# Anderson Localization for Bernoulli and Other Singular Potentials

Rene Carmona*, Abel Klein** and Fabio Martinelli***

Department of Mathematics, University of California, Irvine, CA 92717, USA

**Abstract.** We prove exponential localization in the Anderson model under very weak assumptions on the potential distribution. In one dimension we allow any measure which is not concentrated on a single point and possesses some finite moment. In particular this solves the longstanding problem of localization for Bernoulli potentials (i.e., potentials that take only two values). In dimensions greater than one we prove localization at high disorder for potentials with Hölder continuous distributions and for bounded potentials whose distribution is a convex combination of a Hölder continuous distribution with high disorder and an arbitrary distribution. These include potentials with singular distributions.

We also show that for certain Bernoulli potentials in one dimension the integrated density of states has a nontrivial singular component.

## 1. Introduction

The Anderson tight-binding model is given by the random Hamiltonian $H = -\Delta + v$ on $l^2(\mathbb{Z}^\nu)$, where $(-\Delta)(x, y) = -1$ if $|x - y| = 1$ and zero otherwise, and $v(x), x \in \mathbb{Z}^\nu$, are independent identically distributed random variables with common probability distribution $\mu$.

In one dimension the spectral properties of this random Hamiltonian have been fairly well understood for absolutely continuous $\mu$ with a "nice" density (see for example the year old reviews [1, 2]). In higher dimensions the first results towards localization were due to Fröhlich and Spencer [3]. These were followed by a proof of localization for a hierarchical version of $H$ by Jona–Lasinio, Martinelli and Scoppola [4] based on the methods of [3], and by a short proof of the absence of absolutely continuous spectrum for Anderson's model at high disorder or low energy given by Martinelli and Scoppola [5], again based on [3]. Subsequently proofs of localization for the multidimensional Anderson model in the same region

of the parameters were given by Frohlich, Martinelli, Scoppola and Spencer [6], Delyon, Levy and Souillard [7], and Simon and Wolff [8], again using the results of [3].

In all these papers the potential probability distribution $\mu$ was required to have a bounded density; the inverse of the bound of the density being taken as a measure of the disorder.

In the one dimensional case Delyon, Levy and Souillard [9] and Simon and Wolff [8] were able to extend their proof to any $\mu$ with a nontrivial absolutely continuous component.

The approach followed in [7, 9] and [8] uses the results of [3] and ideas introduced by Carmona for the one dimensional case and subsequently developed and exploited by Kotani [10]. In [6] the work of Frohlich and Spencer [3] is substantially improved following the pattern of the proof of localization for the hierarchical model [4].

In this paper we considerably relax the assumptions on the potential probability distribution $\mu$. In one dimension we allow any probability measure that is not concentrated on a single point and possesses some finite moment. In particular, we solve the longstanding problem of localization for Bernoulli potentials (i.e., potentials that take only two values). For arbitrary dimension we prove localization at high disorder for potentials with Hölder continuous distributions and for bounded potentials whose distribution is a convex combination of a Hölder continuous distribution with high disorder and a small amount of an arbitrary distribution. These include potentials with singular distributions.

We also show that for certain Bernoulli potentials in one dimension the integrated density of states has a nontrivial singular component. This had been conjectured by Simon and Taylor [11].

We thus need a way of proving localization that does not depend on the absolute continuity of $\mu$ (or of a component of $\mu$). A close look at the proof in [6] reveals that the absolute continuity of $\mu$ is *only* used to prove a key probabilistic estimate needed to control, in a probabilistic sense, the small divisors appearing in the finite volume Green's functions. If $\mu$ has a bounded density, this estimate is easily obtained from Wegner's [12] uniform bound on the expectation of the density of states in a finite volume by a simple use of Chebyshev's inequality [3]. But in [7] and [8], the absolute continuity of $\mu$, in addition to having been used in [3], is *also* used, and in a fundamental way, in order to implement the so-called Kotani's trick.

In this paper we prove that key probabilistic estimate for a large class of potential distributions. When the potential distribution is not absolutely continuous, this estimate is very subtle and requires a fine analysis of the finite volume Hamiltonians corresponding to typical configurations of the potential.

Such an estimate has also been proved in one dimension by Klein, Martinelli and Perez [13] under hypotheses on $\mu$ that allow for singular continuous distributions. They use the methods of Campanino and Klein [14] to prove a uniform bound on the expectation of the Green's function in a finite volume from which the desired estimate follows by Chebyshev's inequality. Thus they also obtain localization for some singular potential distributions.

This paper is organized as follows: In Sect. 2 and 3 we state our results for the one

dimensional and for the multidimensional case, respectively. In Sect. 4 we briefly review the strategy of [6] and recall the key probabilistic estimates. We also state the probabilistic estimates we prove, which, combined with the methods of [6], give our results on localization stated in Sect. 2 and 3. The proofs for the one dimensional case are in Sect. 5. This section contains an Appendix where LePage's result [15] on the local Hölder continuity of the integrated density of states is extended to the case where the potential is no longer assumed to be bounded. Section 6 contains the proofs for the multidimensional case. It includes an extension of Wegner's bound on the density of states. In Sect. 7 we show that some of our estimates fail for the Anderson model on the Cayley tree and briefly discuss the consequences in the case when the potential $v$ can be either zero or infinite (quantum percolation).

## 2. Statement of Results in the One Dimensional Case

Our result on localization is

**Theorem 2.1.** *Let $v = 1$. Suppose the support of $\mu$ is not concentrated in a single point and $\int |v|^\eta d\mu(v) < \infty$ for some $\eta > 0$. Then the spectrum of $H$ is pure point with probability one and the corresponding eigenfunctions are exponentially localized.*

Some comments are in order. In the one dimensional case an important ingredient in the Fröhlich, Martinelli, Scoppola and Spencer method [6] is Furstenberg's theorem, which gives the strict positivity of the Lyapunov exponent for probability distributions $\mu$ like in Theorem 2.1. This strict positivity gives the initial step in the inductive construction of singular sets which was introduced in [3]. The lack of a theorem a la Furstenberg in higher dimensions is the reason for the extra assumptions of higher disorder or low energy. This strict positivity of the Lyapunov exponent was generally considered to be sufficient for the almost sure existence of a dense set of eigenvalues with exponentially decaying eigenfunctions and no other type of spectrum. In particular, as early as 1974 Pastur [16] gave a simple argument to show that it implied the non-existence of absolutely continuous spectrum for almost every potential. Nevertheless, as we have already mentioned, the proofs of nonexistence of singular continuous spectrum and the exponential decay of the eigenfunctions required the absolute continuity of $\mu$ [17, 18, 6, 7, 8, 1, 2] (or of a nontrivial component of $\mu$ [8, 9]).

The situation is even worse in the continuous case of the Schrödinger operator, i.e., $H = -\Delta + v$ on $L^2(\mathscr{R})$. Indeed, the independence assumption on the potential does not have an analog unless the potential is taken to be piecewise constant. Exponential localization has been proved under very restrictive conditions (see the review [1]). Surprisingly, the situation is somewhat better for the problem on the half-line. Kotani [10] has proved exponential localization for almost every boundary condition at the endpoint if the Lyapunov exponent is almost everywhere positive. But the situation for the whole line, or for the half-line with a preassigned boundary condition is not so simple.

Consequently the case of potentials with singular distributions was a challenging open problem. The special case of Bernoulli potentials, i.e., $\mu = p\delta(v - a) + (1 - p)\delta(v - b)$ with $0 < p < 1$, $a, b \in \mathscr{R}$, $a \neq b$, was particularly puzzling. In the

continuous case, when the potential $v(x)$ is taken to be constant on the intervals $(n, n + 1)$, and the $v(n)$, $n \in \mathbb{Z}$, are independent identically distributed Bernoulli random variables, the solution is known. A proof was first given by Carmona [19] for the part of the spectrum above the supremum of the potential. This proof was completed by Kotani [20] (see the comments in [1]), who also pointed out that the proof applied to the whole spectrum because the change of variables which is crucial in the proof can be performed when the energy is above the potential in a bounded interval. Since this always happens with nonzero probability, an ergodicity argument gives the exponential localization for the whole spectrum. But this proof cannot be extended to the lattice case.

    After proving Theorem 2.1 and thus obtaining exponential localization for Bernoulli potentials, we take on the problem of the regularity properties of the corresponding integrated density of states and prove a conjecture of Simon and Taylor [11], who gave a very appealing heuristic argument.

**Theorem 2.2.** *Let* $v = 1$, $\mu = p\delta(v) + (1 - p)\delta(v - b)$, *with* $0 < p < 1$ *and* $b > 0$. *Then the integrated density of states has a nontrivial singular component for b large enough.*

## 3. Statement of Results in the Multidimensional Case

We will say that a probability measure $\mu$ is Hölder continuous of order $\rho > 0$ if

$$[\delta_\rho(\mu)]^{-1} \equiv \inf_{\tau > 0} \sup_{|b - a| \leq \tau} |b - a|^{-\rho} \mu([a, b]) < \infty.$$

    In this case we will call $\delta_\rho(\mu)$ the $\rho$-disorder of $\mu$.

    If $\hat{\mu} \in L^p$ for some $1 \leq p < \infty$, it follows that $\mu$ is Hölder continuous of order $1/p$ and $[\delta_{1/p}(\mu)]^{-1} \leq K_p \|\hat{\mu}\|_p$ for some constant $K_p < \infty$ which depends only on $p$. If $1 \leq p \leq 2$, $K_p \leq 1$. In particular, if $\mu$ is absolutely continuous with a density in $L^q$, $1 < q \leq \infty$, it follows that $\mu$ is Hölder continuous of order $1/q'$, where $1/q + 1/q' = 1$, and $[\delta_{1/q'}(\mu)]^{-1} \leq \|d\mu/dv\|_q$. Thus if $\mu$ has a bounded density our notion of disorder coincides with the one used by Fröhlich and Spencer [3].

    It is worth mentioning that there are singular continuous $\mu$ which are Hölder continuous of some order $\rho > 0$ [21, 22].

    We will prove for such $\mu$ an estimate similar to the one obtained from Wegner's bound on the density of states for $\mu$ with a bounded density. As it will be explained in Sect. 4 this will imply that the results of [3, 6] hold with the hypothesis of $\mu$ being absolutely continuous with a bounded density replaced by $\mu$ Hölder continuous of some order $\rho > 0$. In particular we will have

**Theorem 3.1.** *Let* $v$ *be arbitrary and suppose* $\mu$ *is Hölder continuous of order* $\rho > 0$, *i.e.,* $\delta_\rho(\mu) > 0$. *Then the spectrum of H is pure point with exponentially localized eigenfunctions, with probability one, provided* $\delta_\rho(\mu)$ *is large enough.*

    We will also prove localization for $\mu$ which contain an arbitrary component.

**Theorem 3.2.** *Let* $v$ *be arbitrary. Suppose* $\mu = p\mu_1 + (1 - p)\mu_2$, *where* $0 < p \leq 1$ *and* $\mu_1$, $\mu_2$ *are probability measures with compact support such that* $\mu_1$ *is Hölder continuous of order* $\rho > 0$ *and* $\mu_2$ *is otherwise arbitrary. Then the spectrum of H is pure point, with exponentially localized eigenfunctions, with probability one, provided* $\delta_\rho(\mu_1)$ *is large enough and p is sufficiently close to one.*

## 4. The Road Map to Localization

In this section we will briefly review some of the main aspects of the proof of Anderson localization given by Fröhlich, Martinelli, Scoppola and Spencer [6], state the probabilistic estimates we prove, and show how Theorems 2.1, 3.1 and 3.2 then follow.

For each positive integer $l$ we will denote by $B_l$ the cube in $\mathbb{Z}^\nu$ centered at the origin with sides of length $2l$ parallel to the coordinate axes. By $\partial B_l$ we will denote its (inner) boundary, i.e., $\partial B_l = \{y \in B_l \text{ such that there exists } x \in B_l^c \text{ such that } |x - y| = 1\}$. For each energy $E$ (real or complex) $G_{B_l}(E; x, y)$ will denote the Green's function of the restriction $H_{B_l}$ of $H = -\Delta + v$ to $l^2(B_l)$ i.e.,

$$G_{B_l}(E; x, y) = (H_{B_l} - E)^{-1}(x, y)$$

for $x, y \in B_l$.

The basic estimates needed in [6] in order to prove localization in a neighborhood of an energy $E_0$ are:

(I) *Let $E_0 \in \sigma(H)$. There exist*

$$a = a(E_0) < 1, \quad \delta_l = \delta(l, E_0) > 0, \quad 0 \le \varepsilon_l = \varepsilon(\nu, l, E_0) \le 1,$$

*such that*

$$\mathbf{P}\{ \sum_{y \in \partial B_l} |G_{B_l}(E; x, y)| < a \text{ for all } |x| \le l/2 \} \ge 1 - \varepsilon_l \tag{5.1}$$

*for all $E \in [E_0 - \delta_l, E_0 + \delta_l]$, all $l$, and $\varepsilon_l \to 0$ as $l \to \infty$.*

(II) *Let $E_0 \in \sigma(H)$. There exists $\tau = \tau(E_0) > 0$ such that for any $0 < \beta < 1$ and $\sigma > 0$ there exists $l_0 = l_0(E_0, \beta, \sigma)$ and $\alpha = \alpha(E_0, \beta, \sigma) > 0$ such that*

$$\mathbf{P}\{\text{dist}(E, \sigma(H_{B_l})) \le e^{-\sigma l^\beta}\} \le e^{-\alpha l^\beta}$$

*for all $|E - E_0| \le \tau$ and $l \ge l_0$.*

In the one dimensional case (I) is a consequence of Furstenberg's theorem (see [6]; see also Klein, Martinelli and Perez [13] where (I) is derived from a bound on the averaged squared modulus of the Green's function obtained by a rigorous replica trick method).

In higher dimensions the proof of (I) is usually quite hard. The work of Fröhlich and Spencer [3] shows that, using (II), there exists $\varepsilon_0(\nu) > 0$ such that if $\varepsilon(\nu, l, E_0) < \varepsilon_0(\nu)$ for some $l_0$, then $\varepsilon(\nu, l, \varepsilon_0) \le Cl^{-k}$ for some constant $C$ and $k > 0$ as $l \to \infty$. The number $a = a(E_0) < 1$ can be even chosen depending on $l$ as $a = a(l) = e^{-ml}$ for some $m > 0$. In their work [3] they show that for large disorder the starting scale $l_0$ can be chosen equal to one, if we have a slightly stronger version of (II):

$$\mathbf{P}\{\text{dist}(E, \sigma(H_{B_l})) \le e^{-\sigma l^\beta}\} \le \delta^{-1} e^{-\alpha l^\beta}$$

where $\delta$ depends on the disorder and $\delta \to \infty$ as the disorder goes to $\infty$, and $\alpha$ is independent of the disorder. For the case of low energy this has been extended by Holden and Martinelli [23], who showed that for $E_0$ close to the left edge of the spectrum of $H$, provided the latter is finite, the scale $l_0$ can be chosen proportional to $[E_0 - \inf \sigma(H)]^{-1/2}$.

The main result of [6] is:

*If* (I) *and* (II) *are satisfied for some* $E_0 \in \sigma(H)$, *then in a neighborhood of* $E_0$ *the spectrum of* $H$ *is pure point with exponentially localized eigenfunctions with probability one.*

In one dimension it also follows from [6] that the rate of decay of the eigenfunctions with energy in $[E_0 - \delta, E_0 + \delta]$ can be bounded from below by $\frac{1}{2}\gamma(E_0)$, provided $\delta$ is small enough. Here $\gamma(E_0)$ is the Lyapunov exponent for $E_0$. This fact will be used in Sect. 5 in the proof of Theorem 2.2.

The strategy of the Fröhlich, Martinelli, Scoppola and Spencer's proof is as follows. Using (I), the lattice $l\mathbb{Z}^\nu$ is decomposed into two disjoint subsets:

$$l\mathbb{Z}^\nu = S_0(E_0, \nu) \cup S_0(E_0, \nu)^c$$

such that for all blocks $B_l(x) = x + B_l$, $x \in S_0^c$, the event described in (5.1) for $x = 0$ occurs. Since $\varepsilon(\nu, l, E_0) \to 0$ as $l \to \infty$, it follows that for $l$ large enough the set $S_0(E_0, \nu)$ *does not contain* an infinite cluster of nearest neighbors sites with probability one. Using simple perturbation expansions closely related to the random walk expansions for the Green's function of $-\Delta + \nu$ on $l^2(\mathbb{Z}^\nu)$ (see [3]) it is shown that, for $l$ large enough, the Green's function of the restriction of $H$ to the union of an arbitrary number of blocks $B_l(x)$ with $x \in S_0(E_0, \nu)^c$ decays exponentially *uniformly* in $E \in [E_0 - \delta(l, E_0), E_0 + \delta(l, E_0)]$. This in turn implies that any generalized eigenfunction (i.e., any polynomially bounded solution of $(H - E)\varphi = 0$) with energy in this interval is exponentially localized on the blocks $B_l(x)$ with $x \in S_0(E, \nu)$. In order to control the global localization properties of such eigenfunctions one has therefore to study the tunneling among the blocks in $S_0(E_0, \nu)$. This is the heart of the proof. Using (II) it is shown that with probability one the configurations of the random potential satisfy certain *non-resonance* conditions which are sufficient to prevent tunneling over too large scales. The basic *deterministic* argument behind the analysis of tunneling and which illustrate the mechanism leading to localization is the following (see [2]):

Let $A, B$ be two subsets of $\mathbb{Z}^\nu$ such that
(i) $A \cup B \subset \Lambda$, dist $(A, B) \geq L$;
(ii) (diam $A$) $\vee$ (diam $B$) $\leq L^{4/5}$;
(iii) for every $E \in [E_0 - \delta, E_0 + \delta]$, $\delta > 0$, $|G_{\Lambda \backslash A \cup B}(E; x, y)| \leq e^{-m|x-y|}$ with $m > 0$;
(iv) dist $(\sigma(H_A), \sigma(H_B)) \geq e^{-\sqrt{L}}$.

Then if there exists an eigenvalue of $H_A$ in $[E_0 - \delta, E_0 + \delta]$ the corresponding eigenfunctions are exponentially localized *either* in $A$ or in $B$. Using (II) with $\beta = \frac{1}{2}$, $\sigma = 1$ one finds that (iv) occurs with probability greater than $1 - e^{-\alpha\sqrt{L}}$.

It should be emphasized that (II) is the only place in the proof where the smoothness of the potential probability distribution $\mu$ played a role. It was originally derived in [3] from a result of Wegner [12] that required $\mu$ to be absolutely continuous with a bounded density.

In this paper we will prove (II) for a wide class of potential probability distributions $\mu$. In view of the previous discussion Theorems 2.1, 3.1 and 3.2 will then follow from Theorems 4.1, 4.2 and 4.3, respectively.

**Theorem 4.1.** *Let $v = 1$. Suppose the support of $\mu$ is not concentrated in a single point and $\int |v|^\eta d\mu(v) < \infty$ for some $\eta > 0$. Let $I$ be a compact interval. For any $0 < \beta < 1$ and $\sigma > 0$ there exists $l_0 = l_0(I, \beta, \sigma)$ and $\alpha = \alpha(I, \beta, \sigma) > 0$ such that*

$$\mathbf{P}\{\operatorname{dist}(E, \sigma(H_{B_l})) \leqq e^{-\sigma l^\beta}\} \leqq e^{-\alpha l^\beta}$$

*for all $E \in I$ and $l \geqq l_0$.*

**Theorem 4.2.** *Let $v$ be arbitrary and suppose $\mu$ is Hölder continuous of order $\rho > 0$ i.e., $\delta_\rho(\mu) > 0$. Then for $0 < \delta < \delta_\rho(\mu)$ we can find $\eta = \eta(\mu, \delta) > 0$ such that*

$$P\{\operatorname{dist}(E, \sigma(H_{B_l})) \leqq \varepsilon\} \leqq \delta^{-1} 2^\rho (2l + 1)^{v(1 + \rho)} \varepsilon^\rho$$

*for all $E \in \mathbb{R}$ and for all $\varepsilon > 0$ and $l$ such that $\varepsilon(2l + 1)^v < \eta$.*

**Theorem 4.3.** *Let $v$ be arbitrary. Suppose $\mu = p\mu_1 + (1 - p)\mu_2$, where $0 < p \leqq 1$ and $\mu_1, \mu_2$ are probability measures with compact support such that $\mu_1$ is Hölder continuous of order $\rho > 0$ and $\mu_2$ is otherwise arbitrary. Let $0 < \delta < \delta_\rho(\mu_1)$ and suppose $\varepsilon_l > 0$ are such that $\varepsilon_l^a(2l + 1)^v < 1$ for some $0 < a < \frac{1}{2}$ and all $l$ large enough. Let us fix $E_0 > 0$. Then for all $|E| \leqq E_0$ there exists $0 \leqq p(v) \leqq 1$ such that for all $l$ large enough we have*

$$P\{\operatorname{dist}(E, \sigma(H_{B_l})) \leqq \varepsilon_l\} \leqq \delta^{-1} 2^\rho (2l + 1)^{v(1 + \rho)} \varepsilon_l^\alpha + \varepsilon_l^d$$

*for some $\alpha = \alpha(\delta, \rho, a) > 0$ and $d = d(p) > 0$ for $p(v) < p \leqq 1$, where $p(1) = p(2) = 0$ and $0 < p(v) < 1$ for $v \geqq 3$. Moreover $d(p) \to \infty$ as $p \to 1$.*

Theorem 4.1 is proven in Sect. 5 and Theorems 4.2 and 4.3 are proven in Sect. 6.

## 5. The One Dimensional Case

Most of the problems encountered in the one dimensional case can be reformulated in terms of properties of random matrices. Indeed, if we consider the eigenvalue problem

$$-u(n + 1) - u(n - 1) + v(n)u(n) = Eu(n) \tag{5.1}$$

for a fixed energy $E$, the solution for $n \geqq 0$ can be expressed by means of the propagator or transfer matrix

$$T_n^{(E)} = \begin{bmatrix} v(n) - E & -1 \\ 1 & 0 \end{bmatrix} \tag{5.2}$$

via the formula

$$\begin{bmatrix} u(n + 1) \\ u(n) \end{bmatrix} = T_n^{(E)} \cdots T_1^{(E)} \begin{bmatrix} u(1) \\ u(0) \end{bmatrix}, \tag{5.3}$$

with a similar formula for $n \leqq 0$.

Since we restrict ourselves to random potentials given by a sequence $v(n), n \in \mathbb{Z}$, of independent identically distributed random variables, the sequence $T_n^{(E)}, n \in \mathbb{Z}$, of random matrices is also independent identically distributed. The asymptotic properties of products like those appearing in (5.3) have been known for a long time.

In particular, for each fixed energy $E$, the subadditive ergodic theorem tells us that

$$\lim_{n \to \pm \infty} \frac{1}{|n|} \log \| T_n^{(E)} \cdots T_1^{(E)} \|$$

exists almost surely and defines a non-random number $\gamma(E)$ called the Lyapunov exponent of the problem (5.1). Moreover, Oceledec multiplicative ergodic theorem gives, for each energy $E$, the almost sure existence of a (random) unit vector $\mathbf{x}_{E,v}^{+;-}$ in $\mathbb{R}^2$ such that

$$\lim_{n \to \pm \infty} \frac{1}{|n|} \log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|$$

equals $-\gamma(E)$ if $\mathbf{x}$ and $\mathbf{x}_{E,v}^{+;-}$ are colinear and $\gamma(E)$ if not. In view of (5.3) this raises the question of the positivity of $\gamma(E)$. This is given by a famous theorem due to Furstenberg as long as the common distribution of the $v(n)$ is not concentrated on a single point. These results have been very important in the understanding of the localization phenomenon in one dimensional disordered systems and are part of the folklore of the subject. We refer the interested reader to the reviews [1, 2, 24] for complete statements, proofs and references. Part B of [24] is a very helpful source of information on the relation between asymptotic properties of products of random matrices and spectral properties of random Schrödinger operators on $\mathbb{Z}$.

Our objective is to prove Theorem 4.1. To prove the desired estimates in finite intervals we will have to go first to the infinite volume limit to get smoothness and then return to finite intervals where the smoothness was not available due to the singularity of the distribution of the potential.

We start with a Lemma whose idea and proof are borrowed from Lemma 3.1 of [15] (see also Proposition BIII2.2 of [24] for a similar estimate). As before $\mu$ denotes the common probability distribution of the $v(n)$'s.

**Lemma 5.1.** *Suppose $\mu$ is not concentrated on a single point and there exists $\eta > 0$ such that*

$$\int |v|^\eta d\mu(v) < \infty. \tag{5.4}$$

*Then, for each compact energy interval $I$ there exist $\alpha_1 = \alpha_1(I) > 0$, $\delta = \delta(I) > 0$ and a positive integer $n_1 = n_1(I)$ such that*

$$E\{ \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} \leqq e^{-\alpha_1 n}$$

*for all $E \in I$, $n \geqq n_1$ and unit vectors $\mathbf{x} \in \mathbb{R}^2$.*

*Proof.* As we already pointed out, $\gamma(E) > 0$ under the above hypotheses. Moreover, it is a continuous function of the energy $E(\gamma(E)$ is actually locally Hölder continuous and we will use this fact later). Thus $\gamma = \gamma_I = \inf\{\gamma(E); E \in I\} > 0$. Using the inequality $e^y \leqq 1 + y + y^2 e^{|y|}$, we obtain

$$\| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \leqq 1 - \delta \log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|$$
$$+ \delta^2 |\log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^2| \exp\{\delta |\log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \||\}.$$

Since $\| T_j^{(E)} \| = \| [T_j^{(E)}]^{-1} \|$, we always have

$$|\log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \| | \leq \log \| T_n^{(E)} \| + \cdots + \log \| T_1^{(E)} \|.$$

We can thus use Hölder's inequality to obtain

$$\begin{aligned}
\mathbf{E}\{ \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} &\leq 1 - \delta \mathbf{E}\{ \log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \| \} \\
&+ \delta^2 n [\mathbf{E}\{ (\log \| T_1^{(E)} \|)^4 \}]^{1/2} [\mathbf{E}\{ \| T_1^{(E)} \|^{2\delta} \}]^{n/2},
\end{aligned}$$

which is finite by (5.4) if we choose $0 < \delta \leq \eta/2$. It follows that

$$E\{ \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} \leq 1 - \delta \mathbf{E}\{ \log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \| \} + \delta^2 n C_1 C_2^n$$

for some finite constants $C_1 = C_1(I)$, $C_2 = C_2(I)$.

Since we know that

$$\lim_{n \to \infty} \frac{1}{n} \mathbf{E}\{ \log \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \| \} = \gamma(E)$$

and that, as remarked in the proof of Lemma BII4.8 of [24], this convergence is uniform in $E \in I$ and $\mathbf{x}$ in the unit sphere of $\mathbb{R}^2$, we can pick $n_0 = n_0(I)$ such that

$$E\{ \| T_{n_0}^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} \leq 1 - \tfrac{1}{2} n_0 \delta \gamma + \delta^2 n_0 C_1 C_2^{n_0} \leq 1 - \varepsilon \tag{5.5}$$

for some $\varepsilon > 0$ provided $\delta$ is chosen small enough. We now use the fact that (5.5) is independent of the unit vector $\mathbf{x}$. If $n$ is any positive integer, then $n = k n_0 + r$ for some integer $k \geq 0$ and $0 \leq r < n_0$. Then

$$E\{ \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} \leq \mathbf{E}\{ \| [T_{kn_0+1}^{(E)}]^{-1} \cdots [T_n^{(E)}]^{-1} \|^\delta \} \mathbf{E}\{ \| T_{kn_0}^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \}$$

$$\leq [\mathbf{E}\{ \| [T_1^{(E)}]^{-1} \|^\delta \}]^{n_0} \mathbf{E}\left\{ \left\| T_{kn_0}^{(E)} \cdots T_{(k-1)n_0+1}^{(E)} \frac{T_{(k-1)n_0}^{(E)} \cdots T_1^{(E)} \mathbf{x}}{\| T_{(k-1)n_0}^{(E)} \cdots T_1^{(E)} \mathbf{x} \|} \right\|^{-\delta} \right.$$

$$\left. \cdot \| T_{(k-1)n_0}^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \right\} \leq C(1-\varepsilon) E\{ \| T_{(k-1)n_0}^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} [E(\| [T_1^{(E)}]^{-1} \|^{\eta/2}]^{n_0}$$

by (5.5), where $C = \sup_{E \in I} [E\{ \| [T_1^{(E)}]^{-1} \|^{\eta/2} \}]^{M_0} < \infty$ by (5.4).

Repeating the above argument we get

$$E\{ \| T_n^{(E)} \cdots T_1^{(E)} \mathbf{x} \|^{-\delta} \} \leq C(1-\varepsilon)^k \leq e^{-\alpha_1 n}$$

for all $n \geq n_1$ and $E \in I$, for some $\alpha_1 = \alpha_1(I) > 0$ and $n_1 = n_1(I)$.

This concludes the proof of Lemma 5.1.

The next lemma is the second ingredient in the proof of Theorem 4.1. Its proof is reminiscent of Simon's version of an old argument of Halperin (see the Appendix of [11]).

**Lemma 5.2.** *Suppose $\mu$ is not concentrated on a single point and (5.4) holds. Then for each compact energy interval $I$ there exists $\rho = \rho(I) > 0$ and $C = C(I) < \infty$ such that for each $E \in I$ and $0 < \varepsilon < 1$ we have*

$$\mathbf{P} \{ \text{there exists } E' \in (E - \varepsilon, E + \varepsilon) \text{ and } \varphi' \in l^2(B_l), \| \varphi' \| = 1,$$
$$\text{such that } (H_{B_l} - E')\varphi' = 0 \quad \text{and} \quad [|\varphi'(-l)|^2 + |\varphi'(l)|^2]^{1/2} \leq \varepsilon \} \leq C l \varepsilon^\rho. \tag{5.6}$$

*Proof.* The integrated density of states $N(E)$ of the operator $H$ on $l^2(\mathbb{Z})$ is locally Hölder continuous. This result was proven by LePage under the extra assumption of boundedness of the potential (see [15; Proposition 3.2]). As explained in [24], the proof relies on the so-called Thouless formula (whose proof usually requires the boundedness of the potential) and the local Hölder continuity of the Lyapunov exponent (whose proof requires only the $\mu$-integrability of $\log(1 + |v|)$). We show in the Appendix to this section that both Thouless formula and the local Hölder continuity of the integrated density of states still hold in the present situation.

Thus there exist $\rho = \rho(I) > 0$, $C_1(I) < \infty$ such that

$$|N(E) - N(E')| \leq C_1 |E - E'|^\rho \tag{5.7}$$

for all $E, E'$ at a distance from $I$ not larger than two.

Now let us fix $E \in I$ and $l$, and for each $i = 0, 1, 2, \ldots$ let $A_i = A_i(E, l, \varepsilon)$ denote the event that there exists at least one eigenvalue, say $E_i$, of the operator $H_{B_l(i(2l+3))}$, such that

$$|E - E_i| \leq \varepsilon, \tag{5.8}$$

and such that at least one of the corresponding eigenfunctions, say $\varphi_i$, with $\|\varphi_i\| = 1$, satisfies

$$[|\varphi_i(i(2l+3) - l)|^2 + \|\varphi_i(i(2l+3) + l)|^2]^{1/2} \leq \varepsilon. \tag{5.9}$$

We denote by $p = p(E, l, \varepsilon)$ the common probability of the events $A_i$. Notice that $A_i$ is the event in the box $B_l(i(2l+3))$ described in (5.6) for the box $B_l$, and hence $p$ equals the left-hand side of (5.6).

Now let $i_1, i_2, \ldots i_k \in \{0, 1, \ldots, n - 1\}$, be distinct and such that $A_{i_j}, j = 1, \ldots, k$, occur. For each such $j$ let $\tilde{\varphi}_j$ be the function on $\mathbb{Z}$ which is equal to $\varphi_{i_j}$ on $B_l(i_j(2l+3))$ and zero elsewhere. Then $\{\tilde{\varphi}_1, \ldots, \tilde{\varphi}_k\}$ is an orthonormal system satisfying

(i)                                 $\|(H_{K(l,n)} - E)\tilde{\varphi}_j\| \leq 2\varepsilon,$

because of (5.8) and (5.9), and

(ii)                    $\langle \tilde{\varphi}_j, H_{K(l,n)}\tilde{\varphi}_{j'} \rangle = \langle H_{K(l,n)}\tilde{\varphi}_j, \ H_{K(l,n)}\tilde{\varphi}_{j'} \rangle = 0$

for $j \neq j'$ by construction. Here

$$K(l, n) = \{-l, -l + 1, \ldots, (n-1)(2l+3) + l\},$$

and $H_{K(l,n)}$ is the restriction of $H$ to $l^2(K(l, n))$.

According to Simon's version of Temple's inequality (see Lemma A3.2 of [11]), this implies the existence of at least $k$ distinct eigenfunctions of $H_{K(l,n)}$ with corresponding eigenvalues in $[E - 2\varepsilon, E + 2\varepsilon]$. Thus $\#\{i \in \{0, 1, \ldots, n-1\};\ A_i$ occurs$\} \leq \#\{$eigenvalues of $H_{K(l,n)}$ in $[E - 2\varepsilon, E + 2\varepsilon]\}$. If we divide each side by $n$ and take the limit as $n \to \infty$, we get $p \leq (2l + 3)[N(E + 2\varepsilon) - N(E - 2\varepsilon)]$, where we used the fact that $N(E)$ is continuous. But then, by (5.7), we have $p \leq (2l + 3)C_1 2^\rho \varepsilon^\rho \leq Cl\varepsilon^\rho$ for some $C = C(I) < \infty$.

This finishes the proof of Lemma 5.2.

We are now ready to prove Theorem 4.1. Let $\beta, \sigma > 0$ and the compact interval $I$ be fixed. For each positive integer $l$ we set $n_l = [\tau l^\beta] + 1$, where $[r]$ denotes the

integer part of $r$ and $\tau > 0$ is to be chosen later. Let $\delta = \delta(I)$ and $\alpha_1 = \alpha_1(I)$ be the constants given in Lemma 5.1, and let $\kappa = \tau\alpha_1/2\delta$. For each $E \in I$ and $\theta > 0$ we set

$$A_\theta^{(E,l)} = \{ \| T_{-l+n_l}^{(E)} \cdots T_{-l} \mathbf{x}_0 \| > e^{\theta l^\beta} \},$$
$$B_\theta^{(E,l)} = \{ \| [T_{l-n_0}^{(E)}]^{-1} \cdots [T_l^{(E)}]^{-1} \mathbf{y}_0 \| > e^{\theta l^\beta} \},$$

where $\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{y}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Using Chebyshev's inequality and Lemma 5.1 we have

$$\mathbf{P}\{[A_\kappa^{(E,l)}]^c\} \leq \exp\{\kappa\delta l^\beta - \alpha_1 n_l\} \leq \exp\{-\tfrac{1}{2}\tau\alpha_1 l^\beta\}, \tag{5.10}$$

and similarly for $[B_\kappa^{(E,l)}]^c$, provided $l$ is large enough.

We can write

$$\mathbf{P}\{\text{dist}(E, \sigma(H_{B_l})) < e^{-\sigma l^\beta}\}$$
$$\leq \mathbf{P}(\{\text{dist}(E, \sigma(H_{B_l})) < e^{-\sigma l^\beta}\} \cap \bigcap_{|E'-E|<e^{-\sigma l^\beta}} [A_{\kappa/2}^{(E',l)} \cap B_{\kappa/2}^{(E',l)}])$$
$$+ \mathbf{P}(A_\kappa^{(E,l)} \cap B_\kappa^{(E,l)} \cap \bigcup_{|E'-E|<e^{-\sigma l^\beta}} [A_{\kappa/2}^{(E',l)}]^c) + \mathbf{P}(A_\kappa^{(E,l)} \cap B_\kappa^{(E,i)} \cap \bigcup_{|E'-E|<e^{-\sigma l^\beta}} [B_{\kappa/2}^{(E',l)}]^c)$$
$$+ \mathbf{P}([A_\kappa^{(E,l)}]^c) + \mathbf{P}([B_\kappa^{(E,l)}]^c) = \text{(i)} + \text{(ii)} + \text{(iii)} + \text{(iv)}.$$

Let us assume that the event whose probability is (i) occurs. Then we must have either

$$|G_{[-l, -l+n_l]}(E'; -l, -l+n_l)| \leq 2e^{-1/2\kappa l^\beta} \tag{5.11}$$

or

$$|G_{[-l, -l+n_l-1]}(E'; -l, -l+n_l-1)| \leq 2e^{-1/2\kappa l^\beta} \tag{5.12}$$

for each $E' \in (E - e^{-\sigma l^\beta}, E + e^{-\sigma l^\beta})$, because for each positive integer $n$ and every energy $E'$ we always have

$$G_{[1,n]}(E'; 1, n) = [\psi_E(n+1)]^{-1},$$

where $\psi_{E'}(n+1)$ is the first component of the vector $T_n^{(E')} \cdots T_1^{(E')} \mathbf{x}$, and we assumed that the event $A_{\kappa/2}^{(E',l)}$ occurs.

Similarly, we must also have that either

$$|G_{[l-n_l,l]}(E'; l, l-n_l)| \leq 2e^{-1/2\kappa l^\beta} \tag{5.13}$$

or

$$|G_{[l-n_l+1,l]}(E'; l, l-n_l+1)| \leq 2e^{-1/2\kappa l^\beta} \tag{5.14}$$

If $|E' - E| < e^{-\sigma l^\beta}$ and $E'$ is an eigenvalue of $H_{B_l}$ with $\varphi'$ a corresponding normalized eigenvector, then, depending on whether (5.11) or (5.12) holds, we use either the formula

$$\varphi'(-l) = G_{[-l, -l+n_l]}(E'; -l, -l+n_l)\varphi'(-l+n_l+1)$$

or

$$\varphi'(-l) = G_{[-l, -l+n_l-1]}(E'; -l, -l+n_l-1)\varphi'(-l+n_l)$$

to estimate $\varphi'(-l)$. In either case we obtain

$$|\varphi'(-l)| \leq 2e^{-1/2\kappa l^\beta}.$$

Similarly, using either (5.13) or (5.14), we get

$$|\varphi'(l)| \leq 2e^{-1/2\kappa l^\beta}.$$

Thus

(i) $\leq \mathbf{P}\{$there exists $E' \in (E - e^{-\sigma l^\beta}, E + e^{-\sigma l^\beta})$ and $\varphi' \in l^2(B_l)$, $\|\varphi'\| = 1$, such that $(H_{B_l} - E')\varphi' = 0$ and

$$[|\varphi'(-l)|^2 + |\varphi'(l)|^2]^{1/2} \leq 2\sqrt{2}e^{1/2\kappa l^\beta}\} \leq Cl(e^{-\sigma l^\beta} \wedge 2\sqrt{2}e^{-1/2\kappa l^\beta})^\rho \qquad (5.15)$$

by Lemma (5.2).

Now let us suppose that the event whose probability is (ii) occurs. Notice that

$$T_i^{(E')} = T_i^{(E)} + (E' - E)J,$$

where

$$J = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}.$$

There exists $E'$ with $|E' - E| < c^{-\sigma l^\beta}$ such that the event

$$[A_{\kappa/2}^{(E',l)}]^c \cap A_\kappa^{(E,l)}$$

occurs, so we have

$$e^{1/2\kappa l^\beta} \geq \| T_{-l+n_l}^{(E')} \cdots T_{-l}^{(E')}\mathbf{x}_0 \|$$

$$\geq \| T_{-l+n_l}^{(E)} \cdots T_{-l}^{(E)}\mathbf{x}_0 \| - \sum_{k=1}^{n_l+1} \sum_{\substack{\{i_1 < i_2 < \cdots < i_{n_l+1-k}\} \\ \subset \{-l,\ldots,-l+n_l\}}} e^{-k\sigma l^\beta} \| T_{i_{n_l+1-k}}^{(E)} \|$$

$$\cdots \| T_i^{(E)} \| \geq e^{\kappa l^\beta} - \Gamma,$$

where $\Gamma$ denotes the double sum above. Thus $\Gamma \geq \frac{1}{2}e^{\kappa l^\beta}$ if $l$ is large enough. Using Chebyshev's inequality and the inequality $(\Sigma |a_i|)^{\delta'} \leq \Sigma |a_i|^{\delta'}$ with $\delta' = 1 \wedge \eta$, we obtain

(ii) $\leq 2^{\delta'}e^{-\delta'\kappa l^\beta} \sum_{k=1}^{n_l+1} \binom{n_l+1}{k} e^{-\delta'k\sigma l^\beta}[\mathbf{E}\{ \| T_1^{(E)} \|^{\delta'}\}]^{n_l+1-k}$

$\leq 2^{\delta'}e^{-\delta'\kappa l^\beta}[(e^{-\delta'\sigma l^\beta} + M)^{n_l+1} - M^{n_l+1}]$

$\leq 2^{\delta'}e^{-\delta'\kappa l^\beta}(n_l+1)e^{-\delta'\sigma l^\beta}(1 + M)^{n_l} = 2^{\delta'}e^{-\delta'\kappa l^\beta}e^{-\delta'\sigma l^\beta}(\tau l^\beta + 2)(1 + M)^{\tau l^\beta + 2},$

where

$$M = M(I) = \sup_{E \in I} \mathbf{E}\{ \| T_1^{(E)} \|^{\delta'}\} < \infty$$

by our assumption on $\mu$ (i.e., by (5.4)). Thus, if we choose $\tau$ small enough, we get

(ii) $\leq e^{-\alpha_2 l^\beta}$ \hfill (5.16)

for some $\alpha_2 = \alpha_2(I, \sigma, \beta) > 0$ and all $l$ large enough.

Similarly, we get

(iii) $\leq e^{-\alpha_2 l^\beta}.$ \hfill (5.17)

Thus the Theorem follows from (5.10), (5.15), (5.16) and (5.17).

This concludes the proof of Theorem 4.1.

We now turn to the question of the regularity properties of the integrated density of states $N(E)$. As it was already discussed, $N(E)$ is locally Hölder continuous. In addition, Kotani and Souillard [25] have shown that $N(E)$ is absolutely continuous if $\mu$ possesses an absolutely continuous component. We will show that for some potential distributions $\mu$ $N(E)$ has a nontrivial singular component. Our result can be proven for $\mu$ concentrated on finitely many points under some extra conditions, but we will restrict ourselves to the Bernoulli case as stated in Theorem 2.2.

So we now turn to the proof of Theorem 2.2. We let $\mu = p\delta(v) + (1-p)\delta(v-b)$ with $p \in (0,1)$ *fixed* and $b > 0$. We will use the same notation as before except for a subscript $b$ to emphasize the dependence on the parameter $b$.

In the present situation we have

$$\sigma(H_b) = [-2,2] \cup [b-2,b+2].$$

We take $b > 4$ so the two intervals are disjoint. Notice that then it is easy to see that

$$N_b([-2,2]) = p,$$

*so it follows that*

$$N_b([b-2,b+2]) = 1 - p.$$

Using Thouless formula (i.e., (A.2)), we have

$$\gamma_b(E) \leqq p \log 4 + (1-p)\log(b+4) \tag{5.18}$$

for any $E \in [-2,2]$. If we now use Fubini's theorem to interchange integrals we get

$$\int_{-2}^{2} \gamma_b(E)dE = \int_{-2}^{2} dN_b(E')\int_{-2}^{2} \log|E-E'|dE + \int_{b-2}^{b+2} dN_b(E')\int_{-2}^{2}\log|E-E'|dE.$$

The first term is bounded uniformly in $b > 0$ and the second one is bounded from above by $4(1-p)\log(b+4)$ and from below by $4(1-p)\log(b-4)$. This implies that

$$\lim_{b\to\infty} (\log b)^{-1}\int_{-2}^{2} \gamma_b(E)dE = 4(1-p) \tag{5.19}$$

Now, for each fixed $\alpha > 0$ we have

$$|\{E \in [-2,2]; \gamma_b(E) \geqq \gamma\}| \geqq [p\log 4 + (1-p)\log(b+4)]^{-1}\int_{[-2,2]\cap\{\gamma_b(E)\geqq\gamma\}}\gamma_b(E)dE$$

because of (5.18), where $|A|$ denotes the Lebesgue measure of the set $A$.

Since

$$\int_{[-2,2]\cap\{\gamma_b(E)<\gamma\}}\gamma_b(E)dE \leqq 4\gamma,$$

we get

$$\liminf_{b\to\infty}|\{E\in[-2,2]; \gamma_b(E)\geqq\gamma\}| \geqq 4$$

by (5.19), and thus

$$\lim_{b\to\infty} |\{E\in[-2,2]; \gamma_b(E)\geqq\gamma\}| = 4. \tag{5.20}$$

We now fix $\gamma > 2\log 2$. According to (5.20), for each $\varepsilon > 0$ we can choose $b = b(\gamma, \varepsilon) > 0$ so large that

$$|\{E \in [-2, 2]; \gamma_b(E) \geqq 8\gamma\}| \geqq 4 - \varepsilon.$$

For such fixed $b$, we use the continuity of the function $\gamma_b(E)$ to construct an open set $I \subset [-2, 2]$ such that $|I| \geqq 4 - \varepsilon$ and $\gamma_b(E) \geqq 4\gamma$ for all $E \in I$.

According to the remark in Sect. 4 about the rate of decay of the eigenfunctions we know that, almost surely, we have that for each $E \in I$, *either E is not an eigenvalue or the corresponding normalized eigenfunction, say* $\varphi_E$, *satisfies*

$$|\varphi_E(x)| \leqq e^{-2\gamma|x|} \tag{5.21}$$

*for all* $|x|$ *large enough.*

Thus, almost surely if $E \in I$ is an eigenvalue, (5.21) holds for $x = \pm (l + 1)$ if $l$ is large enough. As a consequence, if we define the function $\tilde{\varphi}_E$ by $\tilde{\varphi}_E(x) = \varphi_E(x)$ for $-l \leqq x \leqq l$ and $\tilde{\varphi}_E(x) = 0$ otherwise, we have

$$1 = \|\varphi_E\|^2 \leqq \|\tilde{\varphi}_E\|^2 + 2\sum_{x=l+1}^{\infty} e^{-4\gamma x}.$$

Thus

$$\|\tilde{\varphi}\|^2 \geqq 1 - 2(1 - e^{-4\gamma})^{-1}e^{-4\gamma(l+1)}, \quad \|(H_{B_l} - E)\tilde{\varphi}_E\| \leqq e^{-2\gamma(l+1)},$$

and consequently

$$d(E, \sigma(B_l)) < e^{-\gamma(l+1)}[1 - 2(1 - e^{-4\gamma})^{-1}e^{-4\gamma(l+1)}]^{-1/2} < e^{-\gamma l}$$

for $l$ large enough.

We can thus conclude that, with probability one, all the eigenvalues in $I$ are contained in the set

$$I_\gamma = I \cap \bigcap_{k \geqq 1} \bigcup_{l > k} S_b^{(\gamma)}(l), \tag{5.22}$$

where $S_b(l)$ is the set of all possible eigenvalues of $H_{B_l}$ when we take all the possible configurations of the potential in $B_l$, and the superscript $(\gamma)$ means that we are taking an $e^{-\gamma l}$ neighborhood, i.e.,

$$S_b^{(\gamma)}(l) = \{E; \text{dist}(E, S_b(l)) < e^{-\gamma l}\}.$$

But

$$|S_b^{(\gamma)}(l)| \leqq (2l + 1)2^{2l+1}2e^{-\gamma l} = 4(2l + 1)e^{-(\gamma - 2\log 2)l}$$

which is summable in $l$ because of our choice of $\gamma$. Thus (5.22) implies that $|I_\gamma| = 0$.

On the other hand $N_b(I_\gamma) = N_b(I)$ by our construction since all the eigenvalues in $I$ are actually in $I_\gamma$. Since $I \subset \sigma(H_b)$ and $I$ is open we must have that $N_b(I) > 0$. Hence $N_b(I_\gamma) > 0$.

This finishes the proof of Theorem 2.2.

It is worth remarking that the first part of the above proof, at least up to (5.20), is an attempt to show that for each fixed $E$ one has $\lim_{b \to \infty} \gamma_b(E) = \infty$. Indeed, our

argument shows that, for each $\varepsilon > 0$ and $\gamma > 0$,

$$\lim_{b \to \infty} |\{E \in [-2, 2]; \gamma_b(E) \geqq \gamma (\log b)^{1-\varepsilon}\}| = 4.$$

However, using Halperin's argument [11] and Thouless formula, it can be shown that for a dense set of energies $\gamma_b(E) \leqq K \log b$ with $K < (1 - p)$. This shows that the function $\gamma_b(E)(\log b)^{-1}$, as a function of $E$, "tends" to be highly noncontinuous.

## Appendix to Section 5

In this Appendix we extend LePage's result [15] on the local Hölder continuity of the integrated density of states to the case where the potential is no longer assumed to be bounded. The idea of the proof is LePage's original one. We merely improve on some of the steps and here again we benefited greatly from Part B of [24].

**Theorem A.1.** *Suppose that the probability distribution $\mu$ of the potential at one site is not concentrated on a single point and*

$$\int \log(1 + |v|) d\mu(v) < \infty. \tag{A.1}$$

*Then the integrated density of states $N(E)$ and the Lyapunov exponent $\gamma(E)$ are related by Thouless formula:*

$$\gamma(E) = \int \log|E - E'| dN(E'), \tag{A.2}$$

*and the function $E \to N(E)$ is locally Hölder continuous, i.e., for any compact interval $I$ there exist $\rho = \rho(I) > 0$ and $C = C(I) < \infty$ such that*

$$|N(E_1) - N(E_2)| \leqq C|E_1 - E_2|^\rho$$

*for all $E_1, E_2 \in I$.*

*Proof.* Let us write

$$N_l(E) = \frac{1}{2l + 1} \#(-\infty, E) \cap \sigma(H_{B_l})$$

for the (random) empirical distribution function of the eigenvalues of the operator $H_{B_l}$. We recall that, without any assumption on $\mu$, the measures $dN_l(E)$ converge weakly to the deterministic measure $dN(E)$ with probability one (see for example [1, 2] and, as usual in the Birkhoff ergodic theorem, the expectations converge accordingly.

For each configuration of our random potential we obviously have

$$v - 2 \leqq H_{B_l} \leqq v + 2. \tag{A.3}$$

Notice that these upper and lower bounds are diagonal operators. Let us denote by $N_{l,\pm}$ the empirical distribution functions of the diagonal operators $v \pm 2$ restricted to $B_l$. For any function $f$ we have

$$\int f(E) dN_{l,\pm}(E) = \frac{1}{2l + 1} \sum_{n=-l}^{l} f(v(n) \pm 2). \tag{A.4}$$

Moreover, (A.3) implies that

$$N_{l,+}(E) \leq N_l(E) \leq N_{l,-}(E). \tag{A.5}$$

Notice that (A.4) and (A.5) remain valid after taking expectations. If we first use (A.4) and the rightmost inequality of (A.5) with the function $f_m(E) = \log(1 - E)\chi_{\{E \leq -m\}}$, and then use (A.4) and the leftmost inequality of (A.5) with the function $g_m(E) = \log(1 + E)\chi_{\{E \geq m\}}$, we obtain, for each $m > 0$,

$$\int_{|E| \geq m} \log(1 + |E|)d\mathbf{E}\{N_l(E)\} \leq 2\int_{|v| \geq m-2} \log(1 + |v|)d\mu(v) \tag{A.6}$$

because the expectation of the right-hand side of (A.4) is nothing but $\int f(v \pm 2)d\mu(v)$.

The crucial point of (A.6) is the uniform integrability of the function $\log(1 + |E|)$ with respect to the measure $d\mathbf{E}(N_l(E))$ which is implied by (A.1).

Now let $E$ be a complex number with $\operatorname{Im} E \neq 0$. Then

$$\int \log|E - E'|d\mathbf{E}\{N_l(E')\} = \frac{1}{2l+1}\mathbf{E}\{\log|P_l(E)|\}, \tag{A.7}$$

where $P_l$ is the unit (random) polynomial of degree $2l + 1$ whose roots are the eigenvalues of $H_{B_l}$. As $l \to \infty$, the left-hand side of (A.7) converges to $\int \log|E - E'|dN(E')$ because of the uniform integrability given by (A.6) and (A.1). The right-hand side of (A.7) converges to $\gamma(E)$ because $P_l(E)$ is an entry of the matrix $T_{-l+1}^{(E)} \cdots T_l^{(E)}$. This proves Thouless formula for $\operatorname{Im} E \neq 0$ and we conclude as usual by using the fact that both sides of (A.2) are subharmonic functions to obtain (A.2) for all $E$.

We now argue the local Hölder continuity. For $m > 0$ we set

$$\psi_m(E) = N(E)\chi_{[-3m,3m]}(E).$$

$\psi_m$ is an $L^2$-function. Let us denote by $\tilde{\psi}_m$ its Hilbert transform. A simple integration by parts and (A.2) give

$$\tilde{\psi}_m(\lambda) = -\frac{1}{\pi}\psi_m(3m)\log|3m - \lambda| + \frac{1}{\pi}\psi_m(-3m)\log|3m + \lambda|$$

$$+ \gamma(\lambda) - \int_{|E|>3m} \log|E - \lambda|dN(E). \tag{A.8}$$

We will show that $N$ is Hölder continuous in $[-m, m]$ by showing that the Hilbert transform of $\tilde{\psi}_m$ is Hölder continuous in $[-m, m]$. We have

$$\pi\tilde{\tilde{\psi}}_m(E) = \lim_{\delta \downarrow 0}\int_{\delta < |E-\lambda| < m} \frac{\tilde{\psi}_m(\lambda)d\lambda}{E - \lambda} + \int_{|E-\lambda| \geq m} \frac{\tilde{\psi}_m(\lambda)d\lambda}{E - \lambda} = \varphi_1(E) + \varphi_2(E).$$

A simple calculation using only the square integrability of $\tilde{\psi}_m$ shows that

$$|\varphi_2(E_1) - \varphi_2(E_2)| \leq C_1|E_1 - E_2|$$

for all $E_1, E_2 \in \mathbb{R}$ with $C_1 = C_1(m, \|\tilde{\psi}_m\|_2) < \infty$. For $E \in [-m, m]$, $\varphi_1(E)$ involves only values of $\tilde{\psi}_m(\lambda)$ for $\lambda \in [-2m, 2m]$ and in this interval we can read the Hölder continuity of $\tilde{\psi}_m(\lambda)$ from (A.8). Here we used the local Hölder continuity of Lyapunov exponent $\gamma(E)$. This fact was already the crucial step in LePage's original

proof. In the present generality it is argued in Proposition BII4.7 of [24]. Moreover, all terms in the right-hand side of (A.8) are Hölder continuous of order 1, except perhaps for $\gamma(E)$, which is, say, Hölder continuous of order $\alpha > 0$ in $[-2m, 2m]$. In the same way one usually shows that the Hilbert transform preserves local Hölder continuity, one can verify that

$$|\varphi_1(E_1) - \varphi(E_2)| \leq C_2 |E_1 - E_2|^\tau$$

for $E_1, E_2 \in [-m, m]$, where $C_2 = C_2(m) < \infty$ and $\tau = 2(\alpha \wedge 1)$.

This completes the proof.

## 6. The Multidimensional Case

Let $\varepsilon > 0$. It is easy to see that

$$\mathbf{P}\{\text{dist}(E, \sigma(H_{B_l})) \leq \varepsilon\} \leq \mathbf{P}\left\{\text{Im Tr } G_{B_l}(E + i\varepsilon) \geq \frac{1}{2\varepsilon}\right\}$$

$$\leq \sum_{x \in B_l} \mathbf{P}\{\text{Im } G_{B_l}(E + i\varepsilon, x, x) \geq [2\varepsilon(2l + 1)^\nu]^{-1}\}.$$

Thus Theorem 4.2 follows from

**Lemma 6.1.** *Let $\nu$ be arbitrary and suppose $\mu$ is Hölder continuous of order $\rho > 0$, i.e., $\delta_\rho(\mu) > 0$. Let $W$ be a given (deterministic) potential. For $x \in B_l$ and $v \in \mathbb{R}$ define $H_{B_l}^{(x,v)} = -\Delta + W + v\delta_x$ restricted to $l^2(B_l)$, and let $G_{B_l}^{(x,v)}(E + i\varepsilon)$ be the corresponding Green's function. Then for any $0 < \delta < \delta_\rho(\mu)$ we can find $\eta_1 = \eta_1(\mu, \delta) > 0$ such that*

$$\mu\{v; \text{Im } G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) \geq \theta^{-1}\} \leq \delta^{-1}\theta^\rho$$

*for all $0 < \theta < \eta_1$.*

*Proof.* It follows from the resolvent equation that

$$G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) = G_{B_l}^{(x,0)}(E + i\varepsilon; x, x) - G_{B_l}^{(x,v)}(E + i\varepsilon; x, x)vG_{B_l}^{(x,0)}(E + i\varepsilon; x, x)$$

so

$$G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) = [[G_{B_l}^{(x,0)}(E + i\varepsilon; x, x)]^{-1} + v]^{-1}.$$

Let us write $[G_{B_l}^{(x,0)}(E + i\varepsilon; x, x)]^{-1} = c + id$ with $c, d \in \mathbb{R}$. Then

$$\text{Im } G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) = -\frac{d}{(c + v)^2 + d^2},$$

and hence

$$\text{Im } G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) \geq \theta^{-1} \Leftrightarrow d^2 + \theta d + (c + v)^2 \leq 0 \Rightarrow |c + v| \leq \tfrac{1}{2}\theta.$$

Thus

$$\mu\{v; \text{Im } G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) \geq \theta^{-1}\} \leq \mu\{v; |c + v| \leq \tfrac{1}{2}\theta\}.$$

Now use the Hölder continuity of $\mu$. Given $0 < \delta < \delta_\rho(\mu)$, there exists $\eta_1 = \eta_1(\mu, \delta) > 0$ such that $\mu([a, b]) \leq \delta^{-1}|b - a|^\rho$ for all $a \leq b \in \mathbb{R}$ with $|b - a| \leq \eta_1$. Thus,

for any $a \in \mathbb{R}$ and $0 < \theta \leqq \eta_1$ we have

$$\mu\{v; |a + v| \leqq \tfrac{1}{2}\theta\} \leqq \delta^{-1}\theta^\rho.$$

Thus proves the lemma.

The above considerations give a simple proof of Wegner's estimate on the density of states [12]. If we assume that $\mu([a,b]) \leqq \delta^{-1}|b - a|$ for all $a < b \in \mathbb{R}$, we have

$$\int \operatorname{Im} G_{B_l}^{(x,v)}(E + i\varepsilon; x, x) d\mu(v) \leqq \delta^{-1} \int_{-\infty}^{\infty} \frac{-d}{(c + v)^2 + d^2} dv = \delta^{-1}\pi.$$

So

$$\mathbf{E}(\operatorname{Im} G_{B_l}(E + i\varepsilon; x, x)) \leqq \delta^{-1}\pi$$

for all $\varepsilon > 0$ and $x \in B_l$. This implies Wegner's estimate.

A similar proof of Wegner's estimate has been independently obtained by Simon [26] using the ideas of [8].

Lemma 6.1 can be used to prove an extension of Wegner's estimate.

**Theorem 6.2.** *Let $v$ be arbitrary and suppose there exists $\delta > 0$ and $0 < \rho < 1$ such that $\mu([a,b]) \leqq \delta^{-1}|b - a|^\rho$ for all $a < b \in \mathbb{R}$. Then*

$$\mathbf{E}(\operatorname{Im} G_{B_l}(E + i\varepsilon; x, x)) \leqq \delta^{-1}(1 - \rho)^{-1}\varepsilon^{\rho - 1}$$

*for all $\varepsilon > 0$, $l$ and $x \in B_l$. As a consequence we have*

$$\mathbf{P}\{\operatorname{dist}(E, \sigma(H_{B_l})) \leqq \varepsilon\} \leqq 2(1 - \rho)^{-1}\delta^{-1}(2l + 1)^\nu \varepsilon^\rho$$

*for all $E \in \mathbb{R}$, $\varepsilon > 0$ and $x \in B_l$.*

*Proof.* Let us fix $\varepsilon > 0$ and $x \in B_l$. By Lemma 6.2

$$F(\alpha) = \mathbf{P}\{\operatorname{Im} G_{B_l}(E + i\varepsilon; x, x) \geqq \alpha\} \leqq \delta^{-1}\alpha^{-\rho}$$

for all $\alpha > 0$. But since

$$\operatorname{Im} G_{B_l}(E + i\varepsilon; x, x) \leqq \|G_{B_0}(E + i\varepsilon)\| \leqq \varepsilon^{-1},$$

we have that $F(\alpha) = 0$ for $\alpha \geqq \varepsilon^{-1}$. Thus

$$\mathbf{E}(\operatorname{Im} G_{B_l}(E + i\varepsilon; x, x)) = \int_0^\infty F(\alpha)d\alpha \leqq \int_0^{1/\varepsilon} \delta^{-1}\alpha^{-\rho}d\alpha = \delta^{-1}(1 - \rho)^{-1}\varepsilon^{\rho - 1}.$$

This completes the proof since the second statement follows from the first in the standard way.

We now turn to the proof of Theorem 4.3. For each $l$ let us fix $\varepsilon = \varepsilon_l$, $0 < \varepsilon < 1$, such that $\varepsilon^a(2l + 1)^\nu < 1$ for some $0 < a < \tfrac{1}{2}$. Let $0 < \delta < \delta_\rho(\mu_1)$. For each $E \in \mathbb{R}$ and positive integer $l$ we define

$$\Omega_{E,l} = \{\operatorname{dist}(E, \sigma(H_{B_l})) \leqq \varepsilon\}, \quad \Psi_{E,l} = \{\operatorname{Im} \operatorname{Tr} G_{B_l}(E + i\varepsilon) \geqq \tfrac{1}{2}\varepsilon^{-1}\}.$$

Then $\Omega_{E,l} \subset \Psi_{E,l}$, and hence $\mathbf{P}(\Omega_{E,l}) = \mathbf{P}(\Omega_{E,l} \cap \Psi_{E,l})$.

In order to estimate $\mathbf{P}(\Omega_{E,l})$ we will divide the potential configuration $\{v_x\}_{x \in B_l}$ into its "nice" and "bad" parts as follows:

Let $\{\xi_x\}_{x\in B_l}$ be the Bernoulli process on $B_l$ with parameter $p$, i.e., $\{\xi_x\}_{x\in B_l}$ are i.i.d.r, v.'s such that $\xi_x = 0$ with probability $p$ and $\xi_x = 1$ with probability $1 - p$. We will denote by $\tilde{\mathbf{E}}$ the corresponding expectation. For any integrable function $f$ of the potential $v$ in $B_l$ we have

$$\mathbf{E}\{f(v)\} = \tilde{\mathbf{E}}\left\{\int f(v) \prod_{\substack{x\in B_l \\ \xi_x = 0}} d\mu_1(v(x)) \prod_{\substack{x\in B_l \\ \xi_x = 1}} d\mu_2(v(x))\right\}.$$

Thus, with $\mathbf{P}$ denoting the joint probability over $v$ and $\xi$, we have

$$P(\Omega_{E,l} \cap \Psi_{E,l}) \leq \mathbf{P}(\Omega_{E,l} \cap \{\xi_x = 0 \quad \text{and} \quad \text{Im } G_{B_l}(E + i\varepsilon; x, x) \geq \tfrac{1}{2}(2l + 1)^{-\nu}\varepsilon^{-\tau} \quad (6.1)$$
for some $x\in B_l\}) + \mathbf{P}(\Omega_{E,l} \cap \{\text{Im } G_{B_l}(E + i\varepsilon; x, x) < \tfrac{1}{2}(2l + 1)^{-\nu}\varepsilon^{-\tau}$
for all $x\in B_l$ such that $\xi_x = 0\})$.

Here $\tfrac{1}{2} < \tau < 1 - a$.

The first term in (6.1) can be estimated in the same way as we proved Theorem 4.2 using Lemma 6.1 and is thus bounded by $\delta^{-1}2\rho(2l + 1)^{\nu(1 + \rho)}\varepsilon^{\tau\rho}$ if $\varepsilon^{\tau}(2l + 1)^{\nu} < \eta$, where $\eta = \eta(\mu_1, \delta)$ is chosen as in the proof of Theorem 4.2.

The estimate of the second term in (6.1) will be reduced to a sort of percolation problem.

If $v\in\Omega_{E,l}$ we can find $\lambda\in\sigma(H_{B_l})$ such that $|E - \lambda| \leq \varepsilon$; let $\varphi$ be a corresponding normalized eigenfunction. If

$$\text{Im } G_{B_l}(E + i\varepsilon; x, x) \leq \tfrac{1}{2}(2l + 1)^{-\nu}\varepsilon^{-\tau},$$

it follows that

$$|\varphi(x)| \leq (2l + 1)^{-\nu/2}\varepsilon^{1/2(1 - \tau)} \equiv \varepsilon_0 = \varepsilon_0(\varepsilon, l, \tau).$$

Thus

$$\mathbf{P}(\Omega_{E,l} \cap \{\text{Im } G_{B_l}(E + i\varepsilon; x, x) < \tfrac{1}{2}(2l + 1)^{-\nu}\varepsilon^{-\tau} \quad \text{for all} \quad x\in B_l \quad \text{such that} \quad \xi_x = 0\})$$
$$\leq \mathbf{P} \quad \{\text{there exists a normalized eigenfunctions } \varphi \text{ of}$$

$$H_{B_l} \text{ such that } |\varphi(x)| \leq \varepsilon_0 \text{ for all } x\in B_l \text{ such that } \xi_x = 0\}. \quad (6.2)$$

We will estimate the right-hand side of (6.2). Up to now our methods were independent of the dimension $\nu$, but now we have to study separately the cases $\nu = 1$, $\nu = 2$ and $\nu \geq 3$. Since the $\nu = 1$ case is just a much simpler version of the $\nu = 2$ case and was already extensively studied in Sect. 5, we will proceed directly to $\nu = 2$.

So let us fix $\nu = 2$. Let $\varphi$ be a normalized eigenfunction of $H_{B_l}$ such that $|\varphi(x)| \leq \varepsilon_0$ for all $x\in B_l$ for which $\xi_x = 0$. Let $x_0\in B_l$ be such that $|\varphi(x_0)| > \sqrt{\varepsilon_0}$. Such a point always exists since

$$\max_{x\in|B_l|}|\varphi(x)| \geq (2l + 1)^{-\nu/2} > \sqrt{\varepsilon_0} \quad \text{if we have} \quad (2l + 1)^{\nu}\varepsilon^{1 - \tau} < 1.$$

We will actually, later on, estimate from below the number of such $x_0$'s and we will also show that we can pick $x_0$ away from $\partial B_l$ so that the following construction is contained in $B_l$.

Let $\Gamma$ be a finite strip parallel to a coordinate axis of width two and length $2[\kappa\sqrt{|\log\varepsilon|}] + 1$ such that $x_0$ is the middle point of one of the sides of $\Gamma$. The constant $\kappa$ will be chosen later. We divide $\Gamma$ into disjoint blocks $\{M_t^j\}_j$ of length $t$, with $t$ even, and width two. The length $t$ will be chosen later uniformly on $l$. We will say that a given block $M_t^j$ is "nice" if $\xi_x = 0$ for all $x \in M_t^j$, otherwise $M_t^j$ is "bad." Clearly, given a block $M_t^j$ we have

$$\mathbf{P}\{M_t^j \quad \text{is "nice"}\} = p^{2t}. \tag{6.3}$$

In order to reduce our probabilistic estimate to a percolation problem we will need the following deterministic results which are consequences of the discretized Schrödinger equation. We fix $E_0 > 0$ and always take $|E| < E_0$. In the following lemmas $v$ is a bounded potential, bounded by $c_1 = \sup\{|w|; w \in \mathrm{supp}\,\mu\} < \infty$. We set $C = 3 + C_1 + E_0 + 1$. For future use we notice that if $|E| \leq E_0$, $|\lambda - E| \leq \varepsilon$ implies $|\lambda| \leq E_0 + \varepsilon < E_0 + 1$. We will also assume that $\Gamma$ is parallel to the first coordinate axis.

**Lemma 6.3.** *Let $\varphi$ be a solution of $H_{B_l}\varphi = \lambda\varphi$ in $B_l$, $|\lambda| \leq E_0 + 1$, and assume that $|\varphi(x)| \leq \varepsilon_0$ for all $x \in M_t$ for some block $M_t$. Then $|\varphi(x)| \leq \varepsilon_0 C^t$ for all $x \in B_l$ which are in the union of the two triangles whose hypotenuse is one of the lines of $M_t$ of length $t$ and do not intersect the other line.*

*Proof.* Let $T_t^+$ and $T_t^-$ be the upper and lower triangles described above. Let $Q_t = M_t \cup T_t^+ \cup T_t^-$.

Let $0 \leq n < t/2$ and suppose that on each segment in $Q_t$ parallel to $M_t$ at a distance $h \leq n < t/2$ from $M_t$ the function $\varphi$ is uniformly bounded by $C^h\varepsilon_0$. Let $S_{n+1}$ be the segment in $T_t^+$ parallel to $M_t$ at a height equal to $n+1$. Then using the equation $H_{B_l}\varphi = \lambda\varphi$, we have that

$$|\varphi(x)| \leq 3C^n\varepsilon_0 + \sup_y|v(y) - \lambda|\,C^n\varepsilon_0 \leq C^{n+1}\varepsilon_0$$

for all $x \in S_{n+1}$. Since the above argument holds for $n = 0, 1, \ldots, t/2 - 1$, the lemma follows.

We now define $Q_t^n$, $n = 0, 1, \ldots$, inductively as follows:

$$Q_t^0 = Q_t, \quad Q_t^{n+1} = Q_t^n \cup \{x \in Z^2 \backslash Q_t^n; \text{ there exists } y \in Q_t^n \text{ with } |x - y| = 1\}.$$

For $x \in Q_t^n$ let $h(x)$ be the height of $x$ with respect to $\Gamma$. We also let $\varepsilon_0' = \varepsilon_0 C^t$.

**Lemma 6.4.** *There exists $m > 0$ large enough such that if for any $x \in Q_t^n$ we have $|\varphi(x)| \leq \varepsilon_0' e^{mn^2}(h(x) + 1)$, then if*

$$|\varphi(y)| > \varepsilon_0' e^{m(n+1)^2}(h(y) + 1)$$

*for some $y \in Q_t^{n+1} \backslash Q_t^n$ we have that $|\varphi(x)| \geq \varepsilon_0'$ for all $x$ in the diagonal part of the boundary of $Q_t^{n+1}$ to which $y$ belongs.*

*Proof.* Let $\bar{x}$ be one of the vertices of $Q_t^{n+1}$ lying on $\Gamma$, say the leftmost upper one. In order to prove the lemma it suffices to show that:

(i) If $|\varphi(\bar{x})| \leq \varepsilon_0' e^{m(n+1)^2}$ then $|\varphi(x)| \leq \varepsilon_0' e^{m(n+1)^2}(h(x) + 1)$ for all $x \in S_0$, where $S_0$ is the upper left side of the boundary of $Q_t^{n+1}$.

(ii) If $|\varphi(\bar{x})| > \varepsilon_0' e^{m(n+1)^2}$ then $|\varphi(x)| \geq \varepsilon_0'$ for all $x \in S_0$.

In order to prove (i), let $x \in S_0$, $x = (x^{(1)}, x^{(2)})$, and let

$$x_1 = (x^{(1)} - 1, x^{(2)}), x_2 = (x^{(1)} + 1, x^{(2)}),$$
$$x_3 = (x^{(1)} - 1, x^{(2)} - 1), y = (x^{(1)} - 1, x^{(2)} + 1),$$

and suppose $|\varphi(x)| \leq \varepsilon_0' e^{m(n+1)^2}(h(x) + 1)$.

Since $(H_{B_l}\varphi)(x_1) = \lambda \varphi(x_1)$, we get

$$|\varphi(y)| \leq |\varphi(x)| + |\varphi(x_2)| + |\varphi(x_3)| + (C_1 + E_0 + 1)|\varphi(x_1)|$$
$$\leq \varepsilon_0' e^{m(n+1)^2}(h(x) + 1) + (C_1 + E_0 + 3)\varepsilon_0' e^{mn^2}(h(x_1) + 1)$$
$$\leq \varepsilon_0' e^{m(n+1)^2}(h(x) + 2) + \varepsilon_0' C e^{mn^2}(h(x) + 1) - \varepsilon_0' e^{m(n+1)^2}$$
$$\leq \varepsilon_0' e^{m(n+1)^2}(h(y) + 1)$$

for all $n$ if $m$ is chosen sufficiently large.

The proof of (i) now follows by induction on $x$ starting from $\bar{x}$.

To prove (ii), let $x, y, x_1, x_2, x_3$ be as above. We take as induction step

$$|\varphi(x)| > \varepsilon_0' e^{m(n+1)^2} - C\varepsilon_0' h(x)(h(x) + 1)e^{mn^2}.$$

Again, since $(H_{B_l}\varphi)(x_1) = \lambda \varphi(x_1)$ we get

$$|\varphi(y)| > \varepsilon_0' e^{m(n+1)^2} - \varepsilon_0' C e^{mn^2}\{h(x)(h(x) + 1) + (h(x) + 1)\}$$
$$= \varepsilon_0' e^{m(n+1)^2} - \varepsilon_0' C e^{mn^2}(h(x) + 1)^2 \geq \varepsilon_0' e^{m(n+1)^2} - \varepsilon_0' C e^{mn^2} h(y)(h(y) + 1).$$

This proves the induction step. To complete the proof it remains to show that for any $x \in S_0$, we have

$$\varepsilon_0' e^{m(n+1)^2} - \varepsilon_0' C(h(x) + 1)h(x)e^{mn^2} > \varepsilon_0'.$$

This is true if

$$e^{m(2n+1)} \geq C(n + t/2 + 1)(n + t/2)$$

for all $n$. This can be satisfied if we choose $m$ large enough.

This proves the lemma.

We now have all the ingredients to estimate the right-hand side of (6.2) when $v = 2$. The main idea is the following: for all potentials in the set described in the right-hand side of (6.2) the eigenfunction $\varphi$ satisfies the assumptions of Lemma 4.3 for all "nice" blocks $M_t$ in $\Gamma$ and therefore is smaller than $\varepsilon_0 C^t \equiv \varepsilon_0'$ in $Q_t^0$. From Lemma 6.4 it follows that $|\varphi(x)| \leq \varepsilon_0' e^{mn^2}(n + t/2 + 1)$ for all $x \in Q_t^n$ or we can find an $n_0$ such that on one of the diagonal sides of $Q_t^{n_0}$, $|\varphi(x)| \geq \varepsilon_0'$. This in turn implies that this particular side of $Q_t^{n_0}$ contains only "bad" sites. But for each "nice" block only the latter possibility can occur since

$$\varepsilon_0' e^{m(\kappa|\log\varepsilon|^{1/2} - t)^2}(\kappa|\log\varepsilon|^{1/2} - t + t/2 + 1) \leq \varepsilon_0 C^t e^{m\kappa^2|\log\varepsilon|}(\kappa|\log\varepsilon|^{1/2}) \leq \varepsilon_0^{1/2} \leq |\varphi(x_0)|$$

for all $k = k(m) > 0$ sufficiently small, independent of $l$, and $l$ sufficiently large (recall $\varepsilon = \varepsilon_l \to 0$ as $l \to \infty$); how large, of course, depending on our choice of $t$.

We have thus proved the following lemma.

**Lemma 6.5.** *There exist $\kappa > 0$ such that for all "nice" block $M_t^j$ in $\Gamma$, with $\Gamma$ of length $2[\kappa|\log\varepsilon|^{1/2} + 1]$, we can find $n_j \leq [\kappa|\log\varepsilon|^{1/2}]$ such that on one diagonal side of $Q_{t,j}^{n_j}$ the random variables $\xi_x$ are all equal to 1, for all $l \geq l_0$ for some $l_0$. $\kappa$ is independent of $t$, but $l_0$ may depend on $t$.*

For each nice block $M_t^j$ in $\Gamma$ we will denote by $D_j$ the side of $Q_t^{n_j}$ where $\xi_x = 1$ for all $x \in D_j$ (if there are several we just choose one).

**Lemma 6.6.** *Let $N$ be the number of "nice" blocks in $\Gamma$. Then the set $\bigcup\limits_{i=1}^{N} D_i$ contains $K \leq N$ disjoint segments $\bar{D}_j$ at $45°$ with $\Gamma$ such that:*

   (i) *One of the ends of $\bar{D}_j$ coincides with the upper or lower vertex of one $Q_{t,j}^{n_j}$ and the other end lies at a distance 1 from $\Gamma$.*

   (ii) *Let $|\bar{D}_j|$ be the number of lattice sites on $\bar{D}_j$. Then*

$$|\bar{D}_j| \geq t/2 \quad and \quad \sum_{j=1}^{K} |\bar{D}_j| \geq Nt/2.$$

*Proof.* Clearly $\bigcup\limits_{j=1}^{N} Q_{t,j}^{n_j} \supset \bigcup\limits_{j=1}^{N} M_t^j$. Let $\{Q_{t,j_i}^{n_{j_i}}\}_{i=1}^{K}$ be a minimal covering of $\bigcup\limits_{j=1}^{N} M_t^j$ by the $Q_{y,j}^{n_j}$'s (i.e., $K$ is minimal) and let $\bar{D}_i$ be that part of the "bad" side of $Q_{t,j_i}^{n_{j_i}}$ which lies at a distance 1 from $\Gamma$. We claim that the $\bar{D}_i$ are disjoint and satisfy (i) and (ii). Indeed, (i) and (ii) hold by construction. To show that $\bar{D}_i \cap \bar{D}_k = \phi$ if $i \neq k$ we observe that by Lemma 6.4, if the "bad" sides of $Q_{t,j_i}^{n_{j_i}}$ and $Q_{t,j_k}^{n_{j_k}}$ intersect each other, then either they meet at a vertex on $\Gamma$ or they lie on the same line. Therefore by construction the corresponding $\bar{D}_i, \bar{D}_k$ do not cross each other and cannot lie on the same line since this would contradict the minimality of the family $\{Q_{t,j_i}^{n_{j_i}}\}_{i=1}^{K}$.

This proves the lemma.

Let $\Omega_{x_0}$ denote the event described in Lemma 6.2. We now estimate its probability.

**Lemma 6.7.** *Let $0 < p_0 \leq p \leq 1$. Then there exists $t_0 = t_0(p_0)$ such that for $t \geq t_0$ we have*

$$\mathbf{P}(\Omega_{x_0}) \leq e^{-d(p_0, p, t)|\log\varepsilon|^{1/2}},$$

*where $0 < d(p_0, p, t) \to \infty$ as $p \to 1$.*

*Proof.* Clearly

$$\mathbf{P}(\Omega_{x_0}) = \sum_S \mathbf{P}\{\Omega_{x_0} \cap \Omega_S\},$$

where $S$ is an arbitrary subset of the set of blocks $M_t^j$ in $\Gamma$ and $\Omega_s = \{S$ is the set of "nice" blocks in $\Gamma\}$. We denote by $|S|$ the number of blocks in $S$. Thus

$$\mathbf{P}(\Omega_{x_0}) = \sum_{S:|S| \leq t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]} \mathbf{P}(\Omega_{x_0} \cap \Omega_s) + \sum_{S:|S| > t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]} \mathbf{P}(\Omega_{x_0} \cap \Omega_s). \quad (6.4)$$

The first term can be estimated by

$$\mathbf{P}\{\text{number of nice blocks in } \Gamma \leq t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]\}$$
$$\leq e^{-c_2 t^{-1}(1-p^{2t})^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]} \quad (6.5)$$

for large $l$ by standard large deviation estimates on the binomial distribution since the expected number of blocks in $\Gamma$ is equal to $t^{-1}p^{2t}2[\kappa|\log\varepsilon|^{1/2}]$. Here $c_2 > 0$ is some constant independent of our parameters.

To estimate the second term in (6.4) we first observe that once the set $S$ of "nice" blocks in $\Gamma$ is specified then the event $\Omega_{x_0}$ no longer depends on the random variables in $\Gamma\backslash S$. Thus we can estimate the second term in (6.4) by

$$\sum_{S;|S|\geq t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]} P(\Omega_s) \sum_{k=1}^{|S|} \sum_{\substack{t_1,\dots,t_k \\ t_j\geq t/2 \\ \sum_{j=1}^k t_j\geq t/2|S|}} \sum_{\substack{x_1,\dots,x_k \\ x_j \text{ is the middle point} \\ \text{of a block in } S}} 4^k(1-p)^{\sum_{j=1}^k t_j}$$

$$\leq \sum_{S;|S|\geq t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]} P(\Omega_s)(1-p)^{1/4t|S|} \sum_{p=1}^{|S|} \binom{|S|}{k} 4^k \left(\sum_{t'\geq t/2}(1-p)^{t'/2}\right)^k$$

$$\leq \sum_{S;|S|\geq t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]} P(\Omega_s)(1-p)^{1/4t|S|}(1+4[1-(1-p)^{1/2}]^{-1}(1-p)^{1/4t})^{|S|}$$

$$\leq (1-p)^{t^{-1}p^{2t}[\kappa|\log\varepsilon|^{1/2}]t/8} \tag{6.6}$$

if $(1-p_0)^{t/8}(1+4[1-(1-p_0)^{1/2}]^{-1})\leq 1$, which is always true for $t\geq t_0 = t_0(k)$ for some $t_0$.

The Lemma now follows from (6.4), (6.5) and (6.6).

In order to complete the estimate of the right-hand side of (6.2) we observe that if $x_\infty\in B_l$ is such that $|\varphi(x_\infty)| = \max_{x\in B_l}|\varphi(x)|$, then $|\varphi(x_\infty)|\geq (2l+1)^{-\nu/2}$. Thus by the argument in Lemma 6.3 we can choose a direction along a coordinate axis such that for each

$$n < n_0 = [2\log C]^{-1}|\log\varepsilon_0(2l+1)^\nu| = [2\log C]^{-1}|\log\varepsilon^{1/2(1-\tau)}(2l+1)^{\nu/2}|,$$

$C$ being the constant in Lemma 6.3, we can find $x_0$ such that

$$n\geq |x_0^2 - x_\infty^2|\leq n+1 \quad \text{and} \quad |\varphi(x_0)| > \varepsilon_0^{1/2}.$$

We can thus find

$$N \approx (2[\kappa|\log\varepsilon|^{1/2}])^{-1}(2\log C)^{-1}|\log\varepsilon_0(2l+1)^\nu|$$

sites $x_j$ such that $|x_i^2 - x_j^2|\geq 2[\kappa|\log\varepsilon|^{1/2}]$ and $|\varphi(x_j)| > \varepsilon_0^{1/2}$ for $j=1,\dots,N$.

Thus we can estimate (we have to be careful here on how close the $x_j$ are to the edges of $B_l$) the right-hand side of (6.2) by

$$\sum_{x_1,\dots,x_n\in B_l} \prod_{j=1}^N P(\Omega_{x_j})\leq (\sum_{x\in B_l} P(\Omega_x))^N$$

$$\leq (2l+1)^{\nu N}e^{-Nd(p_0,p,t)|\log\varepsilon|^{1/2}}\leq e^{-d'(p_0,p,t)|\log\varepsilon|} = \varepsilon^{d'(p_0,p,t)},$$

where $d'$ has the same properties as $d$, in case $(2l+1)^\nu\leq\varepsilon^{-\theta}$ for some $\theta > 0$.

This proves the theorem for $\nu = 2$.

If $\nu\geq 3$, let again $x_\infty$ be a site where $\varphi$ attains its maximum. Using the argument of Lemma 6.3 it is easy to see that if $|\varphi(x)|\geq\varepsilon_1$, then for any coordinate direction, say the $j$-th, there exists $y$ such that $|y - x|\leq 2$, $y^{(j)} > x^{(j)}$ and $|\varphi(y)|\geq\varepsilon_1 C_\nu^{-1}$, where $C_\nu = (4\nu + C_1 + E_0 + 1)$ as in Lemma 6.3. Since $|\varphi(x_\infty)|\geq (2l+1)^{\nu/2}$, we can find a

path $\gamma \subset B_l$ made out of points at a distance one or two of the adjacent ones, and of length

$$|\gamma| \geqq [\log C_v]^{-1}|\log 2\varepsilon_0 (2l+1)^{v/2}|$$

such that $|\varphi(x)| \geqq 2\varepsilon_0$ for all $x \in \gamma$. For potentials $v$ in the set described by (6.2), this implies that $\xi_x = 1$ for all $x \in \gamma$. Using standard percolation arguments we get that

$$\mathbf{P}\{\text{there exists } \gamma = \{x_1, x_2, \dots\} \in B_l, \quad |x_i - x_{i+1}| \leqq 2,$$

$$|\gamma| > (\log C_v)^{-1}|\log 2\varepsilon_0 (2l+1)^{v/2}|; \text{ such that } \xi_{x_i} = 1 \text{ for all } i\} \leqq e^{-d_1(p)|\gamma|},$$

where $d_1(p) \to \infty$ as $p \to 1$, which is

$$\leqq e^{-d_2 d_1(p)|\log \varepsilon|} = \varepsilon^{d_2 d_1(p)}$$

for some other constant $d > 0$, independent of $p$.

This finishes the proof of Theorem 4.3.

## 7. Some Remarks on the Anderson Model on the Cayley Tree

We would now like to conclude with a discussion of the estimate (II) for the Anderson model on the Cayley tree. This model was first considered by Abou-Chacra, Anderson and Thouless [27]; Kunz and Souillard have announced rigorous results for the case when $\mu$ has a smooth density [28]. Here we want to show that in the case when $\mu$ is a Bernoulli measure the estimate (II) does not hold.

We will take the coordination number of our tree $T$, namely the number of branches out of any vertex, to be three. If $x \in T$, we define $T_k(x)$ as the finite subtree of $T$ obtained by exactly $k$ bifurcations in two directions starting from $x$. The root of $T_k(x)$, say $\tau(x)$, will be the nearest neighbour of $x$ not in $T_k(x)$. The endpoints of $T_k(x)$ different from $x$ will be denoted by $x_i^{(k)}$, $i = 1, 2, \dots, 2^k$.

Let $T_i = T_k(x_i), i = 1, 2$, be disjoint and such $\tau(x_1) = \tau(x_2)$, and let $v_{T_i} = \{v(x)\}_{x \in T_i}$, $i = 1, 2$. We will say that $v_{T_1} = v_{T_2}$ if by superimposing $T_2$ with $T_1$ keeping $\tau(x_1) = \tau(x_2)$ the values of $v$ in $T_2$ and $T_1$ coincide.

**Lemma 7.1.** *Let $T_i = T_k(x_i), i = 1, 2$, be as above, and let $v_{T_1} = v_{T_2}, \tau = \tau(x_1) = \tau(x_2)$. Then any eigenvalue $E$ for $H_{T_1} = -\Delta + v_{T_1}$ on $l^2(T_1)$ is also an eigenvalue for $-\Delta + v$ on $l^2(T_1 \cup T_2 \cup T_0 \cup \tau)$, if $v|T_i = v_{T_i}, i = 1, 2$ where $T_0$ is any subtree such that*

(i)                              $T_0 \cap T_1 = T_0 \cap T_2 = \phi$

(ii)                            the root of $T_0$ is $\tau$.

*Proof.* Let $J: T_1 \to T_2$ be the map that associates to $x \in T_1$ the site $J(x) \in T_2$ which coincides with $x$ when we superimpose $T_1$ on $T_2$ keeping $\tau$ fixed. Let $E \in \sigma(H_{T_1})$ and let $\varphi$ be a corresponding normalized eigenfunction. Define $\varphi \in l^2(T_1 \cup T_2 \cup T_0 \cup \tau)$ by

$$\psi(x) = \varphi(x) \quad \text{if} \quad x \in T_1,$$

$$\psi(x) = -\varphi(J^{-1}(x)) \quad \text{if} \quad x \in T_2,$$

$$\psi(x) = 0 \text{ elsewhere.}$$

It is easy to check that $(-\Delta + v)\psi = E\psi$ in $T_1 \cup T_2 \cup T_0 \cup \tau$, and hence $E$ is an eigenvalue of $-\Delta + v$ on $l^2(T_1 \cup T_2 \cup T_0 \cup \tau)$ if $v|_{T_i} = v_{T_i}, i = 1, 2$.

This proves the lemma.

Now let us fix

$$\mu = p\delta(v) + (1 - p)\delta(v - b),$$

where $0 < p < 1, b \in \mathbb{R}$, and let us fix sites $x_1, x_2$ as above and a positive integer $k$. We let $T_1 = T_k(x_1)$ be fixed and let $\bar{v}_{T_1}$ be a possible potential configuration in $T_1$, namely, $\bar{v}_{T_1}(x) \in \{0, b\}$ for all $x \in T_1$. Clearly

$$\mathbf{P}\{v_{T_1} = \bar{v}_{T_1}\} = p^{|\{x; \bar{v}_{T_1}(x) = 0\}|}(1 - p)^{|\{x; \bar{v}_{T_1}(x) = b\}|}. \tag{7.1}$$

By $\tilde{T}_n$ we will denote a tree of the form $\tilde{T}_n = T_n(x_1) \cup T_n(x_2)$, where $T_n(x_1)$ and $T_n(x_2)$ are disjoint and $\tau(x_1) = x_2, \tau(x_2) = x_1$.

Our result is

**Theorem 7.2.** *Let $E$ be an eigenvalue of $-\Delta + \bar{v}_{T_1}$ on $l^2(T_1)$. Then for any $n$ large enough we have*

$$\mathbf{P}\{E \in \sigma(H_{\tilde{T}_n})\} \geq [p^{|\{x; \bar{v}_{T_1}(x) = 0\}|}(1 - p)^{|\{x; \bar{v}_{T_1}(x) = b\}|}]^2.$$

*In particular,* (II) *does not hold for the Cayley tree.*

*Proof.* Using Lemma 7.1 we get

$$\mathbf{P}\{E \in \sigma(H_{\tilde{T}_n})\} \geq P\{v_{T_k}(x_i^{n-k}(x_j)) = \bar{v}_{T_1} \quad \text{and}$$
$$v_{T_k}(x_{i=1}^{n-k}(x_j)) = \bar{v}_{T_1} \quad \text{for some} \quad j = 1, 2 \quad \text{and} \quad i \in \{1, \ldots, n-1\}\},$$

where the $T_k(x_i^{n-k}(x_j))$ are chosen in such a way that their endpoints are a subset of the endpoints of $T_n(x_j), j = 1, 2$.

From (7.2) and (7.1) we get

$$\mathbf{P}\{E \in \sigma(H_{\tilde{T}_n})\} \geq [p^{|\{x; \bar{v}_{T_1}(x) = 0\}|}(1 - p)^{|\{x; \bar{v}_{T_1}(x) = b\}|}]^2.$$

This proves the theorem.

Lemma 7.1 can also be used to discuss the quantum percolation on the Cayley tree. This model has as Hamiltonian the discrete Laplacian on the random subset of $T$ determined by the set of open bounds in the standard bond percolation problem on the tree with

$$\mathbf{P}\{\text{a given bond is open}\} = p$$

and

$$\mathbf{P}\{\text{a given bond is closed}\} = 1 - p.$$

Using Lemma 7.1 we can show that there exists a nonrandom countable dense subset of $\sigma(H)$ which is with probability one contained in the pure point spectrum of $H$. However, this does not exclude the possibility of continuous spectrum.

## References

1. Carmona, R.: Random Schrödinger operators. Proceedings of the Ecole d'ete de probabilite XIV, Saint-Flour, 1984. Lect. Notes in Math. Vol. **1180**. Berlin, Heidelberg, New York: Springer (1986)

2. Spencer, T.: The Schrödinger equation with a random potential. A mathematical review. Proceeding of the Les Houches Summer School 1984 (to appear)
3. Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tightbinding model for large disorder of low energy. Commun. Math. Phys. **88**, 151–184 (1983)
4. Jona-Lasinio, G., Martinelli, F., Scoppola, E.: A quantum particle in a hierarchical potential with tunneling over arbitrarily large scales. J. Phys. **A17**, L635–L638 (1984)
5. Martinelli, F., Scoppola, E.: Remark on the absence of absolutely continuous spectrum in the Anderson model for large disorder or low energy. Commun. Math. Phys. **97**, 465–471 (1985)
6. Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: Constructive proof of localization in the Anderson tight binding model. Commun. Math. Phys. **101**, 21–46 (1985)
7. Delyon, F., Levy, Y., Souillard, B.: Anderson localization for multidimensional systems at large disorder or low energy. Commun. Math. Phys. **100**, 463–470 (1985)
8. Simon, B., Wolff, T.: Singular continuum spectrum under rank one perturbations and localization for random Hamiltonians. Commun. Pure Appl. Math. **39**, 75–90 (1986)
9. Delyon, F., Levy, T., Souillard, B.: Anderson localization for one- and quasi one-dimensional systems. J. Stat. Phys. **41**, 375 (1985)
10. Kotani, S.: Lyapunov exponents and spectra for one-dimensional random Schrödinger operators. Proceedings A.M.S. conference on Random Matrices and their Applications 1984 (to appear)
11. Simon, B., Taylor, M: Harmonic analysis on $SL(2,\mathbb{R})$ and smoothness of the density of states in the one-dimensional Anderson model. Commun. Math. Phys. **101**, 1–19 (1985)
12. Wegner, F.: Bounds on the density of states in disordered systems. Z. Phys. **B44**, 9–15 (1981)
13. Klein, A., Martinelli, F., Perez, J. F.: A rigorous replica trick approach to Anderson localization in one dimension. Commun. Math. Phys. **106**, 623–633 (1986)
14. Campanino, M., Klein, A.: A supersymmetric transfer matrix and differentiability of the density of states in the one-dimensional Anderson model. Commun. Math. Phys. **104**, 227–241 (1986)
15. LePage, E.: Repartition d'etats pour les matrices de Jacobi a coefficients aleatoires, Lecture Notes in Mathematics. Vol. **1064**, 309–367 (1983)
16. Pastur, L. A.: Spectral properties of disordered systems in the one body approximation. Commun. Math. Phys. **75**, 179–186 (1980)
17. Kunz, H., Souillard, B.: Sur le spectre des operateurs aux differences finies aleatoires. Commun. Math. Phys. **78**, 201–246 (1980)
18. Delyon, F., Kunz, H., Souillard, B.: One dimensional wave equation in random media. J. Phys. **A16**, 25–42 (1983)
19. Carmona, R.: One-dimensional Schrödinger operator with random potentials. Physica **124**, 181–188 (1984)
20. Kotani, S.: Private communication
21. Zygmund, A.: Trigonometric Series, vol. II. Cambridge: Cambridge University Press 1959
22. Brown,G., Hewitt, E.: Adv. Math **37**, 27–60 (1980)
23. Holden, H., Martinelli, F.: Absence of diffusion near the bottom of the spectrum for a Schrödinger operator on $L^2(\mathbb{R}^\nu)$. Commun. Math. Phys. **93**, 197–217 (1984)
24. Bougerol, P., Lacroix, J.: Products of random matrices with applications to Schrödinger operators. Boston; Birkhäuser 1985
25. Kotani, S., Souillard, B.: Private communication
26. Simon, B.: Private communication
27. Abou-Chacra, R., Anderson, P. W., Thouless, P. S.: A selfconsistent theory of localization. J. Phys. **C6** 1734 (1973)
28. Kunz, H., Souillard, B.: The localization transition on the Bethe Lattice. J. Physique-Lettres **44**, L411–L414 (1983)

**Note added in proof.** Since the completion of this work, F. Martinelli and L. Micheli have proven that, under the hypothesis of Theorem 2.2, $\lim_{b\to\infty} \inf\{\gamma_b(E); E\in[-2,2]\} = \infty$. If this fact is used in the proof of Theorem 2.2, it follows that the integrated density of states is actually purely singular continuous for $b$ large enough.