AN APPLICATION OF GROUP EXPANSION TO THE ANDERSON-BERNOULLI MODEL

J. BOURGAIN

ABSTRACT. We establish smoothness of the density of states for 1D lattice Schrödinger operators with potential taking values $\pm \lambda$, for $\lambda$ in a class of small algebraic numbers and energy $E \in (-2,2)$ suitably restricted away from $\pm 2$.

0. Introduction

Let $H = \Delta + \lambda V$, where $\Delta$ is the lattice Laplacian on $\mathbb{Z}$ and $V_z = (V_n)_{n \in \mathbb{Z}}$ are independent variables in $\{1,-1\}$. The spectral theory of this operator, referred to as the Anderson-Bernoulli model (A-B for short) has been studied by various authors. It was shown by Halperin [S-T] that for fixed $\lambda$, the integrated density of states (IDS) $\mathcal{N}(E)$ of $H$ is not Hölder continuous of any order $\alpha$ larger than

$$\alpha_0 = \frac{2 \log 2}{\arccosh (1 + \lambda)}.$$  \hspace{0.5cm} (0.1)

Hölder regularity for some $\alpha > 0$ has been established in several papers.

In [Ca-K-M], le Page’s method is used. Different approaches (including the super-symmetric formalism) appear in the paper [S-V-W] that relies on harmonic analysis principles around the uncertainty principle. Recently [B1], the author showed that $\mathcal{N}(E)$ restricted to $\delta < |E| < 2 - \delta$ ($\delta > 0$ fixed) is at least Hölder-regular of exponent $\alpha(\lambda) \to 0$.

It is believed that in fact for $\lambda \to 0$, $\mathcal{N}(E)$ becomes arbitrarily smooth and in particular $\frac{d\mathcal{N}(E)}{dE}$ is bounded for $|\lambda|$ small enough. No result of this type for the A-B model seems presently known. Recall also Thouless formula relating $\mathcal{N}(E)$ with the Lyapounov exponent $L(E)$ of $H$, i.e.

$$L(E) = \int \log |E - E'| d\mathcal{N}(E').$$  \hspace{0.5cm} (0.2)
Since $\mathcal{N}(E)$ is obtained as the Hilbert transform of $L(E)$, their regularity properties may be derived from each other.

The purpose of this Note is to prove the following in support of the above conjecture.

**Theorem.** Let $H_\lambda$ be the A-B model considered above and restrict $|E| < 2 - \delta$ for some fixed $\delta > 0$. Given a constant $C > 0$ and $k \in \mathbb{Z}_+$, there is some $\lambda_0 = \lambda_0(C, k) > 0$ such that $\mathcal{N}(E)$ is $C^k$-smooth on $]-2 + \delta, 2 - \delta[$ provided $\lambda$ satisfies the following conditions

1. $|\lambda| < \lambda_0$
2. $\lambda$ is an algebraic number of degree $d < C$ and minimal polynomial $P_d(x) \in \mathbb{Z}[X]$ with coefficients bounded by $(\frac{1}{\lambda})^C$
3. $\lambda$ has a conjugate $\lambda'$ of modulus $|\lambda'| \geq 1$

This seems in particular to be the first statement of Lipschitz behavior of the IDS for an A-B model. Several comments are in order. Firstly, the arithmetic assumptions on $\lambda$ permit to exploit a spectral gap theorem for the projective action $\rho$ of $SL_2(\mathbb{R})$ on $P_1(\mathbb{R})$ that was established in [B-Y] and which is our main tool (cf. also the application in [B2] of the latter result to regularity of Furstenberg measures). This spectral gap property is not a consequence of hyperbolicity but is obtained by an adaptation to $SL_2(\mathbb{R})$ of the arguments from [B-G] on spectral gaps in $SU(2)$, established by methods from arithmetic combinatorics (we will not elaborate on these aspects here; see also §4). In its abstract setting, the result from [B-Y] may be formulated as follows. We identify $P_1(\mathbb{R})$ with the torus $T = \mathbb{R}/\mathbb{Z}$.

**Proposition 1. [B-Y].**

Given a constant $0 < c < 1$, there is $R_0 \in \mathbb{Z}_+$ such that the following holds. Let $R > R_0$ and $\mathcal{G} \subset SL_2(\mathbb{R}), |\mathcal{G}| = R$ generating freely the free group $F_R$ on $R$ generators. Assume moreover

1. $\|g - e\| < R^{-c}$ for $g \in \mathcal{G}$
2. $\mathcal{G}$ satisfies the following ‘non commutative diophantine condition’. Denote $W_\ell(\mathcal{G}) \subset SL_2(\mathbb{R})$ the set of words of length at most $\ell$ written in the $\mathcal{G}$-elements. Then, for all $\ell \in \mathbb{Z}_+$

$$\|g - e\| > R^{-\ell/c} \text{ for } g \in W_\ell(\mathcal{G}) \setminus \{e\}.$$
Then there is a finite dimensional subspace $V$ of $L^2(T)$, that may be taken

$$V = [e(n\theta); |n| < K] \quad (e(n\theta) = e^{2\pi i n\theta})$$

where $K = K(R) \in \mathbb{Z}$ large enough, such that if $f \in L^2(T)$, $\|f\|_2 = 1$ and $f \perp V$, then

$$\left\| \frac{1}{2R} \sum_{g \in G} (\rho_g f + \rho_{g^{-1}} f) \right\|_2 < \frac{1}{2}. \quad (0.8)$$

In the construction from [B-Y], the elements of $G$ have rational entries, more precisely, $G \subset SL_2(\mathbb{R}) \cap \frac{1}{Q} \text{Mat}_{2 \times 2}(\mathbb{Z})$ with $Q \in \mathbb{Z}_+$ satisfying

$$Q^c < |G| < R < Q. \quad (0.9)$$

Obviously $\|g - e\| \geq Q^{-\ell}$ for $g \in W_\ell(G) \{e\}$ and in this way we obtain condition (0.7). In the application in this paper, $G$ will consist of algebraic elements of bounded degree $d < C$ and height bounded by $R^C$. The required diophantine condition follows then from [G-J-S], Proposition 4.3, again invoking simple arithmetic considerations. Presently, the [G-J-S] argument seems the only known one to establish such non-commutative DC and it is a major problem in this area of group expansion to treat non-algebraic generators. This explains why in (0.4), $\lambda$ was assumed algebraic. Let us next explain assumption (0.5), which in some sense is the novel input. Denote for a fixed $E \in ]-2 + \delta, 2 - \delta[$

$$g_+ = \begin{pmatrix} E + \lambda & -1 \\ 1 & 0 \end{pmatrix} \quad g_- = \begin{pmatrix} E - \lambda & -1 \\ 1 & 0 \end{pmatrix}. \quad (0.10)$$

Clearly

$$h_1 = g_+ g_-^{-1} = \begin{pmatrix} 1 & 2\lambda \\ 0 & 1 \end{pmatrix}$$

$$h_2 = g_+^{-1} g_- = \begin{pmatrix} 1 & 0 \\ 2\lambda & 1 \end{pmatrix}. \quad (0.11)$$

We use the following result due to Brenner [Br].

Proposition 2. ([Br]).
If \( \mu \in \mathbb{R}, |\mu| \geq 2 \), then the group generated by the parabolic elements

\[
A = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ \mu & 1 \end{pmatrix}
\]

is free.

As pointed out in [L-U], the same conclusion holds if \( \mu \) is an algebraic number with an algebraic conjugate \( \mu' \) such that \( |\mu'| \geq 2 \). Hence, if \( \lambda \) satisfies (0.5), the elements \( h_1, h_2 \) defined in (0.11) will generate a free group. The set \( \mathcal{G} \) in Proposition 1 is then obtained by considering elements \( h_1^r h_2^r \), \( r = 1, \ldots, R \). Using Proposition 1, we prove that

\[
\|f - \rho g_+ f\|_2 + \|f - \rho g_- f\|_2 > \frac{1}{8} \lambda^\tau \tag{0.12}
\]

if \( f \in L^2(\mathbb{T}), \|f\|_2 = 1, f \in V^\perp \).

Here \( \tau > 0 \) is arbitrary and fixed, \( |\lambda| \) taken sufficiently small depending on \( \tau \) (for our purpose, \( \tau < \frac{1}{2} \) will do). Note that the inequality (0.12), restricted to \( f \in V^\perp, \|f\|_2 = 1 \), is considerably stronger than the general inequality (cf. [S-V-W], Theorem 4.1)

\[
\|f - \rho g_+ f\|_2 + \|f - f g_- f\|_2 > c|\lambda| \tag{0.13}
\]

if \( f \in L^2(\mathbb{T}), \|f\|_2 = 1 \).

From (0.12), we derive a restricted spectral gap for the operator

\[
\frac{1}{4}(I + \rho g_+ + \rho g_-) \quad \text{i.e.}
\]

\[
\left\| \frac{1}{3}(f + \rho g_+ f + \rho g_- f) \right\|_2 \leq (1 - c \lambda^{2\tau}) \|f\|_2 \quad \text{for} \quad f \in V^\perp \tag{0.14}
\]

and (0.14) is then processed further to derive certain smoothing estimates for the convolution powers (cf. [B2]), from which eventually the regularity of the Lyapounov exponent is derived.

Some comments about the energy restriction \( |E| < 2 - \delta \). At some stage of our analysis, we make use of the Figotin-Pastur transformation, setting

\[
E = 2 \cos \kappa \quad (0 < \kappa < \pi) \tag{0.15}
\]

and conjugating the cocycle by the matrix

\[
S = \frac{1}{(\sin \kappa)^{\frac{\tau}{2}}} \begin{pmatrix} 1 & -\cos \kappa \\ 0 & \sin \kappa \end{pmatrix}. \tag{0.16}
\]
This gives
\[ Sg\pm S^{-1} = \begin{pmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{pmatrix} \pm \lambda \begin{pmatrix} 1 & \frac{\cos \kappa}{\sin \kappa} \\ 0 & 0 \end{pmatrix} \] (0.17)
which for small \( \lambda \) are perturbations of a rotation. We did not explore here how to handle the edges of the spectrum.

Finally, let us point out that while \( \lambda \) is taken small, we do not let \( \lambda \to 0 \) in the above Theorem and the regularity estimates on \( N(E) \) degenerate in the limit \( \lambda \to 0 \).

1. A spectral gap estimate

In this section, we prove the following

**Proposition 3.** Fix constants \( C > 1, 0 < \tau < \frac{1}{2} \). Let \( \lambda \) be an algebraic number of degree \( d < C \) and with minimal polynomial \( P_d(x) = \sum_{j=0}^{d} a_j x^j \in \mathbb{Z}[X] \). Assume

\[
(1.1) \quad |\lambda|, \lambda_0 = \lambda_0(C, \tau) < \frac{1}{10} \\
(1.2) \quad H = \max |a_j| < \left( \frac{1}{2} \right)^C \\
(1.3) \quad \lambda \) has an algebraic conjugate \( \lambda' \) with \( |\lambda'| \geq 2 \).
\]

Denote
\[
h_1 = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad h_2 = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}
\]
and let \( \rho \) be the projective representation of \( SL_2(\mathbb{R}) \), acting on \( L^2(\mathbb{T}) \). There is a finite dimensional space \( V = [e(n\theta); |n| < K] \), where \( K = K(\lambda) \), such that if \( f \in L^2(\mathbb{T}), \|f\|_2 = 1 \) and \( f \perp V \), then
\[
\|f - \rho h_1 f\|_2 + \|f - \rho h_2 f\|_2 > \frac{1}{4} \lambda^\tau. \quad (1.4)
\]

By (0.11), Proposition 3 implies (0.12) for \( \lambda \) satisfying assumption (0.5) of the Theorem.

**Proof of Proposition 3.**

The argument relies on Proposition 1 and 2 stated in Section 0.

Let \( f \) be as above (with \( K \) to be specified) and assume
\[
\|f - \rho h_1 f\|_2 < \varepsilon_0, \|f - \rho h_1 f\|_2 < \varepsilon_0. \quad (1.5)
\]
Denoting $W_{\ell}(h_1, h_2)$ the words of length at most $\ell$ written in $h_1, h_2$ and their inverses, it follows from (1.5) that
\[
\|f - \rho_g f\|_2 < \ell \varepsilon_0 \quad \text{for} \quad g \in W_{\ell}(h_1, h_2).
\] (1.6)

By Proposition 2 and (1.3), $h_1, h_2$ are generators of the free group $F_2$. Let
\[
R = \lfloor |\lambda|^{-\tau} \rfloor \quad \text{(1.7)}
\]
and define for $r = 1, \ldots, R$
\[
g_r = h_1^r h_2^r = \begin{pmatrix} 1 & r\lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r\lambda & 1 \end{pmatrix}.
\] (1.8)

Then $G = \{g_1, \ldots, g_R\}$ are free generators of $F_R$ and clearly satisfy
\[
\|1 - g\| < \lambda^{\frac{1}{2}} \quad \text{for} \quad g \in G.
\] (1.9)

In order to apply Proposition 1, we need to verify the DC (0.7). This is basically Proposition 4.3 from [G-J-S], but we recall the argument since the quantitative aspects of the estimate matter here.

Take $N \in \mathbb{Z}_+$, $N \leq H$ such that $N\lambda = \mu \in \mathcal{O} = \mathcal{O}_{\mathbb{Q}(\lambda)}$ (the integers of the number field $\mathbb{Q}(\lambda)$). If $w \in W_{\ell}(G)$, the entries of $w - 1$ are, by (1.8), of the form $f(\lambda)$ with $f(x) \in \mathbb{Z}[X]$ of degree $D \leq 2\ell$ and coefficients bounded by $(2 + R)^{2\ell}$. Let $\lambda = \lambda_1, \lambda_2, \ldots, \lambda_d$ be the conjugates of $\lambda$ and set $\mu_j = N\lambda_j \ (1 \leq j \leq d)$ which are the conjugates of $\mu$. Thus $N^D f(\lambda_j) = f_1(\mu_j)$ where $f_1(X) = N^d f(\frac{X}{N}) \in \mathbb{Z}[X]$. Assuming $f(\lambda) \neq 0$, it follows that $\prod_{j=1}^d f_1(\mu_j) \in \mathbb{Z} \setminus \{0\}$ and hence
\[
|f_1(\mu)| \geq N^{-(d-1)D} \prod_{j=2}^d |f(\lambda_j)|^{-1}. \quad \text{(1.10)}
\]

Since $|\lambda_j| \leq H + 1, |f(\lambda_j)| \leq (2 + R)^{2\ell} (H + 1)^{2\ell}$ and by (1.10), (1.7), (1.2)
\[
\|w - 1\| \geq |f(\lambda)| \geq N^{-dD}[(2 + R)(1 + H)]^{-2\ell(d-1)} > R^{-4(C+1)d\ell} = R^{-C'\ell}
\]

Taking $|\lambda| < \lambda_0(C, \tau)$, we get $R > R_0$ and the conclusion of Proposition 1 applies with some $K$ depending on the size of $\lambda$.

From (0.8), it follows in particular that for some $g \in G \subset W_{2R}(h_1, h_2)$
\[
\frac{1}{2} < \|f - \rho_g f\|_2 < 2R\varepsilon_0
\]
implying (1.4). This proves Proposition 3. □

In the sequel, we will use (0.12) for some fixed \( \tau < \frac{1}{2} \).

2. Smoothing estimates

For \( g \in SL_2(\mathbb{R}) \), denote by \( \tau_g \) the action on \( \mathcal{P}_1(\mathbb{R}) \), identified with the circle \( \mathbb{T} = \mathbb{R}/\mathbb{Z} \). Thus if \( g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \), \( ad - bc = 1 \), then

\[
e^{i\tau_g(\theta)} = \frac{(a \cos \theta + b \sin \theta) + i(c \cos \theta + d \sin \theta)}{[(a \cos \theta + b \sin \theta)^2 + (c \cos \theta + d \sin \theta)^2]^{1/2}}
\]

and \( \rho_g f = (\tau_{g^{-1}}')^{1/2} (f \circ \tau_g) \). Recall that

\[
\tau_g'(\theta) = \frac{\sin^2 \tau_g(\theta)}{(c \cos \theta + d \sin \theta)^2} = \frac{1}{[a \cos \theta + b \sin \theta]^2 + (c \cos \theta + d \sin \theta)^2}
\]

hence

\[
\|g\|^2 \leq \tau_g' \leq \|g\|^2 \quad \text{and} \quad |\tau_g^{(s)}| \leq c_s \|g\|^{2s} \text{ for } s \in \mathbb{Z}_+.
\]

Assume \( |E| < 2 - \delta \) and perform the Figotin-Pastur transformation (0.15)-(0.17) denoting \( \tilde{g}_\pm = Sg_\pm S^{-1} \). Since \( \rho_{\tilde{g}} = \rho_S \rho_{g_\pm} \rho_{S^{-1}} \), it follows from (0.12) that

\[
\|f - \rho_{\tilde{g}_+} f\|_2 + \|f - \rho_{\tilde{g}_-} f\|_2 > \frac{1}{8} \lambda^\tau
\]

provided \( \|f\|_2 = 1 \). Since \( \tau_S \) acts on \( \mathbb{T} \) as a smooth diffeomorphism, the space \( V \) may clearly be redefined as to ensure that (2.4) holds for \( f \in V^\perp, \|f\|_2 = 1 \). Observe also that by (0.17) and our assumption \( |E| < 2 - \delta \), \( \delta \) fixed, \( \tilde{g}_\pm \) are \( O(\lambda) \) perturbations of a circle rotation. Hence, by (2.2)

\[
\|\tilde{g}_\pm\| < 1 + C\lambda
\]

\[
\tau_{\tilde{g}_\pm}' = 1 + O(\lambda).
\]

Denoting

\[
\tilde{T}_1 = \frac{1}{3}(I + \rho_{(\tilde{g}_+)^{-1}} + \rho_{(\tilde{g}_-)^{-1}})
\]

(2.4) implies that

\[
\|\tilde{T}_1 f\|_2 < 1 - \frac{1}{2300} \lambda^{2\tau} \text{ if } f \in V^\perp, \|f\|_2 = 1.
\]
Since \( \rho(\tilde{g} \pm)^{-1} f = ((\tau_{\tilde{g}})^{\pm})^{1/2} (f \circ \tau_{\tilde{g}}) \), (2.6) clearly implies (assuming \( \lambda \) small enough)

\[
\| T f \|_2 \leq \left( 1 - \frac{1}{2301} \right)^{2r} \| f \|_2 \text{ for } f \in V^\perp
\]

(2.9)

where \( V = \{ e(n\theta) ; |n| < K \} \) and we defined

\[
\tilde{T} f = \frac{1}{3} \left( f + (f \circ \tau_{\tilde{g}}) + (f \circ \tau_{\tilde{g}}) \right).
\]

(2.10)

For simplicity, we drop the \( \sim \) notation in the next considerations.

Our next goal is to deduce from the contractive estimate (2.9) further bounds on \( T^m \) acting on various spaces. Note that obviously

\[
\| T^m f \|_\infty \leq \| f \|_\infty.
\]

(2.11)

Let \( g \in W_t(g_+, g_-), n \in \mathbb{Z}, n' \in \mathbb{Z}_* \). By change of variable and partial integration, we obtain

\[
\left| \int e(n' \tau_g(x) + nx) dx \right| = \left| \int e(n' y + n \tau_{g^{-1}}(y)) \tau_{g^{-1}}(y) dy \right|
\]

\[
\ll_{r} \frac{1}{|n'|^r} \| e(n \tau_{g^{-1}})^{r} \tau_{g^{-1}} \|_{C^r}
\]

\[
\ll_{r} \frac{1}{|n'|^r} (|n|^r \| g \|^{2(r+1)}) \quad \text{(by (2.3))}
\]

\[
\ll_{r} \frac{|n|^r}{|n'|^r} (1 + C|\lambda|)^2(r+1)\ell
\]

(2.12)

since \( \| g \| < (1 + C\lambda)^\ell \) from (2.5).

**Lemma 1.**

\[
\| T^m f \|_2 \leq C(\lambda) \| f \|_2.
\]

(2.13)

**Proof.** Denote \( P_K \) the orthogonal (= Fourier) projection on \( V \) and decompose \( f = f^{(1)} + f^{(2)}, f^{(1)} = P_K f, f^{(2)} \perp V \).

Thus

\[
\| f^{(1)} \|_\infty \leq \sqrt{2K} \| f \|_2 \text{ and } \| f^{(2)} \|_2 \leq \| f \|_2
\]

and

\[
\| T^m f \|_2 \leq \| T^m f^{(1)} \|_2 + \| T^m f^{(2)} \|_2
\]

\[
\leq \| T^m f^{(1)} \|_\infty + \| T^{m-1} f_1 \|_2 \quad (f_1 = T f^{(2)})
\]

\[
\leq \| f^{(1)} \|_\infty + \| T^{m-1} f_1 \|_2 \quad \text{(by (2.11))}
\]

\[
\leq \sqrt{2K} \| f \|_2 + \| T^{m-1} f_1 \|_2
\]

(2.14)
where, by (2.9),
$$
\|f_1\|_2 \leq (1 - c\lambda^{2r})\|f^{(2)}\|_2 \leq (1 - c\lambda^{2r})\|f\|_2.
$$

Repeat (2.14) with $f$ replaced by $f_1$ and iterate to get
$$
\|T^m f\|_2 \lesssim \sqrt{2K\lambda^{-2r}}\|f\|_2
$$
proving (2.13). \hfill \square

There is the following refinement of Lemma 1.

**Lemma 2.** Let $\text{supp} \hat{f} \cap [-2^k, 2^k] = \phi$ with $k > k(\lambda)$.

Then
$$
\|T^m f\|_2 \leq C(\lambda)e^{-\min(c\lambda^{2r}m, r)\|f\|_2} \quad (2.15)
$$
for any given $r \geq 1$ (assuming $\lambda$ small enough).

**Proof.** In view of Lemma 1, it suffices to establish (2.15) for $m < C\lambda^{-2r}r$. Set $F_m = T^m f$ and decompose $F_m = P_k F_m + (F_m - P_k F_m) = F_m^{(1)} + F_m^{(2)}$. Then, using (2.12)
$$
|\hat{F}_m(n)| \leq \max_{g \in W_m} \sum_{|n'| > 2^k} |\hat{f}(n')| |e^{(n' \tau_g)}(n)|
$$
$$
\ll_r |n|^r e^{C|\lambda|\rho m} \sum_{|n'| > 2^k} |\hat{f}(n')| |n'|^{-r}
$$
$$
\ll_r |n|^r e^{Cr|\lambda|\rho} 2^{-k(r-\frac{1}{2})}\|f\|_2
$$
$$
\ll_r |n|^r \left(e^{Cr|\lambda|2r-2} \frac{1}{\sqrt{2}}\right)^r \|f\|_2 < |n|^r e^{-\frac{1}{10}r} \|f\|_2
$$
\quad (2.16)

by the assumption on $m$ and $\lambda$ sufficiently small ($\tau < \frac{1}{2}$).

Thus
$$
\|F_m^{(1)}\|_\infty \leq \sqrt{2K} \|F_m^{(1)}\|_2 \leq CK^{r+1} e^{-\frac{1}{10}rk}\|f\|_2.
$$

Estimate
$$
\|F_{m+1}\|_2 \leq \|TF_m^{(1)}\|_\infty + \|TF_m^{(2)}\|_2
$$
$$
\leq CK^{r+1} e^{-\frac{1}{10}rk}\|f\|_2 + (1 - c\lambda^{2r})\|F_m\|_2 \quad (2.17)
$$
where we used again (2.9).
Iteration of (2.17) with \(m < Cr\lambda^{-2\tau k}\) gives
\[
\|F_m\|_2 \leq \left[Cr(\lambda)e^{-\frac{1}{2}\tau k} + e^{-c\lambda^{2\tau}m}\right]\|f\|_2.
\]
This proves (2.15). \(\square\)

Next, we establish bounds on higher Sobolev norms.

**Lemma 3.** For \(s \in \mathbb{Z}_+, |\lambda| < \lambda(s)\), we have for \(f \in H^s(\mathbb{T})\)
\[
\|T^m f\|_{H^s} \leq C(\lambda)\|f\|_2 + e^{-c\lambda^{2s}m}\|f\|_{H^s}.
\](2.18)

In particular
\[
\|T^m f\|_{H^s} \leq C\|f\|_{H^s}.
\]

**Proof.** Apply Lemma 2 with \(m = m_0(\lambda)\) to specify, \(K_1 = 2^{m_0}\), to obtain
\[
\|T^{m_0} (I - P_{K_1})\|_{2 \to 2} \leq C(\lambda)e^{-c\lambda^{2s}m_0}
\](2.20)
while on the other hand for \(s \in \mathbb{Z}_+\)
\[
\|T^{m_0} (I - P_{K_1})\|_{H^s \to H^s} \leq \|T^{m_0}\|_{H^s \to H^s} < C_s \max_{g \in W_{m_0}} \|g\|^{2s} < C_s e^{C \lambda s m_0}.
\](2.21)

Assuming \(\lambda\) sufficiently small and taking \(m_0 = m_0(\lambda, s)\), interpolation between (2.20), (2.21) will imply that
\[
\|T^{m_0} (I - P_{K_1})\|_{H^s \to H^s} < \frac{1}{10}.
\](2.22)
Set \(F_m = T^m f\). Then
\[
\|F_{m+m_0}\|_{H^s} \leq \|T^{m_0} P_{K_1} F_m\|_{H^s} + \|T^{m_0} (I - P_{K_1}) F_m\|_{H^s}
\]
\[
\leq C(\lambda)K_1\|F_m\|_2 + \frac{1}{10}\|F_m\|_{H^s}
\]
\[
\leq C(\lambda)\|f\|_2 + \frac{1}{10}\|F_m\|_{H^s}.
\](2.23)
Iteration of (2.23) implies (2.18). \(\square\)

Lemmas 1, 2, 3 hold for \(\tilde{T}\) defined in (2.10). If we define now \(T\) by
\[
Tf = \frac{1}{3}(f + (f \circ \tau_{g_+}) + (f \circ \tau_{g_-}))
\](2.24)
clearly $T$ and $\tilde{T}$ are related by

$$\tilde{T} f = (T(f \circ \tau_S)) \circ \tau_{S^{-1}}$$

with $S$ given by (0.16). Thus $\tau_S$ intertwines $T^m$ and $(\tilde{T})^m$, Lemma 3 remains valid for the original $T$ given by (2.24).

Let $\mu$ be the probability measure on $SL_2(\mathbb{R})$ defined by

$$\mu = \frac{1}{2}(\delta_{g^+} + \delta_{g^-}). \quad (2.25)$$

The Furstenberg measure $\nu$ is the (unique) $\mu$-stationary measure on $P_1(\mathbb{R}) \simeq T$, i.e. satisfying

$$\nu = \sum_g (\tau_g)_* [\nu] \mu(g). \quad (2.26)$$

For $f \in C^1(T)$, one has large deviation inequalities (cf. [B-L]) of the form

$$\left\| \sum_g (f \circ \tau_g) \mu^{(\ell)}(g) - \int f \, d\nu \right\|_\infty \leq C e^{-c(\lambda)\ell} \|f\|_{C^1}. \quad (2.27)$$

Since

$$T = \frac{1}{3} I + \frac{2}{3} \sum \tau_g \mu(g)$$

$$T^\ell = 3^{-\ell} \sum_{m=0}^{\ell} \binom{\ell}{m} 2^m \left( \sum (\tau_g)_* \mu^{(m)}(g) \right), \quad (2.28)$$

Combined with (2.27), this gives

**Lemma 4.**

$$\|T^\ell f - \int f \, d\nu\|_\infty \leq C(\lambda)e^{-c(\lambda)\ell} \|f\|_{C^1}. \quad (2.29)$$

**Proof.** L.h.s. of (2.29) is bounded by

$$C\|f\|_{C^1} 3^{-\ell} \sum_{m=0}^{\ell} \binom{\ell}{m} 2^m e^{-c(\lambda)m} \leq C\|f\|_{C^1} \left( \frac{2}{3} + \frac{1}{3} e^{-c(\lambda)} \right)^\ell. \quad \square$$

**Lemma 5.** For $s \geq 1$ and $f \in H^{s+1}$

$$\|(T^\ell f)'\|_{H^s} \leq C(\lambda)e^{-c(\lambda)\ell} \|f\|_{H^{s+1}}.$$
Proof. Choose some $\ell_1 < \ell$ and write

$$
\| (T^\ell f)' \|_{H^s} \leq \| T^\ell f - \int f \, d\nu \|_{H^{s+1}}
\leq \| T^{\ell_1} (T^{\ell-\ell_1} f - \int f \, d\nu) \|_{H^{s+1}}
\leq C(\lambda) \| T^{\ell-\ell_1} f - \int f \, d\nu \|_2 + e^{-c(\lambda)\ell_1} \| T^{\ell-\ell_1} f \|_{H^{s+1}} \quad \text{(by Lemma 3)}
\leq C(\lambda) e^{-c(\lambda)(\ell-\ell_1)} \| f \|_{C^1} + C(\lambda) e^{-c(\lambda)\ell_1} \| f \|_{H^{s+1}} \quad \text{(by Lemmas 4, 3)},
$$

and (2.30) follows by taking $\ell_1 \sim \frac{\ell}{2}$. 

\[\square\]

3. Smoothness of Lyapunov exponent and density of states

Recall Thouless’ formula

$$
L(E) = \int \log |E - E'| \, dN(E')
$$

which shows that the Lyapunov exponent $L(E)$ and the IDS $N(E)$ are related by the Hilbert transform. Hence it suffices to consider smoothness of $L(E)$.

Recall also that if $\eta$ is the site distribution of $H$, then

$$
L(E) = \int \log \left\| \begin{pmatrix} E - V & -1 \\ 1 & 0 \end{pmatrix} \frac{\cos \theta}{\sin \theta} \right\| \eta(dv) \nu_E(d\theta)
= \int A \log \left\| \begin{pmatrix} E \pm \lambda & -1 \\ 1 & 0 \end{pmatrix} \frac{\cos \theta}{\sin \theta} \right\| \nu_E(d\theta) \quad \text{(3.1)}
$$

in the Bernoulli case. Denote

$$
\Phi_E(\theta) = A \log \left\| \begin{pmatrix} E \pm \lambda & -1 \\ 1 & 0 \end{pmatrix} \frac{\cos \theta}{\sin \theta} \right\| \quad \text{(3.2)}
$$

which is a smooth function in $(\theta, E)$.

By (3.1) and Lemma 4,

$$
\| L(E) - (T_E)^\ell \Phi_E \|_\infty < C e^{-c\ell} \quad \text{(3.3)}
$$

noting the dependence of $T$ on $E$ (constants in the sequel may depend on $\lambda$).

Proof of the Theorem.
By the preceding, it suffices to show that $L(E)$ is a $C^k$-function of $E$, assuming $\lambda_0$ in (0.3) sufficiently small.

By (3.3), it will suffice to establish bounds on $\partial_E^{(k)} (T_E^\ell \Phi_E)$ that are uniform in $\ell$.

Returning to (0.10), let $\mathcal{G} = \{g_+^{(E)}, g_-^{(E)}, 1\}$. For $g_1, \ldots, g_\ell \in \mathcal{G}$, the chain rule gives

$$
\partial_E (\Phi_E \circ \tau_{g_1 \ldots g_\ell}) = (\partial_E \Phi_E) \circ \tau_{g_1 \ldots g_\ell} + \sum_{m=1}^\ell \left[(\Phi_E \circ \tau_{g_1 \ldots g_{m-1}})' \circ \tau_{g_m \ldots g_\ell}\right][\partial_E \tau_{g_{m+1}} \circ \tau_{g_{m+2} \ldots g_\ell}] 
$$

(3.4)

where $\partial_E \tau_g = -\sin^2 \tau_g$. Averaging (3.4) gives therefore

$$
\partial_E (T_E^\ell \Phi_E) = T_E^\ell (\partial_E \Phi_E) - \sum_{m=1}^\ell T_E^{\ell-m+1} [(T_E^{m-1} \Phi_E)' \sin^2 \theta].
$$

(3.5)

Thus

$$
|(3.5)| < C + \sum_{m=1}^\ell \| (T_E^{m-1} \Phi_E)' \|_\infty
$$

and applying Lemma 5 with $f = \Phi_E$ and $s = 1$ shows that $\| (T_E^m \Phi_E)' \|_\infty \leq C e^{-cm}$.

For $s = 2$, one obtains by iteration of (3.5) expansions of the form

$$
T_E^{m_1} \left( \sin^2 \theta (T_E^{m_2} (\sin^2 \theta (T_E^{m_3} \Phi_E)'))' \right)
$$

(3.6)

where $\ell = m_1 + m_2 + m_3$.

Again from Lemma 5, applied consecutively for $s = 1, s = 2$,

$$
|(3.6)| \lesssim \| (T_E^{m_2} (\sin^2 \theta (T_E^{m_3} \Phi_E)'))' \|_{H^1} \lesssim e^{-cm_2} \| (T_E^{m_3} \Phi_E)' \|_{H^2} \lesssim e^{-c(m_2+m_3)}.
$$

The continuation of the process is clear.
4. Further comments

1. One could conjecture a restricted spectral gap of the form (0.12) to be valid without arithmetical assumptions on $\lambda$. This would enable us to show that the density of states of the A-B model is $C^k$-smooth provided the coupling $\lambda \neq 0$ is sufficiently small (at least with $E$ restricted as in the above Theorem). Note that algebraic hypothesis on $\lambda$ appear in two places. Firstly in the expansion result from [B-Y], where it is used to establish the non-commutative diophantine property of the group (see also [B-G]). In fact weaker properties (such as positive box dimension at appropriate scales) would suffice. But the only available technique so far is that from [G-J-S] using arithmetic heights. Secondly, our application of Brenner’s result is based on algebraic conjugation. The conclusion from Proposition 2 is known to fail for certain values of $\mu$ and a complete understanding of which are the ‘free’ values of $\mu$ seems not available at the present.

2. The A-B model may in some sense be viewed as a non-commutative version of the classical Bernoulli convolution problem about which there is an extensive literature. Recall that for $0 < \lambda < 1$, one considers the measure $\nu_\lambda$ obtained from the random series

$$\sum_{n=0}^{\infty} v_n \lambda^n$$

where $\{v_n\}$ is a sequence of independent $\pm 1$-valued Bernoulli variables, $\mathbb{P}(v_0 = 1) = \mathbb{P}(v_0 + 1) = \frac{1}{2}$. As pointed out in [L-V], $\nu_\lambda$ is $\mu_\lambda$-stationary, where $\mu_\lambda$ is the probability measure supported on the two similarities $x \to \lambda x \pm 1$ putting $1/2$ mass on each. A major problem about the measures $\nu_\lambda$ is their absolute continuity. Starting from the work of Erdős, several results on this issue were obtained. In particular Solomyak [Sol] proved that $\nu_\lambda$ is absolutely continuous for almost all $\lambda > \frac{1}{2}$, while Erdős observed that $\nu_\lambda$ is singular if $\lambda^{-1}$ is a Pisot number. Returning to the A-B model, the situation turns out to be quite different, as our Theorem applies in particular if $\lambda^{-1}$ is a sufficiently large Pisot number and in this case the Furstenberg measure is absolutely continuous with $C^k$-density. The latter statement follows easily from the above analysis indeed (cf. also [B2]). Let $f \in L^\infty(\mathbb{T})$, $|f| \leq 1$ and $\text{supp} \hat{f} \subset [-2^{k+1}, -2^k] \cup [2^k, 2^{k+1}]$. By (2.26), (2.28), $\langle \nu, f \rangle = \langle \nu, T^m f \rangle$ for all $m$. Taking now $m$ large enough and applying the above Lemmas 3 and
2, it follows that
\[
\|T^{2m} f\|_\infty \leq C \|T^{2m} f\|_{H^1} \leq C \|T^m f\|_1 \leq e^{-rk} < C^{-k}
\] (4.2)
where \(C_\lambda\) can be made arbitrarily large for \(\lambda\) small enough. Hence we obtain
\[
|\langle \nu, f \rangle| < C_\lambda^{-k},
\]
from where the smoothness claim for \(\frac{d\nu}{d\theta}\).

Acknowledgment: The author is grateful to the mathematics department of UC Berkeley for their hospitality.

References

[B-L] P. Bougerol, J. Lacroix, Products of random matrices with applications to Schrödinger operators, Birkhauser 1985.

[B1] J. Bourgain, On the Furstenberg measure and density of states for the Anderson-Bernoulli model at small disorder, J. Analyse Math, Vol. 117 (2012), 273–295.

[B2] J. Bourgain, Finitely supported measures on \(SL_2(\mathbb{R})\) which are absolutely continuous at infinity, Geom. Funct. Anal. Lecture Notes, to appear

[B-G] J. Bourgain and A. Gamburd, On the spectral gap for finitely-generated subgroups of \(SU(2)\), Invent. Math. 171 (2008), 83–121.

[B-Y] J. Bourgain and A. Yehudayoff, Expansion in \(SL_2(\mathbb{R})\) and monotone expanders, GAFA, Vol. 23 (2013), 1–41.

[Br] J.L. Brenner, Quelques groupes tibris de matrices, CR Acad. Sc, Paris 241 (1955), 1689–1691.

[Ca-K-M] R. Carmona, A. Klein, G. Martinelli, Anderson localization for Bernoulli and other singular potentials, Comm. Math. Phys. 108 (1987), 41–66.

[G-J-S] A. Gamburd, D. Jacobson, P. Sarnak, Spectra of elements in the group ring of \(SU(2)\), JEMS, 1, 51–85.

[L-U] R.C. Lyndon, J.L. Ullman, Groups generated by two parabolic linear fractional transformations, Canad. J. Math, 21 (1969), 1388–1403.

[L-V] E. Lindenstrauss, P. Varju, Spectral gap in the group of Euclidean isometries, preprint 7/013.

[S-T] B. Simon, M. Taylor, Harmonic analysis on \(SL_2(\mathbb{R})\) and smoothness of the density of states in the one-dimensional Anderson model, Comm. Math. Pys. 101 (1985), 1–10.

[S-V-W] C. Shubin, T. Vakilian, T. Wolf, Some harmonic analysis questions suggested by Anderson-Bernoulli models, Geom. Funct. Anal. 8 (1998), 932–964.

[Sol] B. Solomyak, On the random series \(\sum \pm \lambda^n\) (an Erdos problem), Annals Math. (2), 142 (1995), no 3, 611–625.

Institute for Advanced Study, Princeton, NJ 08540

E-mail address: bourgain@ias.edu