On some operators connected with the Anderson model

Carol Shubin and Ramin Vakilian
Department of Mathematics
California State University
Northridge, Ca 91330

Thomas Wolff
Department of Mathematics
253-37 Caltech
Pasadena, Ca 91125

Remark added May 1997 This paper was a preliminary version of the paper “Some harmonic analysis questions suggested by Anderson-Bernoulli models” by the same authors. It is being posted here because section 5 contains proofs of some minor results which are stated but not proved in the paper submitted for publication. Specifically, Proposition 4.8 in the version submitted is proved differently here, and a finite volume version of Proposition 4.8 is also obtained and is applied to estimate the localization length for Anderson-Bernoulli models at small disorder.

The purpose of this paper is to prove some new lemmas on Fourier transforms, Theorems 1 and 2 below, and to apply these lemmas to some questions about the one dimensional Anderson model.

We follow the approach to the Anderson model in [3],[12], where results on localization and smoothness of the density of states are proved by means of an explicit formula for the Green’s function and then Fourier analysis to estimate the terms in the formula. Roughly speaking, Theorems 1 and 2 allow us to obtain sharper forms of these latter estimates. In particular, this makes the approach applicable also to questions about Bernoulli type models such as Holder continuity of the density of states.

Let

\[ \rho(x) = \min(1, \frac{1}{|x|}) \]  

and make the following definition: a set \( E \subset \mathbb{R}^n \) is \( \epsilon \)-thin if

\[ |E \cap D(x, \rho(x))| \leq \epsilon |D(x, \rho(x))| \]

for all \( x \in \mathbb{R}^n \), where \( D(x, r) \) is the disc centered at \( x \) with radius \( r \) and \( | \cdot | \) is Lebesgue measure. We also let \( E^c \) be the complement of the set \( E \).
Theorem 1 There are $\epsilon > 0$ and $C < \infty$ such that if $E$ and $F$ are $\epsilon$-thin sets in $\mathbb{R}^n$ then for any $f \in L^2$

$$\|f\|_2 \leq C(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)})$$ (2)

Here $\hat{f}(\xi) \overset{def}{=} \int_{\mathbb{R}^n} f(x)e^{-2\pi i x \cdot \xi}dx$ is the Fourier transform of $f$. Theorem 1 is a variant on the harmonic analyst’s uncertainty principle, i.e., on the fact that $f$ and $\hat{f}$ cannot both be concentrated on small sets. There are numerous related results in the literature, e.g. Logvinenko and Sereda [14] gave sharp conditions on $F$ under which 2 holds if $E = D(0, 1)$, and Amrein and Berthier [1] showed that 2 holds if both $E$ and $F$ have finite volume. Further results may be found in the survey articles [8] and [15] and the book [9]. The form of Theorem 1 is dictated by the application we want to make.

Let $G$ be a function on $\mathbb{R}^n$ with $\|G\|_\infty = 1$ and define an operator $T_G : L^2 \to L^2$ via

$$T_G f = \hat{G} f$$

It is clear that $\|T_G\|_{L^2 - L^2} = 1$. However, in practice one is interested in iterating $T_G$. We will use Theorem 1 to prove the following.

Theorem 2 Suppose that $\mu$ and $\nu$ are probability measures on the line, neither of which is a single Dirac mass\(^1\) and that $Q$ is a nondegenerate quadratic form\(^2\) in $\mathbb{R}^n$. Let $G$ and $H$ be functions on $\mathbb{R}^n$ such that

$$|G(x)| \leq |\hat{\mu}(Q(x))|$$

$$|H(x)| \leq |\hat{\nu}(Q(x))|$$

Then

$$\|T_H T_G\|_{L^2 - L^2} \leq \rho$$

where $\rho < 1$ depends only on $\mu$, $\nu$ and $Q$. Indeed, we can take $\rho = 1 - C^{-1} \lambda^2 \gamma$ where $C > 0$ depends only on $Q$ and $\gamma$ and $\lambda$ are any numbers in $(0, 1)$ such that

$$\mu \times \mu(\{(x, y) \in \mathbb{R} \times \mathbb{R} : |x - y| \geq \lambda\}) \geq \gamma$$

$$\nu \times \nu(\{(x, y) \in \mathbb{R} \times \mathbb{R} : |x - y| \geq \lambda\}) \geq \gamma$$

\(^1\)By this we mean that if $p \in \mathbb{R}$ then $\mu(\{p\}) < 1$ and $\nu(\{p\}) < 1$.

\(^2\)By this we mean that $Q(x) = \langle Ax, x \rangle$ with $A$ an invertible real symmetric matrix.
We emphasize that not just $\gamma$ but also $\lambda$ are supposed less than 1 - our argument does not give very good estimates at high disorder. Of course, any measure other than a single Dirac mass will satisfy (3) for some $\lambda \in (0, 1)$ and $\gamma$.

Here now is what we will prove about the (one-dimensional) Anderson model. In section 4, we give an alternate proof of Le Page’s theorem on Holder continuity of the density of states and extend the proof of localization in [12] to the case of Bernoulli models. See Theorem 3 below and the subsequent remarks. In section 5 we prove various refinements. For example, inside the spectrum of the laplacian the bounds in LePage’s theorem can be taken independent of the disorder as the disorder goes to zero (Proposition 5.2), and the spectrum is pure point even if one randomizes only at a sequence of sites which can become quite sparse at infinity (remark after Proposition 5.6).

Acknowledgments We thank A. Klein for pointing out reference [2].

This paper is the result of a seminar at California State University, Northridge on references [3],[11],[12] during the summer and fall of 1995. The third author is grateful to the CSUN mathematics department and Professors David Klein and William Watkins for office space. Most of the writing was done in May 1996 while the first and third authors were visiting Princeton University.

1.Proof of Theorem 1

We fix a radial, real-valued Schwartz function $\phi : \mathbb{R}^n \to \mathbb{R}$ with $0 \leq \hat{\phi} \leq 1$, $\text{supp} \hat{\phi} \subset D(0, 2)$ and $\hat{\phi} = 1$ on $D(0, 1)$. For $j \in \mathbb{Z}^+ \cup \{0\}$ we let $\phi_j(x) = 2^{jn} \phi(2^{j} x)$. Thus $\|\phi_j\|_1 = \|\phi\|_1$ and $\hat{\phi}_j(\xi) = \hat{\phi}(2^{-j} \xi)$ so that $\text{supp} \hat{\phi}_j \subset D(0, 2^{j+1})$, $\hat{\phi}_j = 1$ on $D(0, 2^{j})$.

Let $\psi_0 = \hat{\phi}$, and when $j \in \mathbb{Z}^+$ let $\psi_j = \hat{\phi}_j - \hat{\phi}_{j-1}$. Thus $\text{supp} \psi_j \subset D(0, 2^{j+1}) \setminus D(0, 2^{j-1})$ when $j \geq 1$, and $\sum_{j=0}^{\infty} \psi_j = 1$, where the convergence may be taken to mean that the sum converges uniformly to 1 and all derivatives converge uniformly to zero. Define now

$$S_N f = \sum_{j=0}^{N} \psi_j \cdot (\phi_j \ast f)$$

$$T_N f = \sum_{j=0}^{N} \psi_j \cdot (f - \phi_j \ast f)$$

Since the $\{\phi_j\}$ form an approximate identity and $\sum \psi_j$ converges to 1 in the sense specified
above, we may conclude that if \( f \in S \) then \( S_N f \) and \( T_N f \) converge in the topology of \( S \). We denote the limit operators by \( S \) and \( T \). Clearly \( S + T \) is the identity operator. One can think of \( S f \) and \( T f \) as \( f \) microlocalized to the regions \( |\xi| \leq |x| \) and \( |\xi| \geq |x| \) respectively.

The proof of Theorem 1 will be based on making appropriate \( L^2 \) estimates for the operators \( S \) and \( T \). The estimates will follow by applying Schur’s test to the integral kernels of the operators. The following lemma contains the necessary calculations.

**Lemma 1.1** Define

\[
A_N(x) = \sum_{j=0}^{N} \psi_j(x) \phi_j(x - y)
\]

\[
B_N(\xi, \eta) = \sum_{j=0}^{N} \hat{\psi}_j(\xi - \eta)(1 - \hat{\phi}_j(\eta))
\]

Then, for a suitable constant \( C \) which is independent of \( N \),

(i) \( \int |A_N(x,y)|dy \leq C \) for all \( x \).

(ii) \( \int |A_N(x,y)|dx \leq C \) for all \( y \).

(iii) \( \int |B_N(\xi,\eta)|d\eta \leq C \) for all \( \xi \).

(iv) \( \int |B_N(\xi,\eta)|d\xi \leq C \) for all \( \eta \).

Furthermore, if \( E \) and \( F \) are \( \epsilon \)-thin then

(v) \( \int_E |A_N(x,y)|dy \leq C\epsilon \) for all \( x \).

(vi) \( \int_F |B_N(\xi,\eta)|d\xi \leq C\epsilon \) for all \( \eta \).

We will first complete the proof of Theorem 1 assuming Lemma 1.1 and will then prove Lemma 1.1.

**Proof of Theorem 1** Note that

\[
S_N f(x) = \int A_N(x,y) f(y) dy
\]

and

\[
\widehat{T_N f}(\xi) = \int B_N(\xi,\eta) \hat{f}(\eta) d\eta
\]
Consequently, by parts (i)-(iv) of Lemma 1.1 and Schur’s test, the operators $S$ and $T$ extend to bounded operators on $L^2$ satisfying $S + T = \text{identity}$. Furthermore, if we let $\chi_E$ denote the indicator function of the set $E$, then using (ii) and (v) (respectively (iii) and (vi)) and Schur’s test, we have

$$
\|S(\chi_E f)\|_2 \leq C \epsilon^{\frac{1}{4}} \|f\|_2
$$  \hfill (4)

$$
\|\chi_F \hat{T} f\|_2 \leq C \epsilon^{\frac{1}{2}} \|f\|_2
$$  \hfill (5)

Suppose now that $f$ is given, with $\|f\|_2 = 1$. Then

$$
\hat{f} = S(\chi_{E^c} f) + S(\chi_E f) + \chi_{F^c} \hat{T} f + \chi_F \hat{T} f
$$

and therefore, using (4),(5),

$$
\|\hat{f} - S(\chi_{E^c} f) - \chi_{F^c} \hat{T} f\|_2 \leq C \epsilon^{\frac{1}{2}}
$$

If $\|f\|_{L^2(E^c)} \leq \alpha$, say, then we conclude that

$$
\|\hat{f}\|_{L^2(F^c)} \leq \|\hat{f} - \chi_{F^c} \hat{T} f\|_2 \\
\leq C(\alpha + \epsilon^{\frac{1}{2}}) \\
\leq \frac{1}{\sqrt{2}}
$$

provided $\epsilon$ and $\alpha$ have been chosen small. So $\|\hat{f}\|_{L^2(F^c)} \geq \frac{1}{\sqrt{2}}$ and the proof is complete.

**Proof of Lemma 1.1**

(i) For fixed $x$ there are at most three values of $j$ for which $\psi_j(x) \neq 0$. Furthermore $\|\psi_j\|_\infty \leq 1$ for any $j$, and $\|\phi_j\|_1 = \|\phi\|_1$ for any $j$. We conclude that the integral in (i) is $\leq 3\|\phi\|_1$.

(ii) Fix $y$ and let $\sum_x$ denote the sum over all $j \in \{0, \ldots, N\}$ such that $\text{dist}(y, \text{supp } \psi_j) \geq 1$. There are at most three values of $j$ with $\text{dist}(y, \text{supp } \psi_j) < 1$, and since $\phi \in S$ we have $|\phi_j(x - y)| \leq C 2^{jn}(1 + 2^j|x - y|)^{-3n}$, say. Hence

$$
\int |A_N(x, y)| dx \leq 3\|\phi\|_1 + \int \sum_x |\psi_j(x)\phi_j(x - y)| dx
$$
\[
\begin{align*}
&\leq 3\|\phi\|_1 + C \int \sum_{j} |\psi_j(x)|2^{jn}(1 + 2^j|x - y|)^{-3n} dx \\
&\leq 3\|\phi\|_1 + C \sum_{j} 2^{-2jn} \|\psi_j\|_1 \\
&= 3\|\phi\|_1 + C \sum_{j} 2^{-jn} \|\psi\|_1 \\
&\leq C
\end{align*}
\]

as claimed.

(iii) and (iv). We rewrite the definition of \(B_N\) as follows:

\[
B_N(\xi, \eta) = \sum_{j=0}^{N} \hat{\psi}_j(\xi - \eta) \sum_{i>j} \psi_i(\eta) \\
= \sum_{i=1}^{\infty} \psi_i(\eta) \sum_{j=0}^{\min(i-1,N)} \hat{\psi}_j(\xi - \eta) \\
= \sum_{i=1}^{\infty} \psi_i(\eta) \phi_{i_*}(\xi - \eta) \tag{6}
\]

where we have set \(i_* = \min(i - 1, N)\). Note the similarity between (6) and the definition of \(A_N(\eta, \xi)\). We may therefore prove (iv) by repeating the proof of (i). For (iii), we further rewrite (6) as

\[
B_N(\xi, \eta) = \sum_{i=1}^{N} \psi_i(\eta) \phi_{i-1}(\xi - \eta) + \sum_{i>N} \psi_i(\eta) \phi_N(\xi - \eta) \\
= \sum_{i=1}^{N} \psi_i(\eta) \phi_{i-1}(\xi - \eta) + (1 - \hat{\phi}_N(\eta)) \phi_N(\xi - \eta)
\]

We have \(\int |(1 - \hat{\phi}_N(\eta)) \phi_N(\xi - \eta)| d\eta \leq \int |\phi_N(\xi - \eta)| d\eta = \|\phi_N\|_1 = \|\phi\|_1\), and the estimate \(\int |\sum_{i=1}^{N} \psi_i(\eta) \phi_{i-1}(\xi - \eta)| d\eta \leq C\) follows by repeating the proof of (ii). This proves (iii).

(v) and (vi). We will only prove (v), since (vi) once again follows by the same argument in view of (6).
Fix $x$ and let $j$ be such that $\psi_j(x) \neq 0$. We claim that

\[
\int_E |\phi_j(x-y)|dy \leq C\epsilon \tag{7}
\]

If we can prove this we are done as in the proof of (i), since there are only three possible values for $j$ and the $\psi_j$’s are uniformly bounded.

To prove the claim, we use the following simple geometric property of the discs $D(x, \rho(x))$: if $t > \rho(x)$, then $D(x, t)$ can be covered by discs of the form $D(x_k, \rho(x_k))$ in such a way that

\[
\sum_k |D(x_k, \rho(x_k))| \leq C|D(x, t)|
\]

This holds since

\[
|y-x| \leq \rho(x) \Rightarrow C^{-1}\rho(x) \leq \rho(y) \leq C\rho(x) \tag{8}
\]

See for example [10], lemma 2.5. We sketch the argument for the convenience of the reader. Property (8) is easily verified. Now choose a maximal set of points $\{x_k\} \subset D(x, t)$ with the property that $|x_k - x_j| \geq \min(\rho(x_j), \rho(x_k))$ for all $j$ and $k$. By maximality, the discs $D(x_k, \rho(x_k))$ cover $D(x, t)$. On the other hand, (8) implies that for a suitable constant $C_0$ the discs $D(x_k, C_0^{-1}\rho(x_k))$ are disjoint and contained in $D(x, C_0t)$, and therefore

\[
\sum_k |D(x_k, \rho(x_k))| \leq \sum_k |D(x_k, C_0^{-1}\rho(x_k))| \leq |D(x, C_0t)| \lesssim |D(x, t)|.
\]

It follows that

\[
|D(x, t) \cap E| \leq \sum_k |D(x_k, \rho(x_k)) \cap E| \leq \epsilon \sum_k |D(x_k, \rho(x_k))| \leq C\epsilon|D(x, t)| \tag{9}
\]

for any $x$ and $t \geq \rho(x)$. Next, if $\psi_j(x) \neq 0$ then $2^j$ is comparable to $\rho(x)^{-1}$. Since $\phi \in S$, we therefore have $|\phi_j(x-y)| \leq C\rho(x)^{-n}(1+\frac{|x-y|}{\rho(x)})^{-2n}$, so we may estimate the integral (7) by

\[
(7) \lesssim \sum_{k \geq 0} 2^{-2nk}\rho(x)^{-n}|D(x, 2^k\rho(x)) \cap E|
\lesssim \sum_{k \geq 0} 2^{-2nk}\rho(x)^{-n}\epsilon(2^k\rho(x))^n
\lesssim \epsilon
\]

where we used (9). \qed
Remarks 1. Theorem 1 is sharp in the sense that the rate function $\rho(x)$ cannot be replaced by one which decays more slowly at $\infty$. We give the counterexamples in the case $n = 1$. Let $\phi$ be a fixed $C^\infty_0$ function and consider the functions

$$\Phi_N(x) \overset{def}{=} \sum_{j=-N}^{N} \phi(N(x-j))$$

Then

$$\tilde{\Phi}_N(\xi) = D_N(\xi)N^{-1}\hat{\phi}(\frac{\xi}{N})$$

where $D_N(\xi) = \frac{\sin(2\pi(N+\frac{1}{2})\xi)}{\sin(\pi\xi)}$ is the Dirichlet kernel. Let $E_N^A = \cup_{j=-AN}^{AN}(j - \frac{A}{N}, j + \frac{A}{N})$ and let $F_N^A$ be the complement of $E_N^A$. Then it is not hard to see the following: for any $\eta > 0$ there is $A < \infty$ such that for any large $N$, we have

$$\|\Phi_N\|_{L^2(F_N^A)} + \|\Phi_N\|_{L^2(E_N^A)} < \eta\|\Phi_N\|_2$$

Namely, $\Phi_N$ will vanish on $F_N^A$ due to the support property of $\phi$, and $|\tilde{\Phi}_N|^2$ will have most of its mass on $E_N^A$ since $|D_N|^2$ is concentrated near integers and $\hat{\phi}(\frac{\xi}{N})$ dies out rapidly when $|\xi|$ is large compared with $N$.

On the other hand, if $\rho$ is positive and continuous and $|x|\rho(x) \to \infty$ as $|x| \to \infty$, then for any $\epsilon$ and $A$, we will have $|D(x, \rho(x)) \cap E_N^A| < \epsilon|D(x, \rho(x))|$ for all $x$ provided $N$ is sufficiently large. This shows that the rate function in Theorem 1 is the optimal one, as claimed.

2. However, we do not know the answer to the following question: does Theorem 1 remain true if “There are $\epsilon > 0$ and $C < \infty\ldots$” is replaced by “For all $\epsilon < 1$ there is $C < \infty\ldots$”?

2. Proof of Theorem 2

Theorem 2 will follow by combining Theorem 1 with the following elementary lemma about “thinness” of the set where the characteristic function of a probability distribution is close to its maximum value.
Lemma 2.1 Let \( \mu \) be a probability measure on \( \mathbb{R} \) which is not a single Dirac mass and let \( E_\delta = \{ \xi \in \mathbb{R} : |\hat{\mu}(\xi)| > 1 - \delta \} \). Then for every \( \epsilon > 0 \) there is \( \delta > 0 \) such that \( \sup_{\alpha \in \mathbb{R}} |(a, a+1) \cap E_\delta| < \epsilon \). In fact, if (3) holds (with \( \lambda \in (0, 1) \)) we can take \( \delta = C^{-1} \gamma \lambda^2 \epsilon^2 \) with \( C \) a numerical constant.

Proof We use the following fact about the cosine function, which holds (uniformly over \( t \) and \( \Gamma \)) since the set in question is a union of intervals of length \( \approx |t|^{-1} \sqrt{\alpha} \) centered at integer multiples of \( t^{-1} \).

\[
\left| \{ \xi \in (a, a+1) : \cos(2\pi t \xi) > 1 - \alpha \} \right| \leq C \alpha^{1/2} \max(1, |t|^{-1}) \tag{10}
\]

Fix \( a \) and let \( E_\delta^a = E_\delta \cap (a, a+1) \). We clearly have

\[
(1 - 2\delta)|E_\delta^a| \leq \int_{E_\delta^a} |\hat{\mu}(\xi)|^2 d\xi \tag{11}
\]

On the other hand

\[
\int_{E_\delta^a} |\hat{\mu}(\xi)|^2 d\xi = \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{E_\delta^a} \cos(2\pi \xi(x-y)) d\xi d\mu(x) d\mu(y)
\]

Let \( \alpha \) satisfy \( C \alpha^{\frac{1}{2}} = \frac{1}{2} |E_\delta^a| \lambda \), with \( C \) as in (10). Then

\[
\int_{E_\delta^a} \cos 2\pi \xi(x-y) d\xi \leq |E_\delta^a \cap \{ \xi : \cos(2\pi \xi(x-y)) > 1 - \alpha \}| + (1 - \alpha)|E_\delta^a \cap \{ \xi : \cos(2\pi \xi(x-y)) \leq 1 - \alpha \}| \tag{12}
\]

If \( |x - y| \geq \lambda \), then (10) implies

\[
|E_\delta^a \cap \{ \xi : \cos(2\pi \xi(x-y)) > 1 - \alpha \}| \leq \frac{1}{2} |E_\delta^a| \lambda \max(1, |x - y|^{-1}) \leq \frac{1}{2} |E_\delta^a|
\]

since \( \lambda \leq 1 \). By (12),

\[
\int_{E_\delta^a} \cos(2\pi \xi(x-y)) d\xi \leq \left( 1 - \frac{\alpha}{2} \right)|E_\delta^a|
\]
when $|x - y| > \lambda$. Hence,
\[
\int_{E^g_\delta} |\mu(\xi)|^2 d\xi \leq (1 - \frac{\alpha}{2}) |E^g_\delta| \gamma + |E^g_\delta|(1 - \gamma)
\]
\[
= (1 - \frac{\alpha \gamma}{2}) |E^g_\delta|
\]
Combining this with (11) we conclude that $\delta \geq \frac{\alpha \gamma}{4}$. By choice of $\alpha$ this means that $\delta \geq C^{-1} \gamma \lambda^2 |E^g_\delta|^2$ where the constant is universal. This is equivalent to the lemma. 

Proof of Theorem 2. Any nondegenerate quadratic form $Q$ has the following property: $Q$ maps each disc $D(x, \rho(x))$ onto an interval $I_x$ with $|I_x| \in [C^{-1}, C]$. Furthermore, if $E \subset I_x$ then
\[
\frac{|Q^{-1}E \cap D(x, \rho(x))|}{|D(x, \rho(x))|} \leq C|E|^{1/2}
\]
This holds essentially because when $|x|$ is large, $|\nabla Q(x)|$ is comparable to $\rho(x)^{-1}$. We leave details to the reader.

We conclude by Lemma 2.1 that for any given $\epsilon$ if $\delta = C^{-1} \gamma \lambda^2 \epsilon^2$ then the sets
\[
\{x \in \mathbb{R}^n : |\hat{\mu}(Q(x))| > 1 - \delta\}
\]
\[
\{\xi \in \mathbb{R}^n : |\hat{\nu}(Q(x))| > 1 - \delta\}
\]
are $\sqrt{\epsilon}$–thin. Here $C$ depends on $Q$ only. Consequently by Theorem 1, we can choose $\beta > 0$ and $\eta > 0$ depending on $Q$ only so that, with $\delta = \beta \gamma \lambda^2$ and
\[
E = \{x : |G(x)| > 1 - \delta\}
\]
\[
F = \{\xi : |H(\xi)| > 1 - \delta\}
\]
every function $f$ must satisfy either
\[
\|f\|_{L^2(E^c)} \geq \eta \|f\|_2
\]
or
\[
\|\hat{f}\|_{L^2(E^c)} \geq \eta \|f\|_2
\]
To finish the proof of the lemma, fix $f$ and consider two cases:

(i) $\|f\|_{L^2(E^c)} \geq \eta \|f\|_2$. 

10
(ii) $\|f\|_{L^2(E^c)} < \eta \|f\|_2$.

In case (i) we have

$$
\|Gf\|_2^2 = \|Gf\|_{L^2(E^c)}^2 + \|Gf\|_{L^2(E)}^2 \\
\leq (1 - \delta)^2 \|f\|_{L^2(E^c)}^2 + \|f\|_{L^2(E)}^2 \\
\leq ((1 - \delta)^2 \eta^2 + 1 - \eta^2) \|f\|_2^2 \\
= \|f\|_2^2 - (2\delta - \delta^2) \eta^2 \|f\|_2^2
$$

so that $\|H\hat{G}f\|_2^2 \leq \|Gf\|_2^2 \leq (1 - \alpha) \|f\|_2^2$ with $\alpha = (2\delta - \delta^2) \eta^2$.

In case (ii) we have also

$$
\|Gf\|_{L^2(E^c)} < \eta \|Gf\|_2
$$

since $\sup_{E^c} |G| \leq \inf_E |G|$. Therefore

$$
\|\hat{G}f\|_{L^2(E^c)} \geq \eta \|\hat{G}f\|_2
$$

Then the argument for case (i) shows that

$$
\|H\hat{G}f\|_2^2 \leq (1 - \alpha) \|\hat{G}f\|_2^2
$$

with $\alpha$ as above. So $\|H\hat{G}f\|_2^2 \leq (1 - \alpha) \|f\|_2^2$ and the proof is complete, since $\alpha \approx \gamma \lambda^2$.

**Remarks**

1. We will also need a slight variant on Theorem 2 where the definition of $T_G$ is modified as follows

$$
T_G f(x) \stackrel{df}{=} \hat{G} f(Ux)
$$

with $U$ being a fixed orthogonal map of $\mathbb{R}^n$. This version can be obtained by applying Theorem 2 as originally stated with $f$ and $Q$ replaced by $f(Ux)$ and $Q(Ux)$. We skip the details.

2. In section 5 we will need another related estimate which can be stated as follows: suppose that $\delta \in (0, 1)$ and that $|G(x)| \leq e^{-\delta|x|^2}$, $|H(x)| \leq e^{-\delta|x|^2}$. Then (uniformly with respect to $\delta \in (0, 1)$)

$$
\|T_H T_G\|_{L^2 \to L^2} \leq 1 - C^{-1} \delta
$$

This can be proved in several ways. For example, we can use the preceding argument: the set where $e^{-\delta|x|^2} > 1 - A^{-1} \delta$ is a disc centered at the origin with radius $\approx A^{-1/2}$ and is
therefore $\epsilon$-thin for any given $\epsilon$ if $A$ is large. Accordingly, if we fix a sufficiently large $A$ and define $E = \{ x : |G(x)| \geq 1 - A^{-1}\delta \}$, $F = \{ x : |H(x)| \geq 1 - A^{-1}\delta \}$, then for a suitable fixed constant $\eta$ every function must satisfy either $\|f\|_{L^2(E)} \geq \eta \|f\|_2$ or $\|\hat{f}\|_{L^2(F)} \geq \eta \|f\|_2$. It then follows as before that $\|T_H T_G f\|_2^2 \leq (1 - \alpha) \|f\|_2^2$ with $\alpha = (2A^{-1}\delta - A^{-2}\delta^2)\eta^2 \approx \delta$, as claimed.

3. Supersymmetric formalism

We will summarize some things to be found for example in [3], with some slight modifications since we want to work directly with two dimensional Euclidean Fourier transforms instead of using Hankel transforms as is done there. We refer to [3] for the definitions of the supersymmetric formalism and will follow the notation there. Let $\Lambda$ be a finite graph and assume for simplicity that there is no edge which connects a vertex to itself. Let $V$ be a real-valued function on $\Lambda$. Then we can consider Schrödinger operators on $l^2(\Lambda)$ defined as follows:

$$H = \Delta + V.$$  
Here $\Delta f(x) \overset{\text{def}}{=} \frac{1}{2} \sum_{y \sim x} f(y)$ where $\sim$ is the adjacency relation on the graph, and $V$ is the operator of multiplication by the function $V$. We may also regard $H$ as a matrix with entries

$$H_{mn} = \begin{cases} V(m) & \text{if } m = n \\ 1/2 & \text{if } m \sim n \\ 0 & \text{otherwise} \end{cases}$$

If $\text{im}(z) > 0$, and $m, n \in \Lambda$, then we let $G(m, n, z)$ be the Green’s function, i.e., the matrix $G(m, n, z)_{m \in \Lambda} \in \Lambda$ which is inverse to $H - z \cdot \text{identity}$. The basic formula (cf. [3], the first equation of section 2) is that

$$G(m, n, z) = 2\pi i \int \psi(m) \overline{\psi(n)} e^{-2\pi i \sum_{x \in \Lambda} \Phi(x) - ((H - z)\Phi)(x)} D_{\Lambda} \Phi$$  
(14)

Here $D_{\Lambda} \Phi = \Pi_{x \in \Lambda} d\Phi(x)$, and other notation is as in [3].

Define a supersymmetric Fourier transform via

$$\mathcal{F}f(\Phi) = \int f(\Phi) e^{-2\pi i \Phi \cdot \Phi} d\Phi$$

3Actually, formula 14 is obtained from the formula in [3] by making a change of variables $\Phi \to \sqrt{2\pi} \Phi$. The supersymmetric integral is invariant under such changes of variable.
It is clear that the integral 14 can be expressed in terms of iterated supersymmetric Fourier transforms, which we will now relate to ordinary two dimensional Euclidean Fourier transforms, following [3]. If \( f \) is a (nice enough) function on the half line \( \mathbb{R}^+ \) then let us define

\[
Df = -\frac{1}{\pi} f'
\]

We may consider \( f(|x|^2) \), a radial function on \( \mathbb{R}^2 \). Its two dimensional (Euclidean) Fourier transform is another radial function, hence of the form \( g(|\xi|^2) \). We denote \( g \) by \( Tf \). Furthermore, we denote the operator \( TD \) (i.e. \( D \) followed by \( T \)) by \( S \). Then we have the following lemma:

**Lemma 3.1** \( DS = T \)

**Proof** By the chain rule,

\[
\nabla(f(|x|^2)) = -Df(|x|^2) \cdot 2\pi x
\]

Take the two dimensional Fourier transform of this formula using the usual rules \( \hat{\nabla F} = 2\pi i\xi \hat{F} \), \((2\pi ixF = \nabla \hat{F})\). Thus

\[
2\pi i Tf(|\xi|^2)\xi = -i\nabla(SF(|\xi|^2)) = 2\pi i DSF(|\xi|^2)\xi
\]

and the lemma follows. \( \square \)

If \( f \) is a function on \( \mathbb{R}^+ \) (assumed analytic and with rapid decay at \( \infty \)) then (cf. [3]) the definitions of functions of supervariables lead to the formula \( f(\Phi^2) = f(\phi^2) + f'(\phi^2)\bar{\psi}\psi \). The formulas for the supersymmetric Fourier transform are

**Lemma 3.2**

(a) If \( F(\Phi) = f(\Phi^2) \) then \( \mathcal{F} F(\Phi) = (Sf)(\Phi^2) \)

(b) If \( F(\Phi) = f(\Phi^2)\bar{\psi} \) then \( \mathcal{F} F(\Phi) = -i(Tf)(\Phi^2)\bar{\psi} \)

and also

(c) \( \int f(\Phi^2)\bar{\psi}\psi d\Phi = \frac{1}{\pi} \int_{\mathbb{R}^2} f(|x|^2)dx \)

**Proof** (a) From the definitions [3], we have

\[
e^{-2\pi i\Phi \cdot \bar{\Phi}} = e^{-2\pi i\phi \cdot \bar{\phi}}(1 - \pi i(\bar{\psi}\psi + \bar{\psi}\psi) + \pi^2 \bar{\psi}\psi\bar{\psi}\psi)
\] (15)
Therefore, from the definition of the integral ([3] p.230)

\[ \mathcal{F}F(\tilde{\Phi}) = \int (f(\phi^2) + f'(\phi^2)\bar{\psi}\psi)e^{-2\pi i\Phi}\bar{\phi}d^2\phi \]

\[ = \int (\frac{-1}{\pi} f'(\phi^2) - \pi f(\phi^2)\bar{\psi}\psi)e^{-2\pi i\Phi}\bar{\phi}d^2\phi \]

\[ = Sf(\phi^2) - \pi T f(\phi^2)\bar{\psi}\psi \]

\[ = Sf(\phi^2) \]

where the last line follows by Lemma 3.1.

(b) Once again we substitute (15) for \( e^{-2\pi i\phi}\bar{\Phi} \) in the definition of \( F \). Since the \( \psi \) variables anticommute, and using the definition of the integral, we get

\[ \mathcal{F}F(\tilde{\Phi}) = \int f(\phi^2)\bar{\psi}e^{-2\pi i\phi}\bar{\phi}d^2\phi \]

\[ = -i \int f(\phi^2)\bar{\psi}e^{-2\pi i\phi}\bar{\phi}d^2\phi \]

which is (b).

(c) Using anticommutativity, then the definition of the integral we get

\[ \int f(\Phi^2)\bar{\psi}\psi d\Phi = \int f(\phi^2)\bar{\psi}\psi d\Phi \]

\[ = -\frac{1}{\pi} f(\phi^2)d^2\phi \]

as claimed. \( \square \)

In the one dimensional case i.e. \( \Lambda = \{-l, \ldots, l\} \) with the usual adjacency relation then integrating out the formula (14) one variable at a time and using Lemma 3.2 gives the following formulas for the Green’s function (see [3],[12]): let \( \beta_j(r) = e^{-2\pi i(V_j-z)r} \) and also let \( \beta_j \) be the operator on functions defined by multiplication by the function \( \beta_j \). Then, when \( m \leq n \) (and \( \text{Im}z > 0 \)),

\[ G(m, n, z) = 2^{m-n+1} \int (\Pi_{j=1}^{m-1}S\beta_j 1)(|x|^2)((\Pi_{j=0}^{n-1}T\beta_{n-j} 1)(\Pi_{j=0}^{l-n-1}S\beta_{l-j} 1)|x|^2)\beta_m(|x|^2)dx \]

(16)
Here, if $O_j$ are (noncommuting) operators, then we use $\prod_{j=1}^n O_j$ to denote the operator $O_n O_{n-1} \ldots O_1$, etc. Thus for example the expression $((\prod_{j=1}^{m-1} S_j) 1)(|x|^2)$ means start with the constant function 1, apply the operator $\beta_1$, then $S$, then $\beta_{-1}$ and so forth and evaluate the resulting expression at $|x|^2$. In the case of the Anderson model with single site distribution $\nu$, taking expectations in the above formula for $G(0,0,z)$ leads to

$$\mathcal{E}(G(0,0,z)) = 2i \int_{\mathbb{R}^2} ((S\Gamma)^i 1(|x|^2))^2 \Gamma(|x|^2) dx$$

where we have let $\Gamma(r) = \hat{\nu}(r) e^{2\pi i r z}$ and have also used $\Gamma$ to denote the operator of multiplication by $\Gamma$.

More generally, suppose that $\Lambda$ is a tree. We will derive a formula analogous to (16), following the treatment for the Bethe lattice in [11]. Define a Gaussian function\(^4\) to be a function $y(\xi)$ of the form

$$y(\xi) = e^{2\pi i \xi r}, \text{ where } \text{Im} \xi > 0$$

It is easy to check that

$$Sy(\xi) = y(-\xi)$$

In particular, the class of Gaussian functions is closed under $S$. Note that it is also closed under forming products.

Now consider a Schrodinger operator $H$ on a finite connected tree $b$ and assume that $O$ is a root of $b$, i.e. a site on $b$ which has only one neighbor. Define

$$Y_b(\Phi) = \int e^{-2\pi i \xi \Phi(O)} e^{-2\pi i \Phi(H-z)\Phi} D\Phi$$

Thus $Y_b$ is a function of the single supervariable $\Phi$. An obvious induction based on the preceding remarks and Lemma 3.2(a) shows it is Gaussian, i.e. that $Y_b = y(\Phi^2)$ for a certain $\xi$ depending on $b$ and $O$ with $\text{Im} \xi < 0$. In particular, $|y(\xi)| \leq 1$.

Now let $\Lambda$ be any finite connected tree and fix two sites $m, n \in \Lambda$. There is a unique path from $m$ to $n$, which we denote by $\gamma$. Thus $\gamma$ consists of sites $m_0, \ldots, m_d$, where $d$ is the distance from $m$ to $n$ and $m_0 = m$, $m_d = n$. We define a branch of $\Lambda$ to be a connected component of $\Lambda \setminus \gamma$. For each branch $b$ there is a unique $m_j \in \gamma$ such that $b \cup \{m_j\}$ is connected; we say that $b$ emanates from $\gamma$ at $m_j$. If $b$ emanates from $\gamma$ at

\(^4\)This terminology results from the fact that one mainly considers the composition of $y(\xi)$ with a quadratic form. We will also use the notation 18 when $\xi$ is real.
m_j there is a unique site O_b \in b which is adjacent to m_j, and O_b is a root of b. By the remarks in the preceding paragraph we have

\[ \int e^{-2\pi i \Phi(m_j)} e^{-2\pi i \sum_{i \in b} \Phi(i) \cdot ((H_b - \gamma) \Phi(i))} D_b \Phi = y^b(\Phi^2_{m_j}) \]

where y^b is a Gaussian function. Here \((H_b \Phi)(i) = \frac{1}{2} \sum_{j \in b} \Phi(j) + V(i) \Phi(i)\) is the Schrödinger operator on b defined by H. Now define

\[ \xi_{m_j} = \Pi_b y^b \]

where the product is taken over all branches which emanate from \gamma at m_j. We will also use \xi_{m_j} to denote the operator of multiplication by the function \xi_{m_j}. Integrating out the basic formula (14) and using Lemma 3.2 (b) and (c), it then follows that

\[ G(m, n, \gamma) = 2i^{1-d} \int_{\mathbb{R}^d} (\Pi_{j=0}^{d-1} (T \beta_{m_j} \xi_{m_j}) 1 \langle |x|^2 \rangle (\beta_{m_d} \xi_{m_d})(|x|^2)) \]  

(19)

where the functions \xi_{m_j} are bounded by 1 and depend only on the values of V at sites on branches of \Lambda emanating from \gamma at m_j. If \Lambda = \{-l, \ldots, l\} then (19) reduces to (16), since branches emanate only from the endpoint sites m and n so that \xi_m = (\Pi_{i=-l}^{m-1} S \beta_{i}) 1, \xi_n = (\Pi_{i=0}^{l-1} S \beta_{l-1}) 1, and \xi_{m_j} = 1 when 0 < j < d.

4. Holder continuity of the density of states

Let \(g_l(z) = \mathcal{E}(G(0, 0, z))\), where G is the Green’s function for the Anderson model on \{-l, \ldots, l\} with single site distribution \(\nu\). We assume that

\[ \int |x| d\nu(x) < \infty \]  

(20)

We denote \(z = E + i\eta, E, \eta \in \mathbb{R}\) and will always assume that \(\eta > 0\). We will prove the following:

Theorem 3 For any \(\epsilon_0 > 0\) there are \(\epsilon_1 > 0\) and \(C < \infty\), depending on \(\nu, \epsilon_0\) and a bound for \(|E|\), so that if \(\eta \leq \frac{1}{2} \) and \(l > \epsilon_0 \log \frac{1}{\eta}\) then \(|g_l(z)| \leq C \eta^{-(1-\epsilon_1)}\).
An immediate corollary is Holder continuity of the integrated density of states for the Anderson model on $\mathbb{Z}$ with single site distribution satisfying (20). Namely, by letting $l \to \infty$ in Theorem 3 we obtain the bound

$$|g(z)| \leq C\eta^{-(1-\epsilon)}$$

for all $z = E + i\eta$ with $0 < \eta < \frac{1}{2}$. Since $g(z)$ is the $\bar{\tau}$-derivative of the harmonic extension of the integrated density of states, this bound is equivalent to Holder continuity (cf. [18], ch. 5). One can also obtain an estimate of the finite volume density of states directly from Theorem 3 using that the expected number of states in the interval $(E - \eta, E + \eta)$ is $\leq C\eta \operatorname{Im} g_l(E + i\eta)$.

Remarks The assumption on $l$ in Theorem 3 is easily seen to be best possible. This argument is in [17]. Namely, if $\nu$ is Bernoulli, then there are only $2^{2l+1}$ possible choices for $H$ on $\{-l, \ldots, l\}$, hence (being generous) at most $(2l+1)2^{2l+1}$ possible eigenvalues. So $g_l$ is the Borel transform of a measure with $\leq (2l+1)2^{2l+1}$ mass points and then it follows that $\sup_E \operatorname{Im} g_l(E + i\eta) \gtrsim ((2l+1)2^{2l+1}\eta)^{-1}$, which is large compared with $\eta^{-(1-\epsilon)}$ for fixed $\epsilon$ if $l$ is small compared with $\log \frac{1}{\eta}$.

LePage’s arguments are adapted to finite volume in [5], Theorem 4.1, where a result closely related to Theorem 3 is proved.

We note also that the proof of Theorem 3 does not really use stationarity. With minor changes it works for independent, non-identically distributed $V_j$’s provided (say) that one assumes a uniform bound in $21$ and uniform lower bound on the disorder, i.e. on $\lambda$ and $\gamma$ in (3). See also section 5 in this connection.

We now prove Theorem 3. This argument is related to the proof of Lipschitz continuity of the integrated density of states given in [3], except that by using Theorem 2 we can make the relevant estimates without assuming decay of $\hat{\nu}$. In the present situation Lipschitz continuity of course fails.

Following [3] we work from formula (17) and define Hilbert spaces $H_0$ and $H_1$ on the half line as follows

$$\|f\|_{H_0}^2 = \int_{\mathbb{R}^2} \left( \frac{|f(|x|^2)|}{|x|} \right)^2 dx$$

$$\|f\|_{H_1}^2 = \int_{\mathbb{R}^2} |f(|x|^2)|^2 + |Df(|x|^2)|^2 dx$$

17
Furthermore we define an operator $R$ as follows: given a (nice) function $f$ on the half line, consider the function on $\mathbb{R}^2$,
\[
f(|x|^2) \frac{x}{|x|^2}
\]
Its Fourier transform is (by standard symmetry properties, cf. [18], ch. 3) again of the form
\[
g(|x|^2) \frac{x}{|x|^2}
\]
and we define $Rf = ig$.

**Lemma 4.1** $Sf = f(0) + Rf$. In particular, if $f$ vanishes at 0 then so does $Sf$, and $Rf$ and $Sf$ coincide.

**Proof** Fix $s > 0$. We have
\[
Sf(s^2) = -\frac{1}{\pi} \int f'(|x|^2)e^{-2\pi i(s,0):x} \, dx
\]
\[
= - \int \frac{d}{dr}(f(r^2))e^{-2\pi i r s \cos \theta} \frac{d\theta}{2\pi}
\]
\[
= f(0) + is \int f(r^2) \frac{\cos \theta}{r} e^{-2\pi i r s \cos \theta} \, r \, dr \, d\theta
\]
\[
= f(0) + Rf(s^2)
\]
since the integral on the next to last line is the value at $(s,0)$ of the Fourier transform of the function $\frac{x}{|x|^2} f(|x|^2)$ \qed

Iterating the lemma, it follows that $(ST)^j 1 = \sum_{j=0}^i (RT)^j 1$. Furthermore, by the Plancherel theorem, $R$ is an isometry on $H^0$ and (using Lemma 3.1) $S$ is an isometry on $H^1$. By Theorem 2 applied to the functions $\frac{x}{|x|^2} f(|x|^2)$ with the quadratic form $Q(x) = |x|^2$, we also have
\[
\|(RT)^j f\|_{H^0} \leq \rho^j \|f\|_{H^0}
\]  \hspace{1cm} (21)
for $j \geq 2$, where $\rho < 1$ depends on $\nu$ only. Furthermore, since 20 implies that $\nu'$ is bounded, it is easily checked that
\[
\|ST f\|_{H^1} \leq A\|f\|_{H^1}
\]  \hspace{1cm} (22)
where $A$ depends on a bound for $|E|$.

In order to exploit (21) and (22), we let $\phi$ be a $C^\infty_0$ function which is equal to 1 in a neighborhood of the origin. We define $\phi_t(x) = \phi(t^{-1}x)$. We also fix an index $k \geq 2$ and may then express $(SG)^{l}1$ for $l > k$ in the following form:

$$(SG)^{l}1 = \sum_{j=k}^{l} (RG)^{j}1 + (ST)^{k}\phi + (RG)^{k}(1 - \phi)$$

(23)

In view of (22), the second term satisfies the following estimate:

$$\| (ST)^{k}\phi \|_{H^1} \leq CA^k$$

(24)

For the remaining estimates we use the fact that

$$|\Gamma(r^2)| \leq e^{-2\pi\eta r^2}$$

(25)

and therefore also

$$\| (1 - \phi)\Gamma \|_{H^0} \leq C(\log \frac{1}{\eta})^{\frac{1}{2}}$$

(26)

provided $\eta \geq \eta$, say. Taking $t = 1$ and using (21), we get the following bound for the third term in (23):

$$\| (RG)^{k}(1 - \phi) \|_{H^0} \leq \rho^k(\log \frac{1}{\eta})^{\frac{1}{2}}$$

(27)

Next, we have the following lemma:

**Lemma 4.2** $\| (RG)^21 \|_{H^0} \leq C(\log \frac{1}{\eta})^{\frac{1}{2}}$

Proof We have $\| (RG)^2(1 - \phi)\Gamma \|_{H^0} \leq \| 1 - \phi \|_{H^0} \leq C(\log \frac{1}{\eta})^{\frac{1}{2}}$. On the other hand, define $g : \mathbb{R}^2 \to \mathbb{R}$ via $g(x) = \frac{1}{|x|}(\phi \Gamma)(|x|^2)$. Then $\|g\|_{L^1(\mathbb{R}^2)} \leq C\eta^{\frac{1}{2}}$. Consequently, $\|g\|_{\infty} \leq C\eta^{\frac{1}{2}}$ or equivalently

$$\| (RG\phi)\Gamma(r^2) \| \leq C\eta^{\frac{1}{2}}r$$

It follows by 25 that

$$\| (\Gamma RG\phi)\Gamma(r^2) \| \leq C\eta^{\frac{1}{2}}r e^{-2\pi\eta r^2}$$

and therefore that

$$\| \Gamma RG\phi \|_{H^0} \leq C(\int_0^{\infty} \eta e^{-4\pi\eta r^2} r dr)^{\frac{1}{2}}$$

19
which is bounded independently of $\eta$. The lemma follows, since $R$ is an isometry. \hfill \Box

We conclude that the summands in the first term of (23) satisfy

$$\|(R\Gamma)^j 1\|_{H^0} \leq C \rho^j (\log \frac{1}{\eta})^{\frac{k}{2}}$$

(28)

Combining (23),(24),(27),(28) we conclude that for any $k \geq 2$ and $l > k$, $(S\Gamma)^l 1 = f + g$ with

$$\|f\|_{H^1} \leq CA^k$$
$$\|g\|_{H^0} \leq C \rho^k (\log \frac{1}{\eta})^{\frac{k}{2}}$$

We now set $k$ equal to $\epsilon_0 \log \frac{1}{\eta}$ with $\epsilon_0$ a small positive constant. Then $(S\Gamma)^l 1 = f + g$, where $\|f\|_{H^1}^2 \leq C \eta^{-\frac{k}{2}}$ and $\|g\|_{H^0}^2 \leq C \eta^{\epsilon_1}$. From (17) and (25), we have

$$|g_l(z)| \leq C (\|f\|_{H^1}^2 + \|g\|_{H^0}^2 \sup_r (r^2 e^{-2\pi \eta r^2}))$$
$$\leq C \eta^{\epsilon_1 - 1}$$

completing the proof of Theorem 3. \hfill \Box

Further remarks 1. Proposition 2 can also be used to refine the main estimate in [12] and extend it to Bernoulli distributions. Given Theorem 2, this argument is identical to the corresponding argument in [12], so we will omit details. If we start from formula (16) (or (19) specialized to the case of $\{-l, \ldots, l\}$), multiply it by its complex conjugate and take expectations, then we obtain the following, where $G_l(m, n, z)$ is the Green’s function on $-l, \ldots, l$ (cf. [12]):

$$\mathbb{E}(|G_l(m, n, z)|^2) = 4 \int_{\mathbb{R}^2 \times \mathbb{R}^2} ((TT)^{m-n} \Phi_n)(x, y) \Phi_m(x, y) \Gamma(x, y) dxdy$$

(29)

Here $T$ is the $\mathbb{R}^2 \times \mathbb{R}^2$ Fourier transform defined via

$$Tf(\xi, \eta) = \int_{\mathbb{R}^2 \times \mathbb{R}^2} e^{-2\pi i (\xi \cdot x - \eta \cdot y)} dxdy$$

and we also use $\Gamma$ to denote multiplication by the function $\Gamma$. Also $\Phi_m$ and $\Phi_n$ are functions which are bounded by 1 (see the discussion before and after 19), and therefore
\[ \|e^{-\frac{1}{|x|^2} + |y|^2}}\Phi_n\|_{L^2(\mathbb{R}^2 \times \mathbb{R}^2)} \leq Ct^{-1} \] and similarly with \( \Phi_n \). Theorem 2 with the quadratic form \( Q(x, y) = |x|^2 - |y|^2 \) implies the estimate \( \|(TT)^2\|_{L^2 \to L^2} \leq \rho \), where \( \rho < 1 \) depends on \( \nu \) only.  

Applying this estimate \( \frac{|m-n|}{2} \) times to the function \( f(x, y) = e^{-\pi \eta (|x|^2 + |y|^2)} \Phi_n(x, y) \), we readily obtain

**Theorem 4**  

For any single site distribution \( \nu \) which is not a single dirac mass, there is an estimate \( \mathbb{E}(|G(m, n, z)|^2) \leq Ce^{-\beta^2 \eta^{-2}} \) for the Green’s function on \( \{-l, \ldots, l\} \), uniformly in \( l \) and \( m, n \in \{-l, \ldots, l\} \) with \( |m| \leq \frac{l}{2} \) and \( |n| \geq \frac{l}{2} \) and \( z = E + i\eta \) with \( 0 < \eta \leq 1 \). Here \( \beta \) and \( C \) are fixed positive constants depending on \( \nu \) only.

The usual “multiscale” arguments lead from this and Theorem 3 to a proof of localization, see the end of [12] and [7], Theorem 2.3.

2. More generally, the considerations in remark 1 can be applied on any tree with subexponential growth. Namely, let \( \xi_{m_j} \) be as in (19) and let \( Z_j(x, y) \) be the expectation of the quantity \( \xi_{m_j}(x)\xi_{m_j}(y) \). The functions \( Z_j \) are evidently bounded by 1. If we multiply (19) by its complex conjugate and then take expectations, and use independence, we obtain

\[
\mathbb{E}(|G(m, n, z)|^2) = 4 \int_{\mathbb{R}^2 \times \mathbb{R}^2} Z_j^*(x, y)((\Pi_{j=0}^{d-1} \Gamma \Gamma) Z_j)1(x, y)\Gamma(x, y) dxdy
\]

Since the \( Z_j^* \) are bounded by 1, the estimate

\[
\mathbb{E}(|G(m, n, z)|^2) \leq C\eta^{-2}e^{-\beta d}
\]

now follows from Theorem 2 just as before.

From this we can also conclude

Let \( T \) be a tree with subexponential growth. Consider the Anderson model on \( T \) with a single site distribution which is absolutely continuous with a bounded density. Then the spectrum is almost surely pure point.

---

\( ^5 \) The \( \mathbb{R}^2 \times \mathbb{R}^2 \) Fourier transform differs from the \( \mathbb{R}^4 \) Fourier transform by an orthogonal change of variable, so Theorem 2 is applicable by remark 1 at the end of section 2.
Namely, the assumption on the single site distribution means that Wegner’s estimate is available. Then it follows by 30 and the argument at the end of [12] that the infinite volume Green’s function $G(x_0, \cdot, E)$ decays exponentially, hence is in $L^2$ by the subexponential growth assumption. Then (using absolute continuity again) the theory of rank one perturbations implies the spectrum is pure point.

Molchanov has previously proved some results of this nature.

5. Weak disorder

In this section, we will make some estimates which are uniform in the disorder as the disorder goes to zero. This section was written after the rest of the paper and after a conversation with A. Klein, who pointed out reference [2], especially the norm (31) below.

First consider Theorem 4. Using the quantitative statement in Theorem 2 and the procedure from [12] in the manner indicated above, we immediately obtain the following statement:

**Lemma 5.1** Assume the single site distribution satisfies (3). Then the number $\beta$ in Theorem 4 may be taken to be $C^{-1} \gamma \lambda^2$, where $C$ is a universal constant.

This dependence on $\lambda$ is the expected one in view of known results, e.g. [16],[2] on the Liapunov exponent when the single site distribution is smooth. See also Proposition 5.5 below.

Now consider Theorem 3. One can show that the infinite volume density of states satisfies a Holder estimate which is uniform in the disorder for small disorder, more precisely the following proposition 5.2. Note it applies in particular to dilations of a fixed single site distribution with compact support.

**Proposition 5.2** Suppose $\lambda \leq 1$ and let $\nu$ be a measure which is supported on $\{x : -A\lambda \leq x \leq A\lambda\}$ and satisfies (3). Then, on compact subintervals $I \subset (-1, 1)$ the integrated density of states satisfies $|k(x) - k(y)| \leq C|x - y|^\alpha$ where $C$ and $\alpha$ depend only on $A, \gamma$ and $I$ (and not for example on $\lambda$).

**Proof** We may assume $\lambda$ is small, e.g. $\lambda < A^{-1} \text{dist}(I, \mathbb{R}\setminus\{-1, 1\})$, since otherwise the result follows from Theorem 3. From the form of the statement, we can then assume $\int x d\nu = 0$, else we absorb the first moment into the energy $E$. Now fix $E \in I$. Define a
Hilbert space $H^1_E$ on the half line as follows:

$$
\|f\|_{H^1_E}^2 = \int_{\mathbb{R}^2} |f(|x|^2)|^2 + |Df(|x|^2)|^2 + 2E \text{ im}(f(|x|^2)\overline{Df(|x|^2)})\,dx \\
= (1 - E^2) \int_{\mathbb{R}^2} |f(|x|^2)|^2\,dx + \int_{\mathbb{R}^2} |(Df - iEf)(|x|^2)|^2\,dx 
$$

(31)

It is shown in [2] that the map $f \to S(yEf)$ is an isometry on $H^1_E$.

**Lemma 5.3** The norm of the operator $S\Gamma$ on the space $H^1_E$ is

$$
\|S\Gamma\| \leq 1 + C(\lambda^2 + \eta) 
$$

where $C$ depends on $I$ and $A$ only.

**Proof** Note first of all that $|D\hat{v}|^2 \leq 4A^2\lambda^2$. We also need the following slightly more refined estimate

$$
|D\hat{v}(\xi)|^2 = 4 \int xy \cos(2\pi\xi(x-y))\,d\nu(x)d\nu(y) \\
= 4 \int xy(\cos(2\pi\xi(x-y)) - 1)\,d\nu(x)d\nu(y) \\
\leq 4A^2\lambda^2 \int 1 - \cos(2\pi\xi(x-y))\,d\nu(x)d\nu(y) \\
= 4A^2\lambda^2(1 - |\hat{v}(\xi)|^2) 
$$

(32)

From (33) and calculus we conclude the following: for any constant $B_1$ there is a constant $B_2$ depending on $B_1$ and $A$ such that

$$
|\hat{v}|^2 + B_1(|\hat{v}| |D\hat{v}| + |D\hat{v}|^2) \leq 1 + B_2\lambda^2 
$$

(34)

Next we prove a pointwise inequality. Let $J(z, w) = |z|^2 + |w|^2 + 2E \text{ im}(z\overline{w})$ and note that $J(az, aw + bz) \leq (|a|^2 + C_E(|ab| + |b|^2))J(z, w)$ for any complex numbers $a, b, z, w$. Here the constant $C_E$ remains bounded as long as $E$ remains in a compact subset of $(-1, 1)$. Using the product rule, it follows that for any function $f$,

$$
J(\hat{v}f, D(\hat{v}f)) \leq (|\hat{v}|^2 + C_E|\hat{v}| |D\hat{v}| + C_E|\hat{v}|^2)J(f, Df) \\
\leq (1 + C\lambda^2)J(f, Df) 
$$

where the last line followed from (34) and $C$ depends on $I$ and $A$. Also, it is easy to show that $J(y_{\eta f}, D(y_{\eta f})) \leq (1 + C\eta)J(f, Df)$ so we may conclude that

$$
J(y_{\eta \hat{v}}f, D(y_{\eta \hat{v}}f)) \leq (1 + C(\lambda^2 + \eta))J(f, Df) 
$$

23
By integrating this inequality we obtain \( \|y_n \hat{\nu} f \|_{H^k_1} \leq (1 + C(\lambda^2 + \eta)) \|f\|_{H^k_1} \) and then the result follows by the abovementioned isometry property from [2].

We define \( \Gamma_0 \) like \( \Gamma \), but using the unperturbed laplacian, i.e. \( \Gamma_0(r) = e^{2\pi i zr} (= y_z(r)) \) and we also use \( \Gamma_0 \) to denote the operator of multiplication by \( \Gamma_0 \). We also let \( g^0_n(z) \) be the Green's function of the Laplacian on \( \{ -n, \ldots, n \} \).

\textbf{Lemma 5.4} There are constants \( \kappa > 0 \) and \( C \) which depend only on the interval \( I \), such that the following holds: if \( \eta > 0 \) is small, \( E \in I \), and \( n \) is sufficiently large (depending on \( I \) and a lower bound for \( \eta \)) then \( |(S\Gamma_0)^n1(|x|^2)| \leq e^{-\kappa|x|^2} \) and \( |D(S\Gamma_0)^n1(|x|^2)| \leq Ce^{-\kappa|x|^2}. \)

\textbf{Proof} This basically follows from the fact that the infinite volume free spectral measure is absolutely continuous with smooth density. Details are as follows. Let \( y = (S\Gamma_0)^n1. \) By the remarks in section 3, we know that \( y \) is a Gaussian function, \( y = y_{\zeta_n} \), and the lemma will follow if we can show that \( |\zeta_n| \) is bounded above and \( \text{im} \zeta_n \) is bounded away from zero. For \( n \) large, \( g^0_n \) is a small perturbation of the infinite volume free Green's function, so by the explicit formula for the latter we know that \( |g^0_n| \) is bounded by a constant and \( \text{im} g^0_n \) is bounded below by a positive constant. On the other hand, we have

\[
\begin{align*}
g^0_n &= 2i \int y_{\zeta_n}(|x|^2)^2 \Gamma_0(|x|^2) dx \\
&= \frac{-1}{2\zeta_n + E + i\eta}
\end{align*}
\]

The bounds for \( |g^0_n| \) now imply \( |2\zeta_n + E + i\eta| \) is bounded and bounded away from 0, in particular \( |\zeta_n| \) is bounded. Taking imaginary parts, we may also conclude that \( \text{im} g^0_n \lesssim \eta + 2\text{im} \zeta_n \) and then (since \( \eta \) is small) that \( \text{im} \zeta_n \) is bounded from below. \( \square \)

We now use a modification of the proof of Theorem 3. Fix \( I \); we will show that if \( E \in I \) and if \( n \) is large enough, then \( |g_n(z)| \leq C\eta^{-(1-\alpha)} \). Here \( C \) and \( \alpha \) depend on \( I \) only; the necessary size for \( n \) depends also on \( \eta \). This will suffice for the proof as discussed in section 4.

Using Lemma 5.3 we have the estimate

\[
|(S\Gamma)^k|_{H^1_{k-\alpha} \rightarrow H^1_{k+\alpha}} \leq C e^{C_1 k(\lambda^2 + \eta)}
\]

(35)
Also, using Theorem 2 and (13), we have

$$\|(ST)^k\|_{H^0 \to H^0} \leq Ce^{-C_2k(\lambda^2 + \eta)} \quad (36)$$

Here $C_1$ and $C_2$ are positive constants. Furthermore, since $\nu$ has zero first moment, we have

$$|\hat{\nu}(t) - 1| \leq C\lambda^2 t^2$$

and therefore

$$|\Gamma(|x|^2) - \Gamma_0(|x|^2)| \leq C\lambda^2 |x|^4 \quad (37)$$

Also, from formula (32) we may conclude that $|D\hat{\nu}(\xi)| \leq C\lambda^2 |\xi|$, hence (using the product rule) that

$$|D\Gamma(|x|^2) - D\Gamma_0(|x|^2)| \leq C(|\hat{\nu}'(|x|^2)| + |\hat{\nu}(|x|^2) - 1|) \leq C\lambda^2(|x|^2 + |x|^4) \quad (38)$$

Just as in section 4, it suffices to show that if $n$ is large enough then $(ST)^n1 = f + g$ with $\|f\|_{H^0}^2 \leq C_4\eta^a$ and $\|g\|_{H^1}^2 \leq C_4\eta^{-(1-\alpha)}$, where $C_4$ and $\alpha$ should now be uniform in $\lambda$. Let $n_0$ be large enough that the conclusion of Lemma 5.4 is valid for $n \geq n_0$, and let $n$ be much larger than $n_0$. We may express $(ST)^n1$ as follows:

$$(ST)^n1 = (ST_0)^n1 + \sum_{k=0}^{n-n_0-1} (ST)^kS(\Gamma - \Gamma_0)(ST_0)^{n-k-1}1$$

$$+ \sum_{k=n-n_0}^{n-1} (ST)^kS(\Gamma - \Gamma_0)(ST_0)^{n-k-1}1 \quad (39)$$

Term $I$ has bounded $H^1_E$ norm by Lemma 5.4. Next consider the terms in the sum $II$. Here $n - k - 1 \geq n_0$, so we can use Lemma 5.4. Using Lemma 5.4 and (37) we have

$$\|\Gamma(\Gamma - \Gamma_0)(S\Gamma_0)^{n-k-1}1\|_{H^0} \leq C\lambda^2 \quad (40)$$

and using in addition (38) and the product rule we also have

$$\|\Gamma(\Gamma - \Gamma_0)(S\Gamma_0)^{n-k-1}1\|_{H^1_E} \leq C\lambda^2 \quad (41)$$

25
Hence, by (35) and (36),
\[
\|(S\Gamma)^k(\Gamma - \Gamma_0)(S\Gamma_0)^{n-k-1}\|_{H^0} \leq C\lambda^2 e^{C_1 k(\lambda^2 + \eta)}
\]
(42)
\[
\|(S\Gamma)^k(\Gamma - \Gamma_0)(S\Gamma_0)^{n-k-1}\|_{H^0} \leq C\lambda^2 e^{-C_2 k(\lambda^2 + \eta)}
\]
(43)

If we let \( \alpha \) be a small positive constant and split the sum \( \Pi \) as \( f + g \) where \( f \) is the terms with \( e^{-C_2 k(\lambda^2 + \eta)} \leq \eta^\alpha \), then (using the formula for the sum of a geometric series) we obtain \( \|f\|_{H^0} \leq C\eta^\alpha \) and \( \|g\|_{H^0} \leq C\eta^{-C_3 \alpha} \), where \( C_3 \) is determined by \( C_1 \) and \( C_2 \). This completes the estimation for term \( \Pi \), provided we have taken \( \eta \) small enough. Term \( \Pi_3 \) is simpler: we know that \( |((S\Gamma_0)^{n-k-1}| \leq 1 \) and therefore 37 implies \( \|(\Gamma - \Gamma_0)(S\Gamma_0)^{n-k-1}\|_{H^0} \leq C(\log \frac{1}{\eta})^{\frac{1}{2}} \). Thus by (36) the \( H^0 \) norm of each term in the sum \( \Pi_3 \) may be made arbitrarily small by taking \( n \) large, hence the \( H^0 \) norm of \( \Pi_3 \) may also be made arbitrarily small, since there are just \( n_0 \) terms. So the proof is complete. \( \square \)

**Remark** The preceding argument does not give very good estimates in finite volume, since in Lemma 5.4 one must take \( n \) large compared with \( \frac{1}{\eta} \). Accordingly the result is not directly applicable to localization type problems. However, better estimates may be obtained a posteriori, as we now show.

**Proposition 5.5** In the situation of Proposition 5.2 the Liapunov exponent is \( \geq C^{-1} \lambda^2 \).

**Proof** We will show that the infinite volume green’s function is almost surely bounded by \( Ce^{-C_6 n \lambda^2} \), where \( C \) (but not \( C_6 \)) may depend on \( \lambda \) and the potential. Namely, if we consider the finite volume green’s functions on the diagonal, i.e. \( g_n(z) \) in our previous notation, then we have

\[
|g_{n+1}(z) - g_n(z)| \leq \int_{\mathbb{R}^2} |((S\Gamma)^{n+1})([x]_n)^2 - ((S\Gamma)^n)_1([x]_n)^2| e^{-2\pi \eta |x|^2} dx
\]
\[
\lesssim \int_{\mathbb{R}^2} |((R\Gamma)^{n+1})([x]_n)^2| e^{-2\pi \eta |x|^2} dx
\]
\[
\lesssim \|(R\Gamma)^{n+1}\|_{H^0}(\int_{\mathbb{R}^2} |x|^2 e^{-2\pi \eta |x|^2} dx)\frac{1}{2}
\]
\[
\lesssim \eta^{-1}\|(R\Gamma)^{n+1}\|_{H^0}
\]

The first line is the basic formula (17). The second line is proved as follows: factor \( |((S\Gamma)^{n+1})^2 - ((S\Gamma)^n)_1| = |(S\Gamma)^{n+1} + (S\Gamma)^n_1||((S\Gamma)^{n+1} - (S\Gamma)^n)_1| \). The first factor is
bounded by 2 since \( (ST)^n \) is a synthesis of Gaussians and therefore bounded by 1, and the second factor is \( (RT)^{n+1} \). The third line is the Cauchy-Schwarz inequality and the fourth line follows by calculating the relevant integral. We know by (37) and Lemma 4.2 (which is still valid in the present context, since its proof did not use any assumption on the potential) that \( \| (RT)^{n+1} \|_{H^0} \lesssim e^{-C_2n\lambda^2 (\log \frac{1}{\eta})^2} \). Now, by Proposition 5.2, there is \( \tilde{n} > n \) such that \( |g_\tilde{n}(z)| \leq C\eta^{n-1} \). Thus

\[
|g_n(z)| \leq \sum_{k=n}^{\tilde{n}-1} |g_k(z) - g_{k+1}(z)| + |g_\tilde{n}(z)| \\
\lesssim \sum_{k=n}^{\infty} \eta^{-1} (\log \frac{1}{\eta})^2 e^{-C_2k\lambda^2} + \eta^{n-1} \\
\lesssim \lambda^{-2} e^{-C_2n\lambda^2} \eta^{-1} (\log \frac{1}{\eta})^2 + \eta^{n-1}
\]

and therefore the following statement:

If \( \eta = n^{-B} \) with a fixed large constant \( B \), and if \( n \) is sufficiently large depending on \( \lambda \), then \( \eta |g_n(z)| \leq n^{-20} \).

Consequently, with \( \eta \) as indicated, the probability of a state in the interval \( (E-\eta, E+\eta) \) is \( \lesssim n^{-19} \). We may therefore use Lemma 5.1 and a comparison of \( G_{\{-n, \ldots, n\}}(\cdot, \cdot, E) \) with \( G_{\{-n, \ldots, n\}}(\cdot, \cdot, E+i\eta) \) via the resolvent identity (see the end of [12], for example) to conclude that if \( n \) is sufficiently large then with probability at least \( 1 - Cn^{-17} \) there is an estimate \( |g_{\{-n, \ldots, n\}}(j, k, E)| \leq C_5 e^{-C_6n\lambda^2} \) for all \( j, k \) with \( |j| \leq \frac{n}{3} \) and \( 2 \frac{n}{3} \leq k \leq n \). Here \( C_5 \) is positive and independent of \( \lambda \). From this estimate, a simple version of multiscale analysis leads to an estimate \( |G(0, n, E)| \leq C e^{-C_6n\lambda^2} \) a.s. for the infinite volume Green’s function. \( \square \)

Remark It is also possible to prove that the eigenfunctions decay at the expected rate, namely the following result:

Corollary If \( \nu \) is supported on \( (-A\alpha, A\alpha) \) and satisfies 3, then for compact \( I \subset (-1, 1) \) the following will hold almost surely: if \( u \) is an eigenfunction with eigenvalue \( E \in I \), then there is a constant \( C_u \) such that \( |u(n)| \leq C_u e^{-\alpha \lambda^2 |n|} \). Here \( \alpha \) is a fixed constant depending only on \( \gamma, A \) and \( I \).

This follows from the finite volume density of states estimate in the preceding proof and Lemma 5.1, using the version of multiscale analysis in [7], Theorem 2.3. We omit
the argument, since we will give an almost identical argument in the proof of Proposition 5.6(ii) below.

Once again the proofs in this section did not use stationarity, and they can also be applied when the disorder parameter $\lambda$ depends on the site. For example, one can prove the following.

**Proposition 5.6** Let $\{\lambda_j\}_{j=-\infty}^{\infty}$ be a bounded sequence of nonnegative numbers and assume the following condition: there is a constant $C$ such that $C^{-1} \lambda_{2k} \leq \lambda_{2k+1} \leq C \lambda_{2k}$ for all integers $k$. Let $\nu$ be a probability measure on $\mathbb{R}$ with compact support and zero first moment and define $d\nu_j(x) = d\nu(\lambda_j^{-1}x)$ (if $\lambda_j = 0$, then $\nu_j$ is a Dirac mass at the origin), and consider the Anderson model with single site distribution at $j$ given by $\nu_j$. Then for $E \in (-1, 1)$

(i) The density of states is locally Hölder continuous.

(ii) If in addition the following hold

\[ \lim_{n \to \infty} \sum_{\frac{n}{2} \leq j \leq n} \lambda_j^2 = \infty \quad (45) \]

\[ \lim_{n \to \infty} \sum_{-n \leq j \leq -\frac{n}{2}} \lambda_j^2 = \infty \quad (46) \]

then the spectrum is pure point.

**Remarks** 1. Related results are proved in [6], although the assumptions there are more restrictive. Conditions for absolutely continuous spectrum are also given in [6] and imply the assumptions (45), (46) are close to optimal.6

2. An interesting case occurs when $\lambda_j = 1$ for $j \in \Lambda$ and zero otherwise, where $\Lambda \subset \mathbb{Z}$. Then the assumptions in (ii) mean that $2k \in \Lambda \iff 2k + 1 \in \Lambda$ and that the intersections of $\Lambda$ with dyadic intervals $\{\frac{n}{2}, \ldots, n\}$ or $\{-n, \ldots, -\frac{n}{2}\}$ have large cardinality for large $n$.

**Proof** The proof of part (i) is of course a modification of the proof of Proposition 5.2. Namely, define $\Gamma^j$ like $\Gamma$ but using the measure $\nu_j$. $\Gamma_0$ is defined as before i.e. corresponds to the case of no potential. Then formula (17) should be replaced by the following

\[ \mathcal{E}(G_{-n, \ldots, n}(0, 0, z)) = 2i \int_{\mathbb{R}^2} (\Pi_{j=0}^{n-1} (\Sigma^{n-j}) 1(|x|^2))(\Pi_{j=0}^{n-1} (\Sigma^{n+j}) 1(|x|^2)) \Gamma^0(|x|^2) dx \quad (47) \]

6Actually, we only need an averaged version of (45), 46, cf. (51 below).
(39) should be replaced by the following
\[
(\Pi_{j=0}^{n-1} S T^{n-j}) 1 = (S T_0)^n 1 + \sum_{k=0}^{n-n_0-1} (\Pi_{j=0}^{k-1} S T^{k-j}) S (T^{k+1} - T_0) (S T_0)^{n-k-1} 1
\]
\[
+ \sum_{k=n-n_0}^{n-1} (\Pi_{j=0}^{k-1} S T^{k-j}) S (T^{k+1} - T_0) (S T_0)^{n-k-1} 1
\]
and (35) and (36) should be replaced by
\[
\| (\Pi_{j=0}^{k-1} S T^{k-j}) \|_{H^1_k \to H^1_k} \leq C e^{C_1 \Sigma_j \leq k (\lambda_j^2 + \eta)}
\]
(48)
\[
\| (\Pi_{j=0}^{k-1} S T^{k-j}) (T^{k+1} - T_0) (S T_0)^{n-k-1} 1 \|_{H^1_k} \leq C \lambda_k^2 e^{C_1 \Sigma_j \leq k (\lambda_j^2 + \eta)}
\]
(49)
Then (42) and (43) are replaced by
\[
\| (\Pi_{j=0}^{k-1} S T^{k-j}) (T^{k+1} - T_0) (S T_0)^{n-k-1} 1 \|_{H^1_k} \leq C \lambda_k^2 e^{-C_2 \Sigma_j \leq k (\lambda_j^2 + \eta)}
\]
We used the assumption about \( \lambda_{2k} \) and \( \lambda_{2k+1} \) being comparable in order to derive (49) from Theorem 2, and we also used that the \( \lambda_j \) are bounded and \( \eta \) small. Now let \( \alpha \) be a small positive number. One can easily check that \( \Sigma_{k \leq m} \lambda_{k+1}^2 e^{C_1 \Sigma_j \leq k (\lambda_j^2 + \eta)} \leq C e^{C_1 \Sigma_j \leq m (\lambda_j^2 + \eta)} \), \( \Sigma_{k \geq m} \lambda_{k+1}^2 e^{-C_2 \Sigma_j \leq k (\lambda_j^2 + \eta)} \leq C e^{-C_2 \Sigma_j \leq m (\lambda_j^2 + \eta)} \), and we may therefore conclude the following assertion for a suitable constant \( C_7 \):
\[
\Sigma_{k \leq m} \lambda_{k+1}^2 e^{C_1 \Sigma_j \leq k (\lambda_j^2 + \eta)} \geq \eta^{-C_7 \alpha} \text{ implies } \Sigma_{k \geq m} \lambda_{k+1}^2 e^{-C_2 \Sigma_j \leq k (\lambda_j^2 + \eta)} \leq \eta^\alpha
\]
Hence if we define \( m \) to be the smallest number for which the former inequality fails then we may repeat the argument following 43, splitting the sum \( \Pi \) according to whether \( k \) is larger or smaller than \( m \). The conclusion is that if \( \alpha \) is small enough then for large positive \( n \)
\[
(\Pi_{j=0}^{n-1} S T^{n-j}) 1 = f + g
\]
with \( \| f \|_{H^1_k} \leq \eta^\alpha \) and \( \| g \|_{H^1_k} \leq \eta^{-1-\alpha} \). The analogous statement for negative \( n \) is of course also true and the result now follows using formula 47.

Now we prove (ii). Given the preceding estimates this is a routine adaptation of the argument in [7]. Since generalized eigenfunctions have polynomial growth it suffices to show the
Claim If $\gamma$ is a compact subinterval of $(-1, 1)$ and $r < \infty$ then with probability 1 the following is true: if $u : \mathbb{Z} \rightarrow \mathbb{R}$ satisfies $(H - E)u = 0$ for some $E \in \gamma$, and if $|u(0)| + |u(1)| = 1$ and $|u(n)| \leq |n|^r$ when $|n| \geq 1$, then $u \in L^2$.

We need a calculus lemma.

**Lemma 5.7** For any $r < \infty$ there is $\epsilon > 0$ making the following true. Assume that $u : \mathbb{Z} \rightarrow \mathbb{R}$ satisfies $|u(n)| \leq n^r$ when $n \geq 1$ and

$$|u(n)| \leq n^{-\alpha} \max_{j \in [\sqrt{n}, n^2]} |u(j)|$$

(50)

for all sufficiently large $n$. Then $u \in L^2$.

**Proof of Lemma 5.7** We can take $\alpha$ to be any number such that $\alpha > r$. With $\alpha$ chosen this way we will show that under the hypotheses of the lemma $|u(n)| \leq C n^{-2r}$ for a suitable constant $C$. This will suffice, since we may obviously assume that $r > \frac{1}{2}$.

Fix $N$ large enough that (i) (50) is valid when $n \geq N$ and (ii) $N^{-2r} \geq 2$, and choose $C$ to be such that $CN^{-2r} > N^r$. Now assume there is a “bad” site $n_0$, i.e. $|u(n_0)| > Cn_0^{-2r}$. Then $n_0 \geq N$ by choice of $C$ so we may apply (50). We conclude there is $n_1 \in [\sqrt{n_0}, n_0^2]$ such that $|u(n_1)| \geq n_0^\alpha |u(n_0)|$. This implies in particular that $|u(n_1)| \geq Cn_0^{-2r} \geq Cn_0^{-r} \geq Cn_1^{-2r}$, i.e. $n_1$ is another bad site. Also $\frac{|u(n_1)|}{n_1^\alpha} \geq n_0^\alpha \frac{|u(n_0)|}{n_0^\alpha} \geq n_0^\alpha \frac{|u(n_0)|}{n_0^\alpha} \geq n_0^\alpha \frac{|u(n_0)|}{n_0^\alpha} \geq 2^{i+1}$.

Iterating this construction, we obtain a sequence of bad sites $\{n_j\}$ with $|u(n_j)| \geq 2^{j} |u(n_0)|$, clearly a contradiction. \(\square\)

Fixing $\gamma$ and $k$, we let $\tilde{\gamma}$ be another compact subinterval of $(-1, 1)$ containing $\gamma$ in its interior, $m$ a suitable large constant and let $B >> m$. Fix a site $n \in \mathbb{Z}$ with $|n| > 100$; we may assume $n > 0$. Let $I^0_n$ be the interval $\{-\sqrt{n}, \ldots, \sqrt{n}\}$ and let $I_n$ be the interval $\{\sqrt{n} + 1, \ldots, n^2\}$. Let $g^n_0$ and $g_n$ be the Green’s functions on $I^0_n$ and $I_n$. The hypothesis 45 implies that

$$\lim_{n \rightarrow \infty} \frac{n^{1/n}}{\log n} = \infty \quad (51)$$

$$\lim_{n \rightarrow \infty} \frac{n}{\log n} = \infty \quad (52)$$

We can estimate

$$\|\Pi_{j=0}^{n-1} 1 - (\Pi_{j=0}^{n} S T^{n-1-j}) 1\|_{H_0} = \|\Pi_{j=0}^{n} R T^{n+1-j}) 1\|_{H_0} \leq C_B n^{-B} (\log \frac{1}{\eta})^{1/2} \quad (53)$$
for any given $B$ by inserting 51 into the formula 49. A straightforward modification of the proof of (44) using (53) and part (i) therefore leads to the following estimate:

$$|g^0_n(0, 0, z)| \lesssim n^{-100} \eta^{-1} (\log \frac{1}{\eta})^\frac{B}{2} + \eta^{\alpha - 1}$$

with $\alpha$ being the Holder exponent from part (i). Now let $\sigma_n$ and $\sigma^0_n$ be the spectra of $H$ on $I_n$ and $I^0_n$ respectively. For $E \in \hat{\gamma}$, let $P(n, E)$ be the probability that the distance from $E$ to $\sigma^0_n$ is less than $n^{-m}$. If $m$ has been chosen large enough then the preceding estimate for $g^0_n$ implies $P(n, E) \lesssim n^{-20}$. Since $\sigma_n$ has cardinality $\leq n^2$, the following statement is now immediate by conditioning on $\sigma_n$: let $p(n)$ be the probability that $\text{dist}(\sigma_n \cap \hat{\gamma}, \sigma^0_n \cap \hat{\gamma}) \leq n^{-m}$. Then $p(n) \lesssim n^{-18}$.

We now argue as in the proof of Theorem 4 or Lemma 5.1, using the generalization of formula 29 analogous to 47 together with (51), (52). We conclude that for any fixed $B$

$$\mathbb{E}(|g^0_n(0, k, E + i\eta)|^2) \lesssim \eta^{-2} n^{-B}$$

$$\mathbb{E}(|g_n(n, k, E + i\eta)|^2) \lesssim \eta^{-2} n^{-B}$$

provided $k$ is close enough to the boundary of $I^0_n$ or $I_n$ respectively. Using the resolvent identity to compare e.g. $g^0_n(0, \cdot, E)$ with $g^0_n(0, \cdot, E + i\eta)$, $\eta \approx n^{-m}$, it follows that with probability at least $1 - n^{-17}$ we have either

$$|g^0_n(0, k, E)| \lesssim n^{-\frac{B}{2}}$$

for all $k$ close to the boundary of $I^0_n$, or

$$|g_n(n, k, E)| \lesssim n^{-\frac{B}{2}}$$

for all $k$ close to the boundary of $I_n$, where the two possibilities correspond to the two cases $\text{dist}(E, \sigma^0_n) \geq \frac{1}{2} n^{-m}$ or $\text{dist}(E, \sigma_n) \geq \frac{1}{2} n^{-m}$. Now the rest of the argument is the same as the relevant part of the proof of Theorem 2.3 in [7]. Taking $B$ large compared with $r$, it follows that 54 is impossible for large $n$ (it would imply $|u(0)| + |u(1)| \lesssim n^{-\frac{B}{2}} \max_{I_n} |u| < 1$), and then 55 leads to a bound of the form

$$|u(n)| \leq C n^{-\frac{B}{2}} \max_{I_n} |u|$$

(56)

Thus we have shown the following: if $n$ is fixed and sufficiently large then with probability at least $1 - n^{-17}$ we have a bound (56) for all $u$ satisfying the hypotheses of the claim. By Borel-Cantelli, with probability 1 we have such a bound for all $u$ satisfying the hypotheses of the claim and all sufficiently large $n$. We are now done by Lemma 5.7. \[\Box\]
References


