_j \in B^u \), for every \( j \in (-t, 0] \), then
\[
r_0 s_0 \cdots r_{-t+1} s_{-t+1} \geq \lambda^{t(1-5\rho)} \tag{B.8}
\]
Proof. For every $0 \leq j < t$, the distance $r_j - s_j < \lambda^{-3}$. Therefore, $s_j \in B^i \implies \lambda^{-2} \leq r_j \leq \lambda^{-1} + \lambda^{-3} < 2\lambda^{-1}$. Hence $s_j \in B^i$ implies that $\lambda^{-4} \leq s_j r_j \leq 2\lambda^{-2} < \lambda^{-1}$, if $\lambda$ is sufficiently large.

Therefore, $\theta_j \not\in \Sigma \implies s_j r_j < \lambda^{-1}$, and $\theta_j \in \Sigma \implies s_j r_j \leq \lambda^4$. This implies that

$$r_0 s_0 \cdots r_{t-1} s_{t-1} \leq \lambda^{(5-\rho)}.$$  

The second part is proved in a similar way. \hfill \Box

**Lemma B.5.** Let $\theta_0 \in \mathbb{T}$ and $s_0 < r_0$. If we suppose that $r_0 - s_0 \geq \lambda^{-7}$, then

$$\frac{|\{ j : 0 \leq j < t, r_j \in B^i \}|}{t} \leq \frac{2}{3} + \frac{3}{2t},$$

for every $t > 0$. Moreover, for any $0 \leq m \leq n$, there is a $0 \leq j \leq 10M_m$ such that

$$\theta_j \not\in \Sigma, r_j \in B^u.$$  \hspace{1cm} (B.9)

Proof. First of all, $\lambda^{-2} \leq s_j < r_j < \lambda^2$ for every $j \geq 0$. This means that $r_j \in B^i \implies s_j \in B^i$. In particular $r_j \in B^i \implies \lambda^2 \leq \frac{1}{r_j s_j} \leq \lambda^4$. Set

$$\pi = \frac{|\{ j : 0 \leq j < t, r_j \in B^i \}|}{t}.$$  

Then for every $t > 0$

$$\frac{\lambda^{-7}}{r_0 s_0 \cdots r_{t-1} s_{t-1}} \leq \frac{1}{r_0 s_0 \cdots r_{t-1} s_{t-1}} \cdot |r_0 - s_0| = |r_t - s_t| \leq \lambda^2,$$

and

$$\frac{1}{r_0 s_0 \cdots r_{t-1} s_{t-1}} \leq \lambda^9.$$  

If $r_j \not\in B^i$, that is $r_j \in (\lambda^{-1}, \lambda^2)$, then

$$\lambda^{-4} \leq \frac{1}{r_j s_j} \leq \lambda^3.$$  

Therefore

$$\lambda^{(6\pi - 4)t} = \lambda^{2\pi t} \lambda^{-4(1-\pi)t} \leq \frac{1}{r_0 s_0 \cdots r_{t-1} s_{t-1}} \leq \lambda^9,$$

implying that $(6\pi - 4)t \leq 9$. This yields the inequality

$$\pi \leq \frac{4t + 9}{6t}.$$  

For the second part, note that $\Theta_m = \mathbb{T} \setminus \Xi$, where $\Xi = \bigcup_{j=0}^m \Sigma_j$, and $\Sigma_j = \bigcup_{i=-M_j}^{M_j} I_j + i\omega$. Setting $I_j = 2M_j + 1$ and $r_j = N_j - 2M_j - 1$, **Lemma B.1** applies to $\Xi$, and gives us the bound

$$\frac{|\{ 0 \leq i < t : \theta_i \in \bigcup_{j=0}^m \Sigma_j \}|}{t} \leq \sum_{j=0}^m \frac{2M_j + 1}{t} + \frac{2M_j + 1}{N_j}.$$  

35
Using the estimates for $M_j$ and $N_j$ in the beginning of section 7, we see that choosing $t = 10M_m$ will ensure the inequality
\[
\frac{|\{0 \leq i < 10M_m : \theta_i \in \bigcup_{j=0}^{m} \Sigma_j \}|}{10M_m} \leq \frac{1}{4},
\]
if $\lambda$ is large enough. For the same $t$, we have that $\pi \leq \frac{3}{t}$, which means that the intersection between the sets $|\{0 \leq i < 10M_m : \theta_i \in \Theta_m \}$ and $|\{0 \leq i < 10M_m : r_i \notin B^s \}$ has relative size
\[
\frac{|\{0 \leq i < 10M_m : \theta_i \in \Theta_m, r_i \notin B^s \}|}{10M_m} \geq \frac{7}{20}.
\]
Since we can make the measure of $\Xi_m \cup \Xi_m'$ arbitrarily small, by making $\lambda$ larger, there have to be two successive iterates $\theta_i$ and $\theta_{i+1}$ that are both in $\Theta_m$, and such that both $r_i$ and $r_{i+1}$ are not in $B^s$. Therefore, lemma 4.2 implies that $r_{i+1} \in B^s$, and we are done.

The measure of $\Xi_m \cup \Xi_m'$ can be made arbitrarily small in a uniform manner (the upper bound of the measure can be made independent of $m$), since the measure of the sets $\bigcup_{-M_j}^{M_j} I_j + i\omega$ decreases super-exponentially in $j$.

In a similar way we obtain the following result.

**Lemma B.6.** Let $\theta_0 \in \mathbb{T}$ and $s_0 < r_0$. If we suppose that $r_0 - s_0 \geq \lambda^{-7}$, then
\[
\frac{|\{j : 0 \leq j < t, s - j \in B^s \}|}{t} \leq \frac{2}{3} + \frac{3}{2t},
\]
for every $t > 0$. Moreover, for any $0 \leq m \leq n$, there is a $0 \leq j \leq 10M_m$ such that
\[
\theta_{-j} \in \Theta_m, s - j \in B^s. \tag{B.10}
\]

### C Derivative estimates

Throughout this section, we will assume that $(\theta_0, r_0), (\theta_0, s_0) \in \mathbb{T} \times B$ are such that
\[
r_j, s_j \in B, \tag{C.1}
\]
for every $j \in \mathbb{Z}$. Since orientation is preserved in the fibres, and each fibre takes only strictly positive values,
\[
0 \leq \Pi_{j,k}(s_0, r_0) \leq \Pi_{j,k}(z_0, r_0) \leq 1 \tag{C.2}
\]
if $s_0 \leq z_0 \leq r_0$. Moreover
\[
\frac{\Pi_{j,s}(r_0, s_0)}{\Pi_{j}(r_0, s_0)} = 1 + (r_j + s_j)(r_{j+1} - s_{j+1})\Pi_{j}(r_0, s_0). \tag{C.3}
\]
Since $\partial r_{k+1} = \lambda \nu'(\theta) + \partial \frac{r_k}{r_k}$, it follows that
\[
\partial \theta(r_{k+1} - s_{k+1}) = \frac{\partial \theta r_k}{r_k} - \frac{\partial \theta s_k}{s_k} = \frac{\partial \theta s_k}{s_k} \frac{r_k}{r_k s_k} - \frac{\partial \theta r_k}{r_k s_k} \frac{s_k}{s_k}
\]
\[
= \frac{1}{s_k} \frac{s_k}{r_k s_k} \partial \theta (r_k - s_k) - \frac{\partial \theta s_k}{s_k} \frac{r_k}{s_k}
\]
\[
= D_k(r_0, s_0) \Pi_k(s_0, r_0) \partial \theta (r_k - s_k) + \partial \theta (s_k) \frac{1}{r_k} = (r_k + s_k)(r_{k+1} - s_{k+1}).
\]

36
It follows by induction that
\[ \partial_\theta(r_{k+1} - s_{k+1}) = D_{0,k}(r_0, s_0) \Pi_{0,k}(s_0, r_0) \partial_\theta(r_0 - s_0) + R_{0,k}(r_0, s_0), \]
where the rest term is
\[ R_{0,k}(r_0, s_0) = \partial_\theta(s_k) \cdot \left( \frac{1}{r_k} + \frac{1}{s_k} \right) (r_{k+1} - s_{k+1}) + \sum_{j=0}^{k-1} D_{j+1,k}(r_0, s_0) \Pi_{j+1,k}(s_0, r_0) \partial_\theta(s_j) \left( \frac{1}{r_j} + \frac{1}{s_j} \right) (r_{j+1} - s_{j+1}) = \]
\[ = (r_{k+1} - s_{k+1}) \left[ \partial_\theta(s_k) \left( \frac{1}{r_k} + \frac{1}{s_k} \right) + \sum_{j=0}^{k-1} \Pi_{j,k}(s_0, r_0) \partial_\theta(s_j) \left( \frac{1}{r_j} + \frac{1}{s_j} \right) \right]. \]

Noting that \( D_{0,k}(r_0, s_0) = \frac{r_{k+1} - s_{k+1}}{r_0 - s_0} \), we obtain the expression
\[ \partial_\theta(r_{k+1} - s_{k+1}) = \left( r_{k+1} - s_{k+1} \right) \frac{\partial_\theta(r_0 - s_0)}{r_0 - s_0} \Pi_{0,k}(s_0, r_0) + R_{0,k}(r_0, s_0), \quad (C.4) \]
Since every \( r_i, s_i \in B = [\lambda^{-2}, \lambda^2] \), and \( s_0 \leq r_0 \), it satisfies the inequality
\[ |R_{0,k}(r_0, s_0)| \leq 2\lambda^4 \cdot k \cdot \max_{0 \leq j \leq k} |\partial_\theta(s_j)|. \quad (C.5) \]

Later on, we will see that this is small in comparison to the first term in (C.4). Before we carry on the analysis, let us consider the implications of this. If we disregard the rest term \( R_{0,k} \), we would have
\[ \partial_\theta(r_{k+1} - s_{k+1}) = \left( r_{k+1} - s_{k+1} \right) \frac{\partial_\theta(r_0 - s_0)}{r_0 - s_0} \Pi_{0,k}(s_0, r_0). \]
This is how the derivative at the \( k \)-th step is related to the initial distance. Since \( r_{k+1} - s_{k+1} \) will be related to our stopping time, we may disregard it as essentially constant. The only problem remaining is therefore the factor \( \Pi_{0,k}(s_0, r_0) \). If it is not bounded away from 0 as \( E \not\supset E_0 \), we may lose the constant. Unfortunately, we may not establish a uniform bound. However, for most parameter values, it will be uniformly bounded; and interestingly, for the parameters where the bound fails, the factor \( r_0 - s_0 \) will be dominant in the limit. First, note that we can rewrite
\[ \Pi_j(s_0, r_0) = \frac{1}{1 + \frac{r_j - s_j}{s_j}}. \]
Since \( \frac{r_j - s_j}{s_j} \) is always positive, we have
\[ \Pi_{0,k}(s_0, r_0) = \exp(-\sum_{j=0}^{k} \log(1 + \frac{r_j - s_j}{s_j})). \]
Therefore, in order to obtain a lower bound for \( \Pi_{0,k}(s_0, r_0) \), we need only an upper bound for the expression
\[ \sum_{j=0}^{k} \log(1 + \frac{r_j - s_j}{s_j}) \leq \sum_{j=0}^{k} \frac{r_j - s_j}{s_j} = (r_{k+1} - s_{k+1}) \sum_{j=0}^{k} \frac{1}{D_{j,k}(r_0, s_0) s_j} \leq \lambda^4 \sum_{j=0}^{k} \frac{1}{D_{j,k}(r_0, s_0)}. \]
In conclusion, we have the bounds
\[ \exp(-\lambda^4 \sum_{j=0}^{k} \frac{1}{D_{j,k}(r_0, s_0)}) \leq \Pi_{0,k}(s_0, r_0) \leq 1. \quad (C.6) \]
In order to accomplish that, we need some better control on \( D_{j,k}(r_0, s_0) \). It turns out that \( D_{j,k} \) behaves like a geometric series for most parameters, but can lose the uniformity in the exponent for certain \textit{bad} parameter values.

37
D  More derivative estimates

Recall that \( r_1 = \lambda^2 v(\theta_0) - E - \frac{1}{r_0} \). Therefore

\[
\begin{align*}
\partial_r r_1 &= -1 + \frac{\partial_r r_0}{r_0} \\
\partial^2_r r_1 &= \frac{\partial^2_r r_0}{r_0^2} - 2 \left( \frac{\partial_r r_0}{r_0} \right)^2 \\
\partial_\theta r_1 &= \lambda^2 v'(\theta_0) + \frac{\partial_\theta r_0}{r_0} \\
\partial^2_\theta r_1 &= \lambda^2 v''(\theta_0) + \frac{\partial^2_\theta r_0}{r_0} - 2 \left( \frac{\partial_\theta r_0}{r_0} \right)^2
\end{align*}
\]

By induction we obtain the formulas

\[
\begin{align*}
\partial_r r_{k+1} &= \frac{\partial_r r_0}{r_0 \cdots r_k} - 1 - \sum_{j=1}^{k} \frac{1}{r_j \cdots r_k} \\
\partial^2_r r_{k+1} &= \frac{\partial^2_r r_0}{r_0 \cdots r_k} - 2 \sum_{j=0}^{k} \frac{(\partial_r r_j)^2}{r_j \cdots r_k} \\
\partial_\theta r_{k+1} &= \frac{\partial_\theta r_0}{r_0 \cdots r_k} + \lambda^2 v'(\theta_0) + \lambda^2 \sum_{j=1}^{k-1} \frac{v'(\theta_{j-1})}{r_j \cdots r_k} \\
\partial^2_\theta r_{k+1} &= \frac{\partial^2_\theta r_0}{r_0 \cdots r_k} - 2 \sum_{j=0}^{k} \frac{(\partial_\theta r_j)^2}{r_j \cdots r_k} + \lambda^2 v''(\theta_0) + \lambda^2 \sum_{j=1}^{k-1} \frac{v''(\theta_{j-1})}{r_j \cdots r_k}.
\end{align*}
\]

Note that we can rewrite, for \( 2 \leq j \leq k \),

\[
\frac{(\partial_\theta r_j)^2}{r_j \cdots r_k} = \frac{1}{r_j \cdots r_k} \left( \frac{\partial_\theta r_0}{r_0 \cdots r_{j-1}} + \lambda^2 v'(\theta_{j-1}) + \lambda^2 \sum_{i=1}^{j-1} \frac{v'(\theta_{i-1})}{r_i \cdots r_{j-1}} \right)^2 = \\
= \left( \frac{\partial_\theta r_0}{r_0 \cdots r_j r_{j+1} \cdots r_k} + \frac{\lambda^2 v'(\theta_{j-1})}{r_j \cdots r_k} + \lambda^2 \sum_{i=1}^{j-1} \frac{v'(\theta_{i-1})}{r_i \cdots r_k} \right)^2.
\]

For \( j = 1 \), we have

\[
\frac{(\partial_\theta r_1)^2}{r_1 \cdots r_k} = \frac{1}{r_1 \cdots r_k} \left( \frac{\partial_\theta r_0}{r_0} + \lambda^2 v'(\theta_0) \right)^2 = \\
= \left( \frac{\partial_\theta r_0}{r_0 \cdots r_1^2 r_k} + \frac{\lambda^2 v'(\theta_0)}{r_1^2 r_k} \right)^2.
\]

For \( j > 1 \), we have

\[
\frac{(\partial_\theta r_j)^2}{r_j \cdots r_k} = \frac{1}{r_j \cdots r_k} \left( \frac{\partial_\theta r_0}{r_0 \cdots r_{j-1}} - 1 - \sum_{i=1}^{j-1} \frac{1}{r_i \cdots r_{j-1}} \right)^2 = \\
= \left( \frac{\partial_\theta r_0}{r_0 \cdots r_j r_{j+1} \cdots r_k} - \frac{1}{r_j^{1/2} r_{j+1} \cdots r_k} - \sum_{i=1}^{j-1} \frac{1}{r_i \cdots r_j r_{j+1} \cdots r_k} \right)^2.
\]

38
Lemma D.1. Suppose that \((\theta_0, r_0)\), \(N > 0\) and \(c_1, c_2 > 0\) are such that \(\lambda^{-c_1/2}, \lambda^{-c_2/2} \leq \frac{1}{2}\), and

\[
\frac{1}{r_k \cdots r_N} \leq \lambda^{-c_1(N-k)/2-1},
\]

\[
\frac{1}{r_k \cdots r_{j-1}^{1/2} r_j \cdots r_N} \leq \lambda^{-c_2(N-k)/2-1},
\]

\[
\frac{1}{r_k \cdots r_{j}^{1/2} r_{j+1} \cdots r_N} \leq \lambda^{-c_2(N-k)/2-3/2},
\]

for every \(0 \leq k \leq j \leq N\). If \(|\partial_{r_{N+1}}\theta_0|, |\partial_{r_{N+1}}r_0| \leq \frac{1}{\lambda^2}\) for \(i = 1, 2\), and \(\lambda\) is large enough, then

\[
|\partial_{E} r_{N+1} + 1| \leq \frac{2}{\lambda^2}
\]

\[
|\partial_{E}^2 r_{N+1}| \leq \frac{16}{\lambda^2}
\]

\[
|\partial_{\theta} r_{N+1} - \lambda^2 v'(\theta_N)| \leq 2||v||_{C_1}
\]

\[
|\partial_{\theta}^2 r_{N+1} - \lambda^2 v''(\theta_N)| \leq \frac{16\lambda^2||v||_{C_1}^2}{\lambda^2} + 8\lambda||v||_{C_1}^2 + 2||v||_{C_2}^2.
\]

**Proof.** We immediately obtain the estimates

\[
|\partial_{E} r_{N+1} + 1| \leq \frac{1}{\lambda^{c_1(N+2)}} + \sum_{j=1}^{N} \frac{1}{\lambda^{c_1(N-j)+2}} \leq \lambda^{-2} \sum_{j=0}^{N} \frac{1}{\lambda^{c_1j}} \leq \frac{1}{\lambda^2(1-\lambda^{-c_1})} \leq \frac{2}{\lambda^2}
\]

\[
|\partial_{\theta} r_{N+1} - \lambda^2 v'(\theta_N)| \leq \frac{\lambda^2}{\lambda^{c_1(N+2)}} + \lambda^2 \sum_{j=1}^{N} \frac{|v'(\theta_{j-1})|}{\lambda^{c_1(N-j)+2}} \leq \lambda^2 ||v||_{C_1} \sum_{j=0}^{N} \frac{1}{\lambda^{c_1j}} \leq 2||v||_{C_1}.
\]

In the same way, we estimate

\[
|\frac{(\partial_{r_j})^2}{r_j \cdots r_N^2}| \leq \left( \frac{\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2(N/2+1)}} + \frac{\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2(N-j)/2+1}} + \lambda^2 ||v||_{C_1} \sum_{i=1}^{j-1} \frac{1}{\lambda^{c_2(N-i)/2+1}} \right)^2 =
\]

\[
= \lambda^4 ||v||_{C_1}^2 \left( \sum_{i=0}^{j} \frac{1}{\lambda^{c_2(N-i)/2}} \right)^2 =
\]

\[
= \lambda^4 ||v||_{C_1}^2 \left( \sum_{i=0}^{j} \frac{1}{\lambda^{c_2(N-i)/2}} \right)^2 \leq
\]

\[
\leq 4\lambda^2 ||v||_{C_1}^2 \lambda^{c_2(N-j)},
\]

for \(2 \leq j \leq N-1\). If \(j = 1\), then

\[
|\frac{(\partial_{r_1})^2}{r_1 \cdots r_N^2}| \leq \left( \frac{\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2(N/2+1)}} + \frac{\lambda^2 v'(\theta_1)}{\lambda^{c_2(N-1)/2+1}} \right)^2 \leq \frac{4\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2(N-1)}}.
\]

39
and if $j = 0$, then
\[
\left| \frac{(\partial_\theta r_0)^2}{r_0 \cdot r_0 \cdots r_N^2} \right| = \left( \frac{\partial_\theta r_0}{r_0^{1/2} \cdots r_0 \cdots r_N} \right)^2 \leq \left( \frac{\lambda^2 ||v||_{C_1}}{\lambda^{2N/2+1}} \right)^2 \leq \frac{\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2N}}.
\]

For the case $j = N$, we instead use the estimate
\[
\left| \frac{(\partial_\theta r_N)^2}{r_j \cdot r_j \cdots r_N^2} \right| \leq \left( \frac{\lambda^2 ||v||_{C_1}}{r_j^{1/2} \cdots r_j \cdots r_N} + \lambda^2 \sum_{i=1}^{N-1} \frac{v'(\theta_{i-1})}{r_i^{1/2} \cdots r_{i-1}^{1/2} r_i \cdots r_N} \right)^2 = \lambda^4 ||v||_{C_1}^2 \sum_{i=1}^{N} \left( \frac{1}{\lambda^{c_2(N-i)}} + \frac{1}{\lambda^{3/2}} + \frac{1}{\lambda^{c_2(N-i)+3/2}} \right) = 4\lambda ||v||_{C_1}^2.
\]

Therefore
\[
\left| \sum_{j=0}^{N} \frac{(\partial_\theta r_j)^2}{r_j \cdot r_j \cdots r_N^2} \right| \leq \sum_{j=0}^{N-1} \left| \frac{(\partial_\theta r_j)^2}{r_j \cdot r_j \cdots r_N^2} \right| + \left| \frac{(\partial_\theta r_N)^2}{r_N \cdot r_N} \right| \leq \sum_{j=0}^{N-1} \frac{4\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2(N-j)}} + 4\lambda ||v||_{C_1}^2 \leq \frac{8\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2}} + 4\lambda ||v||_{C_1}^2.
\]

This means that
\[
|\partial_\theta r_{N+1} - \lambda^2 v''(\theta_N)| \leq \frac{|\partial_\theta^2 r_0|}{r_0^{1/2} \cdots r_0 \cdots r_N} + 2 \sum_{j=0}^{N} \left| \frac{(\partial_\theta r_j)^2}{r_j^{1/2} \cdots r_j \cdots r_N} \right| + \left| \lambda^2 \sum_{j=1}^{N} \frac{v''(\theta_{j-1})}{r_j^{1/2} \cdots r_j \cdots r_N} \right| \leq \frac{\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2(N-1)}} + 2 \sum_{j=0}^{N} \left| \frac{(\partial_\theta r_j)^2}{r_j^{1/2} \cdots r_j \cdots r_N} \right| + \left| \lambda^2 \sum_{j=1}^{N} \frac{||v||_{C_1}^2}{r_j^{1/2} \cdots r_j \cdots r_N} \right| \leq \frac{16\lambda^2 ||v||_{C_1}^2}{\lambda^{c_2}} + 8\lambda ||v||_{C_1}^2 + 2||v||_{C_1}^2.
\]

Similarly, for $2 \leq j \leq N$
\[
\left( \frac{\partial_\theta r_j}{r_j \cdot r_j \cdots r_N} \right)^2 = \left( \frac{\partial_\theta r_0}{r_0^{1/2} \cdots r_0 \cdots r_N} - \frac{1}{r_j^{1/2} \cdots r_j \cdots r_N} - \frac{1}{r_j^{1/2} \cdots r_{j-1}^{1/2} r_j \cdots r_N} \right)^2 \leq \left( \sum_{i=1}^{j} \frac{1}{\lambda^{c_2(N-i)/2+1}} \right)^2 \leq \frac{4}{\lambda^{c_2(N-j)+2}},
\]

and for $j = 1$
\[
\left( \frac{\partial_\theta r_1}{r_1 \cdot r_1 \cdots r_N} \right)^2 = \left( \frac{1}{r_1^{1/2} \cdots r_1 \cdots r_N} \right)^2 \leq \frac{1}{\lambda^{c_2(N+2)}}.
\]
Therefore
\[ |\partial^2 E_{rN+1}^k| \leq 2 \sum_{j=1}^{N} \frac{4}{\lambda^{c_1(N-j)+2}} \leq \frac{16}{\lambda^2}. \]

Similarly, one obtains expressions for the derivatives of backward iterates:

\begin{align*}
\partial E_{r-}(k+1) &= (\partial E r_0) r_2^2 \cdots r_{-(k+1)}^2 + \sum_{j=1}^{k+1} r_{-j}^2 \cdots r_{-(k+1)}^2 \tag{D.1} \\
\partial^2 E_{r-}(k+1) &= (\partial^2 E r_0) r_2^2 \cdots r_{-(k+1)}^2 + 2 \sum_{j=1}^{k} \frac{(\partial E r_j)^2}{r_{-j}} r_{-j}^2 \cdots r_{-(k+1)}^2 + 2 \frac{(\partial E r_{-(k+1)})^2}{r_{-k}} + \sum_{j=1}^{k+1} r_{-j}^2 \cdots r_{-(k+1)}^2 \tag{D.2} \\
\partial \theta_{r-}(k+1) &= (\partial \theta r_0) r_2^2 \cdots r_{-(k+1)}^2 - \lambda^2 \sum_{j=1}^{k+1} v'(\theta_{-j})r_{-j}^2 \cdots r_{-(k+1)}^2 \tag{D.3} \\
\partial^2 \theta_{r-}(k+1) &= (\partial^2 \theta r_0) r_2^2 \cdots r_{-(k+1)}^2 + 2 \sum_{j=1}^{k} \frac{\partial \theta r_j}{r_{-j}} r_{-j}^2 \cdots r_{-(k+1)}^2 + 2 \frac{\partial \theta r_{-(k+1)}}{r_{-k}} - \lambda^2 \sum_{j=1}^{k+1} v''(\theta_{-j})r_{-j}^2 \cdots r_{-(k+1)}^2 \tag{D.4} \\
\partial E^k_{r-}(k+1) &= (\partial E^k r_0) r_2^2 \cdots r_{-(k+1)}^2 + 2 \sum_{j=1}^{k} \frac{(\partial E^k r_j)^2}{r_{-j}} r_{-j}^2 \cdots r_{-(k+1)}^2 + 2 \frac{(\partial E^k r_{-(k+1)})^2}{r_{-k}} - \lambda^2 \sum_{j=1}^{k+1} v''(\theta_{-j})r_{-j}^2 \cdots r_{-(k+1)}^2 \tag{D.5}
\end{align*}

The proof of the next lemma proceeds analogously to the proof of the previous lemma.

**Lemma D.2.** Suppose that \((\theta_0, r_0), N > 0\) and \(c_1, c_2 > 0\) are such that \(\lambda^{-c_1/2}, \lambda^{-c_2/2} \leq \frac{1}{2}\), and

\[
\begin{align*}
r_{-k} \cdots r_{-N} &\leq \lambda^{-c_1(N-k)/2-1}, \\
r^2_{-k} \cdots r^2_{-(j+1)} r^{1/2}_{-j} r_{-j} \cdots r_{-N} &\leq \lambda^{-c_2(N-k)/2-1}, \\
r^2_{-k} \cdots r^2_{-(j+1)} r^{1/2}_{-j} r_{-j} \cdots r^2_{-N} &\leq \lambda^{-c_2(N-k)/2-3/2},
\end{align*}
\]

for every \(0 \leq k \leq j \leq N\). If \(|\frac{\partial E r_0}{r_0 \cdots r_{-N}}|, \frac{\partial^2 E r_0}{r_0 \cdots r_{-N}}| \leq \frac{1}{\lambda^2}\) for \(i = 1, 2\), and \(\lambda\) is large enough, then

\[
\begin{align*}
|\partial E_{r-}| &\leq \frac{2}{\lambda^2} \\
|\partial^2 E_{r-}| &\leq \frac{16}{\lambda^2} \\
|\partial \theta_{r-}| &\leq 2\|v\|_{c_1} \\
|\partial^2 \theta_{r-}| &\leq \frac{16\lambda^2\|v\|_{c_1}^2}{\lambda^{c_2}} + 8\lambda\|v\|_{c_1}^2 + 2\|v\|_{c_2}.
\end{align*}
\]

**References**

[BJ02] J. Bourgain and S. Jitomirskaya. Continuity of the lyapunov exponent for quasiperiodic operators with analytic potential. *Journal of Statistical Physics*, 108(5):1203–1218, Sep 2002.
[Bje07] K. Bjerklöv. Dynamics of the quasi-periodic schrödinger cocycle at the lowest energy in the spectrum. *Commun. Math. Phys.*, 2007.

[Bou00] J. Bourgain. Hölder regularity of integrated density of states for the almost mathieu operator in a perturbative regime. *Letters in Mathematical Physics*, 51(2):83–118, 2000.

[BS08] Kristian Bjerklöv and Maria Saprykina. Universal asymptotics in hyperbolicity breakdown. *Nonlinearity*, 21:557–586, 2008.

[Dam17] David Damanik. Schrödinger operators with dynamically defined potentials. *Ergodic Theory and Dynamical Systems*, 37(6):1681–1764, 2017.

[FGJ17] Gabriel Fuhrmann, Maik Gröger, and Tobias Jäger. Non-smooth saddle-node bifurcations ii: dimensions of strange attractors. *Ergodic Theory and Dynamical Systems*, pages 1–23, 2017.

[FH15] Jordi-Lluís Figueras and Àlex Haro. Different scenarios for hyperbolicity breakdown in quasiperiodic area preserving twist maps. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(12):123119, 2015.

[GS01] Michael Goldstein and Wilhelm Schlag. Hölder continuity of the integrated density of states for quasi-periodic schrödinger equations and averages of shifts of subharmonic functions. *Annals of Mathematics*, 154(1):155–203, 2001.

[HdlL06] A Haro and R de la Llave. Manifolds on the verge of a hyperbolicity breakdown. *Chaos*, 16, 2006.

[Her83] Michael-R Herman. Une méthode pour minorer les exposants de lyapounov et quelques exemples montrant le caractère local d’un théorème d’arnold et de moser sur le tore de dimension 2. *Comment. Math. Helv.*, 58:453—502, 1983.

[OT17] Thomas Ohlson Timoudas. Power law asymptotics in the creation of strange attractors in the quasi-periodically forced quadratic family. *Nonlinearity*, 30(12):4483, 2017.

[Sta97] Jaroslav Stark. Invariant graphs for forced systems. *Physica D*, 109:163–179, 1997.

[WY13] Yiqian Wang and Jiangong You. Examples of discontinuity of lyapunov exponent in smooth quasiperiodic cocycles. *Duke Math. J.*, 162(13):2363–2412, 10 2013.

[You97] L.-S. Young. Lyapunov exponents for some quasi-periodic cocycles. *Ergodic Theory and Dynamical Systems*, 17(2):483–504, 1997.