

1.2. The result. Let μ_σ be the integrated density of states measure (IDS) for H_ω . We have the following theorem:

Theorem 1. *Consider the Anderson model under the conditions that \mathbb{P} has mean 0, variance 1 and support bounded by c_0 . For all $\gamma > 0$ the IDS, μ_σ , restricted to the interval $(-2 + \gamma, -\gamma) \cup (\gamma, 2 - \gamma)$, is Hölder continuous with exponent $1 - 496c_0^3\sigma/\gamma$. More precisely, for $\lambda_0 \in (-2 + \gamma, -\gamma) \cup (\gamma, 2 - \gamma)$, $\sigma \leq 1$ and $\lambda \leq 1$*

$$\mu_\sigma[\lambda_0, \lambda_0 + \lambda] \leq \frac{2}{\sigma^3} \lambda^{1-496c_0^3\sigma/\gamma}.$$

1.3. Why the Anderson model. The Anderson model is used to consider a quantum mechanical particle moving through a disordered solid, feeling potential from atoms at the lattice sites, where the randomness of the potential corresponds to impurities in the solid; see, for example, the discussion in [8]. The particle moving in d -dimensional space is given by a function ψ , and its evolution by $e^{-itH_\omega}\psi_0$. With this view, the operator prescribes the time evolution of the particle, and properties of the spectrum of H_ω , $\Sigma(H_\omega)$, correspond to questions about how electrons move through the wire. A natural question to ask is whether the generalized eigenfunctions are localized or delocalized, which can be thought of as a question about the conductive properties of the solid. When $\sigma = 0$ we imagine a metal with no impurities, which we expect to be a conductor. Indeed, the operator H_0 has spectrum $(-2, 2)$, and its generalized eigenfunctions are not in ℓ^2 . On the other hand, in 1-dimension, for any $\sigma > 0$ one can show that the eigenfunctions become exponentially localized, a phenomenon known as Anderson localization. See for example the results of [6, 10], and [4], the latter covering the case of Bernoulli-potentials.

1.4. The integrated density of states. The integrated density of states (IDS) can be thought of as the average number of eigenfunctions per unit volume in the spectrum. It can be obtained by restricting the operator to a finite box, and then taking the limit of the empirical eigenvalue distribution, see [8]. Understanding the IDS is a first step in the study of the spectral properties of the random operator. When \mathbb{P} is absolutely continuous, much is understood about the IDS. The main tool mathematicians use in this case is the celebrated estimate of Wegner [14]. It bounds the expected number of eigenvalues in a small interval of the spectrum of a Schrödinger operator restricted to a finite box. This bound depends on the infinity norm of the density, and so only exists in the case where the distribution of the noise is absolutely continuous. The lack of this tool in cases where the noise is not absolutely continuous results in a bigger challenge to prove many expected results; even in the simple case where the noise has a Bernoulli distribution, referred to as the Anderson–Bernoulli model, much less is known.

It is natural to ask further questions about the IDS, such as what kind of continuity properties it has, and whether we can describe it more explicitly. One would expect that the IDS should be Hölder continuous for small coupling constants, and that the exponent should improve, specifically approach 1 as $\sigma \downarrow 0$, see [1]. This and more has been known when the noise is absolutely continuous for some time. For example, Minami estimates – bounds on the probability of seeing two eigenvalues in a small interval of the spectrum of a Schrödinger operator – are even more refined than the Wegner estimate, can be proved in the continuous case, and are used in [11] to establish Poisson statistics of the spectrum. On the other hand, when the noise is not absolutely continuous, it is possible for Hölder continuity to fail if σ is not small enough. For example, in [13] Simon and

Taylor formalize a result of Halperin [7] to show that, when the noise is Bernoulli, for any $\sigma > 0$, the IDS cannot be Hölder continuous with exponent greater than

$$2 \log 2 / \operatorname{arccosh}(1 + \sigma).$$

Since the maximum exponent of Hölder continuity is 1 anyway, this result has no content for small sigma. On the other hand, for any $\sigma > 9/8 = \cosh(2 \log 2) - 1$, the exponent of Hölder continuity must be bounded away from 1.

1.5. Hölder continuity. In [12] Hölder continuity is established in the Anderson–Bernoulli model for certain coupling constants, but the exponent in that paper gets worse instead of better as σ decreases. Bourgain [1] establishes that the Hölder continuity doesn’t break down as σ decreases, and the exponent must tend to at least 1/5. This result is improved in [2], where he gives a non-quantitative bound to show that the Hölder exponent converges to 1 as $\sigma \downarrow 0$. Following his argument carefully it seems that his methods yield a bound of the form

$$1 - c |\log(\sigma)|^{-1/2}.$$

In contrast, our result gives that the speed with which the exponent tends to 1 is bounded by

$$1 - c\sigma,$$

where our value of c is explicit. In both our result and Bourgain’s the constant depends on the energy being considered, in particular it gets large at energies near the edge of the spectrum, but also near 0. However, our method applies to a wider class of noise distributions than Bernoulli, specifically our main assumption is that \mathbb{P} has finite support. Our assumptions that \mathbb{P} has mean 0 and variance 1 are for ease of notation.

The breakdown of this work is as follows. In Sect. 2, we use the method of Transfer matrices to view the eigenvalue equation for the finite-level Schrödinger operator as a product of 2×2 matrices, and get some geometric intuition by viewing this matrix product as a random walk in the (upper half) complex plane via projectivization. In Sect. 3, we prove a deterministic result (Theorem 2) relating the number of eigenvalues in a small interval of the finite-level Schrödinger operator to the number of large backtracks of the imaginary part of a random walk (with drift) defined in Sect. 2. We also bound the jumps of the real part of this random walk. In Sect. 4, we use the known Figotin–Pastur recursion, most clearly laid out in [3], and a Martingale argument to bound the probability of large backtracks of random walks like the one in Sect. 2 (Theorem 3). Finally, in Sect. 5, we carefully choose some parameters and apply the results of Sects. 2 and 3 to bound the probability of the number of eigenvalues in a small interval of the finite-level Schrödinger operator, and take a limit to obtain the main result.

2. Preliminaries

2.1. The transfer matrix approach. Consider the 1-dimensional random Schrödinger operator in the Anderson model $H_\omega = H_0 + \sigma V_\omega$. We will be working with the restriction of this operator to a finite box, $H_{\omega,n}$. Since H_ω is tri-diagonal, the eigenvalue equation

$$H_{\omega,n}\phi = \lambda\phi$$

can be solved recursively in order to determine if a given λ is an eigenvalue. Doing so allows us to write down an equivalent formulation of the eigenvalue equation:

$$\begin{bmatrix} \phi_{n+1} \\ \phi_n \end{bmatrix} = T_n^{(\lambda)} T_{n-1}^{(\lambda)} \cdots T_1^{(\lambda)} \begin{bmatrix} \phi_1 \\ \phi_0 \end{bmatrix} \tag{2}$$

where we set $\phi_{n+1} = \phi_0 = 0$ and the T matrices are given by

$$T_i^{(\lambda)} = \begin{bmatrix} \lambda - \sigma \omega_i & -1 \\ 1 & 0 \end{bmatrix}.$$

Note that ϕ_n in this equation is unknown, and that by linearity we may let $\phi_1 = 1$, which is allowed because ϕ_1 can't be 0, since if it were, the recursion would imply that $\phi \equiv 0$. This rewriting of the eigenvalue equation is a common technique when studying the spectrum of Schrödinger operators in the Anderson model, often called the transfer matrix approach. One immediate benefit of this approach is that we can use the transfer matrices to define the Lyapunov exponent, $\gamma_\sigma(\lambda)$, a quantity which captures the speed at which the product of these transfer matrices grows, as follows

$$\gamma_\sigma(\lambda) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|T_i^{(\lambda)}\|.$$

The Lyapunov exponents of Schrödinger operators can give us information about the operators themselves. For example, the authors in [5] give a theorem excluding Hölder continuity of the IDS for operators with large Lyapunov exponents.

2.2. The complex plane. To help with intuition, we will identify the objects we're working with in the upper half of the complex plane (UHP). Specifically, we can view the transfer matrices $T_i^{(\lambda)}$ as automorphisms of the UHP through projectivization. Given some (complex) 2-vector

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

we think of its projectivization as the point

$$\mathcal{P}[v] = \frac{v_1}{v_2}$$

in the complex plane. Then a 2×2 matrix

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

can be thought of as an automorphism of the plane as

$$M \circ v = \mathcal{P} \left[M \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right] = \frac{a\mathcal{P}[v] + b}{c\mathcal{P}[v] + d}.$$

While the UHP will be the most useful model for us to think about our objects geometrically, occasionally things will be easier to understand in the context of the disk. For example, a certain automorphism of the half plane may be most easily understood as a "rotation" if it corresponds to mapping the UHP to the disk with a Cayley transform, applying a rotation to the disk, and then mapping the result back to the UHP. In such cases, we may call such an automorphism a rotation for simplicity.

2.3. *More on transfer matrices.* We will be investigating the spectrum by fixing a particular point, or energy in the spectrum, λ_0 , and looking at the spectrum near this energy. For a fixed λ_0 , define θ , ρ , and z by

$$\begin{aligned} \lambda_0 &:= 2 \cos \theta, \quad 0 \leq \theta \leq \pi \\ \rho &:= \frac{1}{\sqrt{4 - \lambda_0^2}} = \frac{1}{2 \sin \theta} \end{aligned}$$

and

$$z := (\lambda_0 + i/\rho) / 2 = e^{i\theta}.$$

To simplify notation we suppress the λ_0 when it appears in the transfer matrices, writing

$$T_i^{(\lambda_0)} = T_i = \begin{bmatrix} \lambda_0 - \sigma \omega_i & -1 \\ 1 & 0 \end{bmatrix}.$$

Finding eigenvalues near λ_0 means solving equation (2) for $\lambda_0 + \lambda$. If we define

$$Q = \begin{bmatrix} 1 & 0 \\ -\lambda & 1 \end{bmatrix}$$

then $T_i^{(\lambda_0+\lambda)} = T_i Q$, and we can substitute this into equation (2), evaluated at $\lambda_0 + \lambda$, to get

$$\begin{bmatrix} \phi_{n+1} \\ \phi_n \end{bmatrix} = T_n Q T_{n-1} Q \cdots T_1 Q \begin{bmatrix} \phi_1 \\ \phi_0 \end{bmatrix}$$

which we can rearrange, and apply our boundary conditions $\phi_0 = \phi_{n+1} = 0$ to, to obtain

$$(T_1)^{-1} (T_2)^{-1} \cdots (T_n)^{-1} \begin{bmatrix} 0 \\ \phi_n \end{bmatrix} = Q^{T_{n-1} T_{n-2} \cdots T_1} Q^{T_{n-2} T_{n-3} \cdots T_1} \cdots Q \begin{bmatrix} \phi_1 \\ 0 \end{bmatrix} \quad (3)$$

with the notation Q^A being conjugation of Q by A (i.e. $A^{-1} Q A$). This expression is convenient because all of the randomness on the right hand side is in the conjugation, but λ only appears in Q , which has no randomness. This allows us to easily view the process as a random walk. To simplify notation, let $W_i = T_i T_{i-1} \cdots T_1$, call the expression on the left hand side of (3) v_* , i.e.

$$v_* = W_n^{-1} \begin{bmatrix} 0 \\ \phi_n \end{bmatrix}$$

and let V_n be the expression on the right hand side of equation (3) so that (by reversing the sides of the equation) we may rewrite (3) as

$$V_n := \begin{bmatrix} v_{1,n} \\ v_{2,n} \end{bmatrix} = Q^{W_{n-1}} Q^{W_{n-2}} \cdots Q^{W_1} Q \begin{bmatrix} \phi_1 \\ 0 \end{bmatrix} = v_*. \quad (4)$$

The sequence $\{W_k^{-1} \circ z\}_{k=1}^n$ defines a process in the UHP, and the sequence $\{\mathcal{P}[V_k]\}_{k=1}^n$ defines a process on the boundary of the UHP plane. Each V_k is obtained by applying the automorphism $Q^{W_{k-1}}$ to the previous point, starting at the point at infinity, given by the projectivization of

$$p = \begin{bmatrix} \phi_1 \\ 0 \end{bmatrix}.$$

Let s_k be the projectivization of V_k , in other words

$$s_k = \mathcal{P}[V_k] = v_{1,k}/v_{2,k}$$

and, keeping in mind that the process

$$W_n^{-1} \begin{bmatrix} z \\ 1 \end{bmatrix}$$

corresponds to the process $W_n^{-1} \circ z$ in the UHP model, we will split this process up into its real and imaginary parts so that

$$X_n + iY_n := W_n^{-1} \circ z.$$

With the understanding of the process $W_n^{-1} \circ z$ as a process in the UHP, and its separation into real and imaginary parts, we are able to state our main theorems.

2.4. Main theorems. If Y is a real valued process, then whenever Y increases by B , we call this a *backtrack* of Y by an amount B . Note that this terminology makes more sense for processes with drift down. In particular it makes sense for the imaginary parts of random walks in the UHP which converge to the boundary.

Theorem 2. *Let $\lambda_0 \in (-2, 0) \cup (0, 2)$, $n \in \mathbb{N}$, $\lambda > 0$ and $\epsilon > 0$. Fix M , let $0 < \beta \leq (2M)^{-1}$, and assume that $|\Delta X_k|/Y_k = |X_k - X_{k-1}|/Y_k \leq M$ for all $k \leq n$. Then the number of eigenvalues of $H_{\omega,n}$ in the interval $[\lambda_0, \lambda_0 + \lambda]$ can be no more than 1 plus the number of backtracks of the process $\log Y_n + [(\epsilon + \lambda\beta) / \sin \theta + 2M\beta]n$ that are at least as large as $\log(\epsilon\beta/\lambda)$.*

Theorem 3. *Assume $\sin 2\theta \neq 0$. Let $E(\omega_j) = 0$, $E(\omega_j^2) = 1$, $|\omega_j| < c_0$, and $\sigma \leq \frac{2 \sin \theta |\sin 2\theta|}{496c_0^3}$. Also assume $\kappa \leq 6c_0^3 \rho^3 \sigma^3 / |\sin 2\theta|$. Then the probability that the process $\log Y_n + \kappa n$ has a backtrack of size B starting from time 1 is at most*

$$2e^{-B(1-248c_0^3\sigma/2\sin\theta|\sin 2\theta|)}.$$

3. Random Schrödinger Operator and Random Walks

3.1. Walk on the boundary of the UHP. The process V_k can be viewed as a random walk on the boundary of the UHP via projectivization. Since

$$Q \circ v = \frac{v}{1 - \lambda v}$$

there is reason to think of the matrix Q as moving points v on the boundary of the UHP “to the right”. Since λ is small, it certainly does this when v is not too large. If v is very large, it is possible that $Q \circ v < v$, but in this case we will think of Q as having moved v “to the right, past ∞ ”. In this sense, conjugates of Q also move points “to the right” along the boundary of the UHP.

With this in mind, we view the process V_n as a random walk on the boundary of the UHP moving only to the right, so the notion of “how many times this process passes a fixed point” makes sense. On the other hand, since (4) is just a rearrangement

of the eigenvalue equation for the Schrödinger operator $H_{\omega,n}$, we make the following observation: for a fixed n and λ if

$$Q^{W_{n-1}} Q^{W_{n-2}} \dots Q \begin{bmatrix} \phi_1 \\ 0 \end{bmatrix} = v_*$$

then $\lambda_0 + \lambda$ is an eigenvalue of $H_{\omega,n}$. This motivates the following well known fact:

Lemma 4. *The number of eigenvalues of $H_{\omega,n}$ in the interval $[\lambda_0, \lambda_0 + \lambda]$ is equal to the number of times that the process $Q_{\lambda}^{W_{k-1}} Q_{\lambda}^{W_{k-2}} \dots Q_{\lambda}(p)$ passes the point v_* as k goes from 1 to n .*

Note: the idea here is that for a fixed n we plan to count the eigenvalues of $H_{\omega,n}$ by considering each Q^{W_k} as one step in a process, and looking at the behaviour of that process as k goes from 1 to n .

Proof. This proof from [9]. Let $B = [\lambda_0, \lambda_0 + \lambda] \times [0, n]$. By interpolating linearly to continuous time, we may consider the continuous map $f : B \rightarrow S^1$ given by

$$f(\lambda, t) = Q_{\lambda(t-1-\lfloor t-1 \rfloor)}^{W_{\lfloor t-1 \rfloor}} Q_{\lambda}^{W_{\lfloor t-1 \rfloor}} Q_{\lambda}^{W_{\lfloor t-2 \rfloor}} \dots Q_{\lambda}(p).$$

Consider the loop given by going around the perimeter of B , i.e. from $(\lambda_0, 0)$ to $(\lambda_0 + \lambda, 0)$ to $(\lambda_0 + \lambda, n)$ to (λ_0, n) and back to $(\lambda_0, 0)$. Since B is simply connected, the image of f is topologically trivial. Further, $f([\lambda_0, \lambda_0 + \lambda] \times \{0\}) = f(\{\lambda_0\} \times [0, n]) = p$. Therefore, $f(\{\lambda\} \times [0, n])$ and $f([\lambda_0, \lambda_0 + \lambda] \times \{n\})$ must have opposite winding numbers. In other words, the number of times that the process

$$\{V_k\}_{k=1}^n$$

passes the point v_* is equal to the number of times that the process

$$Q_{\lambda^*}^{W_{n-1}} Q_{\lambda^*}^{W_{n-2}} \dots Q_{\lambda^*}(p)$$

passes the point v_* as λ^* is varied from 0 to λ . By the observation above, the latter is clearly the number of eigenvalues in $[\lambda_0, \lambda_0 + \lambda]$. \square

3.2. *Bounding By rotations.* Define

$$V'_t = R^{W_t} V_t \tag{5}$$

where R is given by

$$R = \frac{\lambda}{\sin^2 \theta} \begin{bmatrix} -\cos \theta & 1 \\ -1 & \cos \theta \end{bmatrix},$$

and W_t is the piecewise constant interpolation of W_n , that is $W_t = W_{\lfloor t \rfloor}$. Note that R is chosen so that if we map the UHP to the disk using the version of the Cayley transform sending $z = e^{i\theta}$ to the center of the disk, then R is a rotation about z with speed λ . For this reason we may think of R as a “rotation” even in the UHP. In Theorem 5 we find a relationship between V_k and V_t , and in what follows we will use this relationship to understand V_k through V_t . This is useful because rotations are relatively simple to deal with. This view of R as a “rotation” is also useful in explaining our view of what happens in the projectivization of the V_t process as the point moves past infinity.

Theorem 5. *The process V_k is upper-bounded by the process V_t given by differential equation (5), in the sense that the projectivizations of V_k and V_t are each processes following the point at infinity as it moves along the boundary of the UHP to the right, and for any time $t = k$, the point in the V_t process has moved at least as much as the point in the V_k process has.*

Consider first a simple version of the V_k process where the Q matrices are unconjugated. Call this process \tilde{V}_k , so

$$\tilde{V}_k = Q^k \begin{bmatrix} \phi_1 \\ 0 \end{bmatrix}.$$

Then the \tilde{V}_k process can be described by the finite difference equation

$$\tilde{V}_{k+1} = Q \tilde{V}_k \tag{6}$$

where we set

$$\tilde{V}_0 = \begin{bmatrix} v_{1,0} \\ v_{2,0} \end{bmatrix} = \begin{bmatrix} \phi_1 \\ 0 \end{bmatrix}.$$

Lemma 6. *Solutions to the finite difference equation (6) are equal to solutions to differential equation (7) at integer times.*

$$\tilde{V}'_t = \begin{bmatrix} 0 & 0 \\ -\lambda & 0 \end{bmatrix} \tilde{V}_t =: \Lambda \tilde{V}_t. \tag{7}$$

Proof. The difference equation (6) can be decoupled by considering the rows separately. The first row gives $\tilde{v}_{1,k+1} = \tilde{v}_{1,k}$. This means that $\Delta \tilde{v}_1 = 0$ (where we have dropped the k from this coordinate because the solution tells us that it's autonomous). The second row gives $\tilde{v}_{2,k+1} = -\lambda \tilde{v}_{1,k} + \tilde{v}_{2,k}$. This means that $\Delta \tilde{v}_2 = -\lambda \tilde{v}_1$, (where again we drop the k because our solution from the first row means that this row is also autonomous). On the other hand, the differential equation (7) is already decoupled, and encodes precisely the same information: $\tilde{v}'_1 = 0$, $\tilde{v}'_2 = -\lambda \tilde{v}_1$. \square

We now consider the differential equation (7) instead of the difference equation (6). We would like to work with the projectivization, specifically the process $\tilde{s} = \tilde{v}_{t,1}/\tilde{v}_{t,2}$. Using the quotient rule, we obtain the differential equation governing \tilde{s} , which is:

$$\tilde{s}' = \lambda \tilde{s}^2. \tag{8}$$

Note that \tilde{s} gives (through its solutions at integer times) the projectivization of the \tilde{V}_k process. Ultimately we would like to bound the V_k process by the process given in (5). To that end, we will consider what happens when we replace the matrix Λ in (7) by R . If we replace Λ by R in Eq. (7), then with our understanding of R as a rotation, we can use monotonicity to relate the solutions of the two differential equations.

Lemma 7. *The solution to differential equation (8) is upper bounded¹ by the solution to the differential equation (9), below, which comes from the projectivization of the differential equation obtained by replacing Λ with R in the \tilde{V}_t process:*

$$\tilde{s}' = \frac{\lambda}{\sin^2 \theta} \left(\tilde{s}^2 - 2\tilde{s} \cos \theta + 1 \right). \tag{9}$$

¹ To be more precise, we think of the one-point compactification of the real line, which is topologically a circle. The solutions of the differential equations in the theorem can be lifted to the universal cover, which is the real line, and this is where the upper bound holds.

Proof. The derivative \tilde{s}' is strictly positive in both differential equations, which means in both cases, the solution \tilde{s} is strictly increasing, so it suffices to show that \tilde{s}' is always bigger in (9) than in (8), or that the ratio

$$\frac{\frac{\lambda}{\sin^2 \theta} (\tilde{s}^2 - 2\tilde{s} \cos \theta + 1)}{\lambda \tilde{s}^2}$$

is always at least 1. But we can use calculus to find that this ratio is minimized by $\tilde{s} = 1/\cos \theta$, and has a minimum value of precisely 1. \square

At this point we have shown that the simple version of the V_k process (\tilde{V}_k , where the Q matrices are unconjugated) has its projectivization upper bounded by the solution to the differential equation given above in (9). We will now show that this holds even in the case where the Q matrices are conjugated.

Let s be the projectivization of the process defined by

$$\tilde{V}'_t = \Lambda^{W_t} \tilde{V}_t.$$

In other words, by using s we are now reintroducing the conjugations.

Corollary 8. *The solution to the differential equation governing s is upper bounded by the solution to the differential equation governing the process corresponding to s but with Λ replaced by the rotation matrix R . In other words, the result of Lemma 7 holds true even in the case where the Q matrices are conjugated.*

Proof. Conjugation of Q by a k -independent matrix W is equivalent to replacing the \tilde{V} in the finite difference equation (6) by WV . This new finite difference equation encodes the same information as differential equation (7) applied to WV

$$WV'_t = \begin{bmatrix} 0 & 0 \\ -\lambda & 0 \end{bmatrix} WV_t.$$

In the projectivization, this means that conjugation of the Q matrices corresponds to applying the transformation W to \tilde{s} in differential equations (8) and (9). Since W is a fractional linear transformation, it respects order, so the results of Lemma 7 still apply. Since W_t is a piecewise constant function, by continuity of the solutions, the bound holds even when conjugating by W_t . \square

We may now prove Theorem 5:

Proof. Equation (8) with W_k applied to \tilde{s} is the equation governing the projectivization of the process V_k , and Eq. (9) with W_t applied to \tilde{s} is the equation governing the projectivization of the process V_t . By Corollary 8 the projectivization of V_t bounds the projectivization of V_k . \square

Theorem 5 allows us to consider V_t instead of V_k with the effect that the point on the boundary that we are following will always have moved to the right more than it would have without the replacement. This is useful since R , and therefore R^W are rotations, so R^W has a fixed point, $W^{-1} \circ z$. To figure out where the point p gets moved by the process V_t , we need only follow the sequence of centers of rotations: $W_k^{-1} \circ i$.

3.3. *Movement from a different perspective.* We will now look at the process $s_t = \mathcal{P}[V_t]$ from the perspective of the process $W_t \circ z$. From this perspective, s_t will have discrete jumps at integer times. Write $W_t^{-1} \circ z = X_t + iY_t$ where X_t and Y_t are real and coupled in the following way: $dY_t = Y_t dZ$ and $dX_t = Y_t dU$ for some processes U and Z . Note that U and Z are pure jump processes.

Lemma 9. V_t satisfies the differential equation

$$V'_t = \frac{\lambda}{\sin^2 \theta} \begin{bmatrix} -\cos \theta & 1 \\ -1 & \cos \theta \end{bmatrix}^{\bar{W}_t} V_t = \frac{\lambda}{\sin \theta} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^{A\bar{W}_t} V_t$$

where

$$A = \begin{bmatrix} 1 & -\cos \theta \\ 0 & \sin \theta \end{bmatrix}$$

and

$$\bar{W}_t = \begin{bmatrix} 1 & -X_t + Y_t \cot \theta \\ 0 & Y_t / \sin \theta \end{bmatrix}.$$

Proof. The first equality is nearly a restatement of the definition of V_t from equation (5), but with \bar{W}_t in place of W_t , so to prove the first equality it is sufficient to check that $R^{W_t} = R^{\bar{W}_t}$. The eigenvectors of R are

$$\begin{bmatrix} z \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} \bar{z} \\ 1 \end{bmatrix}.$$

But $W_t^{-1} \circ z = X_t + iY_t$, and we can compute $\bar{W}_t \circ X_t + iY_t = z$, so

$$\bar{W}_t^{-1} \begin{bmatrix} z \\ 1 \end{bmatrix} = cW_t^{-1} \begin{bmatrix} z \\ 1 \end{bmatrix}$$

which means that the eigenvectors of $W_t \bar{W}_t^{-1}$ are also

$$\begin{bmatrix} z \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} \bar{z} \\ 1 \end{bmatrix}$$

so R and $W_t \bar{W}_t^{-1}$ commute. Therefore $R^{W_t \bar{W}_t^{-1}} = R$, and $R^{W_t} = R^{\bar{W}_t}$. The second equality is true because

$$\frac{1}{\sin \theta} \begin{bmatrix} -\cos \theta & 1 \\ -1 & \cos \theta \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^A.$$

□

Now let F_t be V_t seen from the perspective of the $X_t + iY_t$, so we have

$$F_t := A\bar{W}_t V_t = \begin{bmatrix} v_{1,t} - X_t v_{2,t} \\ Y_t v_{2,t} \end{bmatrix}$$

and we can compute dF_t as follows:

$$\begin{aligned} dF_t &= Y_t \begin{bmatrix} -dU \\ dZ \end{bmatrix} v_{2,t} + \begin{bmatrix} v'_{1,t} - X_t v'_{2,t} \\ Y_t v'_{2,t} \end{bmatrix} \\ &= Y_t \begin{bmatrix} -dU \\ dZ \end{bmatrix} v_{2,t} + \begin{bmatrix} 1 - X_t \\ 0 \end{bmatrix} Y_t V'_t dt \\ &= Y_t \begin{bmatrix} -dU \\ dZ \end{bmatrix} v_{2,t} + A \bar{W} V'_t dt \\ &= Y_t \begin{bmatrix} -dU \\ dZ \end{bmatrix} v_{2,t} + \frac{\lambda}{\sin \theta} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} F_t dt. \end{aligned}$$

Once again the differential equation is autonomous, so can be written compactly as:

$$dF = F_2 \begin{bmatrix} -dU \\ dZ \end{bmatrix} + \frac{\lambda}{\sin \theta} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} F dt \tag{10}$$

and taking projectivizations, we define

$$\bar{s}_t := \frac{F_1}{F_2}.$$

Remark 10. The process \bar{s} starts at p and moves along the boundary of the UHP, however it is not well defined because of the discrete jumps at integer times. To ensure that \bar{s} is well defined, we will always use the right-continuous version of the process.

Lemma 11. Fix λ, M, ϵ , and $\beta \leq (2M)^{-1}$. Let X_t and Y_t be real processes coupled by $dY_t = Y_t dZ$ and $dX_t = Y_t dU$, where U and Z are pure jump processes. If $|\Delta X_t|/Y_t \leq M$ (for all t), and the process $\log Y_n + [(\epsilon + \lambda\beta) / \sin \theta + 2M\beta]n$ has no backtracks as large as $\log(\epsilon\beta/\lambda)$, then the process \bar{s} can never pass ∞ .

Proof. First define

$$L := \log(-\bar{s}) = \log(-F_1) - \log F_2$$

This doesn't make sense for $\bar{s} \geq 0$, but for the remainder of the proof we will only be concerned with negative values of \bar{s} , so this causes no problems. We can use (10) to find the differential equation governing L . This differential equation will have three terms, the first two of which come from jumps:

- $dF_1/dU = -F_2$ and $dF_2/dU = 0$. When $F_1 \rightarrow F_1 - F_2 dU$, $\log(-F_1) \rightarrow \log(-(F_1 - F_2 dU))$, so $dL = \log(-(F_1 - F_2 dU)) - \log(-F_1) = \log(1 - dU/\bar{s})$. So dL has a $\log(1 - dU/\bar{s})$ term.
- $dF_2/dZ = F_2$ and $dF_1/dZ = 0$. When $F_2 \rightarrow F_2 + F_2 dZ$, $\log(F_2) \rightarrow \log(F_2 + F_2 dZ)$, so dL has a $-\log(1 + dZ)$ term.
- At non-integer values of t , L is continuous in t , so we may use the quotient rule to compute that dL has a $\frac{\lambda}{\sin \theta} (\bar{s} + 1/\bar{s}) dt$ term.

So the differential equation governing L is

$$dL = \frac{\lambda}{\sin \theta} (\bar{s} + 1/\bar{s}) dt - \log(1 + dZ) + \log(1 - dU/\bar{s})$$

and if we integrate both sides between t_1^- and t_2 we get

$$L_{t_1^-} - L_{t_2} = \int_{t_1}^{t_2} \frac{\lambda}{\sin \theta} (e^L + e^{-L}) dt + \int_{t_1^-}^{t_2} \log(1 + dZ) - \int_{t_1^-}^{t_2} \log\left(1 - \frac{dU}{\bar{s}}\right). \tag{11}$$

Here, the second and third integral correspond to summing the integrands over the jumps of Z and U . Also, note that both sides absorbed a negative sign. Now let $t_2 = \inf\{t : \bar{s} \geq -1/\beta\}$, and let $t_1 = \sup_{t < t_2} \{t : \bar{s} \leq -\epsilon/\lambda\}$. Then we have the following inequalities:

$$\begin{aligned} L_{t_1^-} &\geq \log \epsilon/\lambda \\ L_{t_2} &\leq \log 1/\beta \end{aligned}$$

so that

$$L_{t_1^-} - L_{t_2} \geq \log \epsilon/\lambda - \log 1/\beta. \tag{12}$$

When $t_1 \leq t \leq t_2$ we have:

$$\epsilon/\lambda \geq e^L \geq 1/\beta \tag{13}$$

and

$$\lambda/\epsilon \leq e^{-L} \leq \beta. \tag{14}$$

Since Y_t is piecewise constant $dZ = 0$ at non-integer times, so $Y_{t+1} - Y_t = Y_t dZ$ by the definition of Z , meaning $dZ + 1 = Y_{t+1}/Y_t$ at integer times. Hence

$$\int_{t_1^-}^{t_2} \log(1 + dZ) = \log(Y_{t_2}/Y_{t_1^-}) = \log Y_{t_2} - \log Y_{t_1^-}. \tag{15}$$

Since ΔU is upper bounded by M , $-\bar{s}$ is lower bounded by $1/\beta$ on the interval we are considering, and $\beta \leq (2M)^{-1}$, we have $|dU/\bar{s}| \leq M\beta \leq 1/2$. For $x \leq 1/2$ we can use $-\log(1 - x) < 2x$ to get

$$-\int_{t_1^-}^{t_2} \log(1 - dU/\bar{s}) \leq ([t_2] - [t_1^-]) 2M\beta. \tag{16}$$

We are now able to continue integrating in Eq. (11). Combining (12)–(16), (11) implies that

$$\log \epsilon\beta/\lambda \leq (t_2 - t_1) \frac{\lambda}{\sin \theta} (\epsilon/\lambda + \beta) + \log Y_{t_2} - \log Y_{t_1^-} + ([t_2] - [t_1^-]) 2M\beta$$

and by rearranging, we have:

$$\log Y_{t_2} - \log Y_{t_1^-} + (t_2 - t_1) [(\epsilon + \lambda\beta) / \sin \theta] + ([t_2] - [t_1^-]) 2M\beta \geq \log \epsilon\beta/\lambda.$$

For this inequality to hold, the process $\log Y_n + [(\epsilon + \lambda\beta) / \sin \theta + 2M\beta]n$ must have a backtrack of size at least $\log \epsilon\beta/\lambda$ between $[t_1^-]$ and $[t_2]$. So such backtracks are necessary in order for \bar{s} to move through the range between $-\epsilon/\lambda$ to $-1/\beta$, which is necessary for \bar{s} to pass ∞ . In particular, we get the condition that in order for \bar{s} to pass ∞ , the process $\log Y_n + [(\epsilon + \lambda\beta) / \sin \theta + 2M\beta]n$ must backtrack by at least $\log \epsilon\beta/\lambda$. □

3.4. Proof of Theorem 2.

Proof. Define N_n to be the number of eigenvalues of $H_{\omega,n}$ in the interval $[\lambda_0, \lambda_0 + \lambda]$. By Lemma 4, N_n is equal to the number of times the process $\{\mathcal{P}[V_k]\}_{k=1}^n$ passes the point $\mathcal{P}[v_*]$, and so from Theorem 5 we get that N_n is less than or equal to the number of times the process s_t passes the point $\mathcal{P}[v_*]$, which is no more than 1 plus the number of times the process s_t passes ∞ .

By Lemma 11, in order for the process \bar{s} , and therefore the process s_t to pass ∞ , there must be a backtrack as large as $\log \epsilon\beta/\lambda$ in the process $\log Y_n + [(\epsilon + \lambda\beta) / \sin \theta + 2M\beta]n$. □

Theorem 2 gives a deterministic result relating the number of eigenvalues of a finite level Schrödinger operator to the number of large backtracks of the imaginary part of a random walk. It also relies on the existence of a bound on the jumps of the real part of that random walk. We now prove that such a bound exists.

3.5. Bounding the real part.

Theorem 12. *Let X_n and Y_n be defined as in Sect. 2.3, with $\sigma \in [0, 1]$, θ arbitrary, $|\omega_i| \leq c_0$ and $c_0 \geq 1$. Then for all $k \geq 0$*

$$\frac{|X_{k+1} - X_k|}{Y_k} \leq \frac{\sqrt{5} \sigma c_0^2}{2 \sin^2 \theta}.$$

Proof. Define

$$d_1(x + iy, x' + iy') = \frac{|x - x'|}{y} \tag{17}$$

and also

$$d_2(x + iy, x' + iy') = \frac{(x - x')^2 + (y - y')^2}{yy'}.$$

Lemma 13. d_2 is invariant under Möbius transforms, namely

$$d_2(z, z') = d_2(Tz, Tz') \tag{18}$$

for any T fixing the UHP.

Proof. It suffices to check the following 3 cases:

d_2 is invariant under shifts:

$$d_2(z + d, z' + d) = \frac{((x + d) - (x' + d))^2 + (y - y')^2}{yy'} = d_2(z, z')$$

d_2 is invariant under dilations:

$$d_2(\alpha z, \alpha z') = \frac{\alpha^2(x - x')^2 + \alpha^2(y - y')^2}{\alpha y \alpha y'} = d_2(z, z')$$

d_2 is invariant under inversion:

$$\begin{aligned}
 d_2(1/z, 1/z') &= d_2\left(\frac{x-iy}{|z|^2}, \frac{x'-iy'}{|z'|^2}\right) \\
 &= \frac{(x/|z|^2 - x'/|z'|^2)^2 + (-y/|z|^2 + y'/|z'|^2)^2}{yy'/|z|^2|z'|^2} \\
 &= \frac{|z|^2|z'|^2}{yy'} \left[\frac{x^2}{|z|^4} - \frac{2xx'}{|z|^2|z'|^2} + \frac{(x')^2}{|z'|^4} + \frac{y^2}{|z|^4} - \frac{2yy'}{|z|^2|z'|^2} + \frac{(y')^2}{|z'|^4} \right] \\
 &= \frac{1}{yy'} \left[(x^2 + y^2) \frac{|z'|^2}{|z|^2} + ((x')^2 + (y')^2) \frac{|z|^2}{|z'|^2} - 2(xx' + yy') \right] \\
 &= \frac{1}{yy'} \left[|z'|^2 + |z|^2 - 2(xx' + yy') \right] \\
 &= \frac{(x-x')^2 + (y-y')^2}{yy'} = d_2(z, z').
 \end{aligned}$$

□

Lemma 14.

$$d_1^2 \leq d_2\left(1 + \frac{d_2}{4}\right)$$

Proof. Write $z = x + iy$, $z' = x' + iy'$. Since both d_1 and d_2 are invariant under shifts and dilations of the UHP, we may assume that $x = 0$ and $y = 1$. Then

$$d_1(z, z') = |x'|$$

and

$$d_2(z, z') = \frac{(x')^2 + (1 - y')^2}{y'}.$$

Now we can simplify:

$$d_2(z, z') \left(1 + \frac{d_2(z, z')}{4}\right) - (x')^2 = \frac{((x')^2 + 1 - (y')^2)^2}{(2y')^2} \geq 0$$

so that

$$d_2\left(1 + \frac{d_2}{4}\right) \geq (x')^2 = d_1^2$$

completing the proof. □

Now we have the following:

$$\begin{aligned}
 \frac{|X_k - X_{k+1}|}{Y_k} &= d_1\left(W_k^{-1} \circ z, W_{k+1}^{-1} \circ z\right) \\
 &\leq \sqrt{d_2\left(W_k^{-1} \circ z, W_{k+1}^{-1} \circ z\right) \left(1 + d_2\left(W_k^{-1} \circ z, W_{k+1}^{-1} \circ z\right) / 4\right)}. \quad (19)
 \end{aligned}$$

But we can bound $d_2(W_k^{-1} \circ z, W_{k+1}^{-1} \circ z)$ as follows:

$$\begin{aligned} d_2(W_k^{-1} \circ z, W_{k+1}^{-1} \circ z) &= d_2(W_k^{-1} \circ z, W_k^{-1} T_{k+1}^{-1} \circ z) \\ &= d_2(z, T_{k+1}^{-1} \circ z) \\ &= d_2(T_{k+1} \circ z, z). \end{aligned}$$

When $\omega = 0$ we have that $T_{k+1}^{\omega=0} \circ z = z$, so

$$\begin{aligned} d_2(T_{k+1} \circ z, z) &= d_2(T_{k+1} \circ z, T_{k+1}^{\omega=0} \circ z) \\ &= d_2\left(\frac{(\lambda_0 - \sigma\omega_{k+1})z - 1}{z}, \frac{\lambda_0 z - 1}{z}\right) \\ &= d_2(\lambda_0 - \sigma\omega_{k+1} - \bar{z}, \lambda_0 - \bar{z}). \end{aligned}$$

By invariance under Möbius transforms, this is equal to

$$d_2(-\sigma\omega_{k+1} + i \sin \theta, i \sin \theta)$$

which can be computed to get

$$d_2(W_k^{-1} \circ z, W_{k+1}^{-1} \circ z) = \frac{(\sigma\omega_{k+1})^2}{\sin^2 \theta}.$$

Using this bound in (19) gives

$$\frac{|X_{k+1} - X_k|}{Y_k} \leq \sqrt{\frac{(\sigma c_0)^2}{\sin^2 \theta} \left(1 + \frac{(\sigma c_0)^2}{4 \sin^2 \theta}\right)}$$

and since we have $\sin \theta \leq 1$, $c_0 \geq 1$, and $\sigma \leq 1$, we get

$$\frac{|X_{k+1} - X_k|}{Y_k} \leq \frac{\sqrt{5}\sigma c_0^2}{\sin^2 \theta}.$$

□

4. Bounding Backtracks

4.1. The Figotin–Pastur vector.

Lemma 15. *Let \tilde{M} be a real 2×2 matrix with determinant 1. Then*

$$\text{Im}(\tilde{M}^{-1} \circ i) = \left\| \tilde{M} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^{-2}.$$

Proof. Write

$$\tilde{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

so that we have

$$\begin{aligned} \operatorname{Im} \left(\tilde{M}^{-1} \circ i \right) &= \operatorname{Im} \left(\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \circ i \right) \\ &= \operatorname{Im} \frac{id - b}{-ic + a} \\ &= \frac{\operatorname{Im} ((id - b)(a + ic))}{a^2 + c^2} \\ &= \frac{1}{a^2 + c^2} \\ &= \left\| \tilde{M} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^{-2}. \end{aligned}$$

□

We want to understand the backtracks of the $\log Y_t$ process, which means we want to follow the log of $\operatorname{Im} \left((AW_t)^{-1} \circ i \right)$. Lemma (15) allows us to instead follow $1/||\gamma_t||^2$, where

$$\gamma_t := AW_t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

which is the well-known Figotin–Pastur vector for which a recurrence relation is known. Define r_k and α_k by

$$\gamma_k = \sqrt{r_k} \begin{pmatrix} \cos(\alpha_k) \\ \sin(\alpha_k) \end{pmatrix}$$

so that

$$Y_k^{-1} = r_k / \sin \theta = ||\gamma_k||^2 / \sin \theta$$

(because $\det AW_t = \sin \theta$) and recall that

$$z = e^{i\theta}$$

and

$$\rho = \frac{1}{2 \sin \theta} = \frac{1}{|1 - z^2|}.$$

Then from [3] we have the recurrence relations

$$r_{k+1} = r_k (1 + 2\sigma^2 \omega_{k+1}^2 \rho^2 - 2\sigma \omega_{k+1} \rho \sin (2\alpha_k + 2\theta) - 2\sigma^2 \omega_{k+1}^2 \rho^2 \cos (2\alpha_k + 2\theta)) \quad (20)$$

and

$$e^{2i\alpha_{k+1}} = e^{2i\alpha_k} z^2 - \frac{\sigma \omega_{k+1} i \rho (z^2 e^{2i\alpha_k} - 1)^2}{1 - \sigma \omega_{k+1} i \rho (1 - z^2 e^{2i\alpha_k})} \quad (21)$$

and the non-recursive expression for r_k

$$r_k = \prod_{j=1}^{k-1} \left(1 + 2\sigma^2 \omega_j^2 \rho^2 + 2\sigma \omega_j \rho \sin (2\alpha_{j-1} + 2\theta) - 2\sigma^2 \omega_j^2 \rho^2 \cos (2\alpha_{j-1} + 2\theta) \right).$$

4.2. *Martingales.* In what follows, we will use a martingale argument to bound the probability of a large backtrack of the process $\log Y_n + \kappa n$, with Y_n as in the previous section and κ sufficiently small. We will use a function of Y_n which, raised to the power of $1 - \delta$, is a supermartingale for an appropriate choice of δ . This δ will need to be big enough to make the process a supermartingale, but it can't be too large or else it will ruin the bound we are trying to get. We find lower and upper bounds for δ ; the lower bound is the more important bound, necessary to ensure we are working with a supermartingale, where as the upper bound we choose is for technical reasons, specifically to bound a Taylor expansion cutoff, and could be chosen differently if desired.

Lemma 16. *Assume there are positive constants $c_1 \dots c_7$ so that the following holds. Let X_k be a sequence of random variables such that*

$$E(X_k | \mathcal{F}_{k-1}) = \sigma^2 B_{k-1}, \quad E(X_k^2 | \mathcal{F}_{k-1}) \leq \sigma^2 A_{k-1},$$

where $|A_k| \leq 6\rho^2$, $|B_k| \leq 4\rho^2$, and $|X_k| \leq c_1\sigma$, and where \mathcal{F}_k is the sigma algebra generated by $\omega_1, \dots, \omega_k$. Assume further that there exists a constant \tilde{c} and some functions F_k, G_k such that with $\Delta F_k = F_k - F_{k-1}$ we have

$$|B_k - \Delta F_k - \tilde{c}| \leq c_3\sigma, \quad |A_k - \Delta G_k - \tilde{c}| \leq c_5\sigma, \tag{22}$$

and

$$|\Delta F_k| \leq c_2, \quad |\Delta G_k| \leq c_4. \tag{23}$$

Then for $\kappa \in [0, 1]$, σ satisfying

$$\sigma \leq \max(4c_1, c_6, (c_2 + c_4)^{1/2})^{-1} \tag{24}$$

and for δ satisfying

$$\frac{2\sigma}{\tilde{c}} \left(\frac{2\kappa}{\sigma^3} + c_3 + c_5 + 2c_7 \right) \leq \delta \leq \frac{1}{2} \tag{25}$$

with c_7 as in (31), the following process is a supermartingale:

$$\Pi_k = e^{\sigma^2(1-\delta)(F_{k-1} - (1-\delta/2)G_{k-1})} \prod_{i=1}^k (e^{-\kappa} (1 + X_i))^{\delta-1}.$$

Proof.

$$E(\Pi_k | \mathcal{F}_{k-1}) = \Pi_{k-1} E(e^{\sigma^2(1-\delta)(\Delta F_{k-1} - (1-\delta/2)\Delta G_{k-1})} (e^{-\kappa} (1 + X_k))^{\delta-1} | \mathcal{F}_{k-1})$$

We will write

$$1 + a := E((1 + X_k)^{\delta-1} | \mathcal{F}_{k-1}), \quad 1 + b := e^{\sigma^2(1-\delta)(\Delta F_{k-1} - (1-\delta/2)\Delta G_{k-1})}$$

and it suffices to show that

$$(1 + a)(1 + b) \leq e^{-\kappa}.$$

First we get two bounds on a :

For $\delta \in [0, 1/2]$ and $|x| \leq 1/4$, Taylor expansion gives $|(1 + x)^{\delta-1} - 1| \leq 2|x|$, giving the bound

$$|a| \leq 2c_1\sigma. \tag{26}$$

Taking the Taylor expansion one term further gives

$$(1+x)^{\delta-1} \leq 1 - (1-\delta)(x - (1-\delta/2)x^2) + 3|x|^3.$$

Since $|X_k| \leq c_1\sigma \leq 1/4$ we get the more precise bound on a :

$$a \leq -\sigma^2(1-\delta)(B_{k-1} - (1-\delta/2)A_k) + 3c_1^3\sigma^3. \tag{27}$$

Now we get a bound on b :

For $|x| \leq 1$ we have the two inequalities $|e^x - 1| \leq 2|x|$ and $e^x \leq 1 + x + x^2$. Note that by (23) we have $|\Delta F| + |\Delta G| \leq c_2 + c_4$. The first inequality gives that for $\sigma^2 \leq 1/(c_2 + c_4)$ we have the bound on b :

$$|b| \leq 2(c_2 + c_4)\sigma^2. \tag{28}$$

The second inequality gives more precisely:

$$b \leq \sigma^2(1-\delta)(\Delta F_{k-1} - (1-\delta/2)\Delta G_{k-1}) + \sigma^4(c_2 + c_4)^2. \tag{29}$$

If $\sigma < 1/c_6$, the last term is at most $\sigma^3(c_2 + c_4)^2/c_6$. To bound the product $(1+a)(1+b)$ we use the finer bounds for $a+b$ and the rough bounds for $|ab|$. Combining (26,27,28,29) this way, we get an upper bound of

$$1 + \sigma^2(1-\delta)(\Delta F_{k-1} - B_{k-1} + (1-\delta/2)(A_{k-1} - \Delta G_{k-1})) + \text{error} \tag{30}$$

where

$$\text{error} \leq (3c_1^3 + (c_2 + c_4)^2/c_6 + 4c_1(c_2 + c_4))\sigma^3 := c_7\sigma^3. \tag{31}$$

Now by assumption (22), the quantity (30) is at most

$$1 + \sigma^2(1-\delta)(c_3\sigma + c_5\sigma - \delta\tilde{c}/2) + c_7\sigma^3$$

where the term in the brackets is negative by the lower bound in (25), so by the upper bound in (25) we get that

$$1 + \frac{\sigma^2}{2}(c_3\sigma + c_5\sigma - \delta\tilde{c}/2) + c_7\sigma^3 \leq 1 - \kappa \leq e^{-\kappa},$$

where the first inequality is equivalent to the left inequality of (25). This completes the proof. \square

We will assume (and heavily use) for the rest of the paper that

$$\begin{aligned} \sigma &\leq \frac{2 \sin \theta |\sin 2\theta|}{10c_0^3}, \text{ implying } \sigma \leq \frac{4 \sin^2 \theta}{10c_0^3} = \frac{1}{10\rho^2 c_0^3}, \\ \sigma &\leq \frac{\sin \theta}{5c_0^3} = \frac{1}{10\rho c_0^3} \leq \frac{1}{10\rho c_0}. \end{aligned} \tag{32}$$

The last inequality, combined with the fact that c_0 , an absolute bound on a random variable of variance 1, satisfies

$$c_0 \geq 1$$

gives

$$\sigma c_0 \rho \leq \frac{1}{10}. \tag{33}$$

Lemma 17. *If $E(\omega_j) = 0$, $E(\omega_j^2) = 1$, $|\omega_j| \leq c_0$, then there exist functions F_k and G_k satisfying*

$$|F_k| \leq 4\rho^3, \quad |G_k| \leq \frac{2\rho^2}{|\sin 2\theta|}$$

so that for σ satisfying (32), $\kappa \in [0, 1]$ and δ satisfying

$$\frac{\kappa}{\sigma^2 \rho^2} + 242 \frac{c_0^3 \rho \sigma}{|\sin 2\theta|} \leq \delta \leq \frac{1}{2}$$

we have that with

$$r_k = \prod_{j=1}^{k-1} \left(1 + 2\sigma^2 \omega_j^2 \rho^2 + 2\sigma \omega_j \rho \sin(2\alpha_{j-1} + 2\theta) - 2\sigma^2 \omega_j^2 \rho^2 \cos(2\alpha_{j-1} + 2\theta) \right)$$

the following process is a supermartingale with respect to $\mathcal{F}_k = \sigma(\omega_1, \dots, \omega_k)$

$$e^{(F_{k-1} - (1-\delta/2)G_{k-1})\sigma^2(1-\delta)} (e^{-\kappa k} r_k)^{(\delta-1)}. \tag{34}$$

Proof. First compute

$$\begin{aligned} E \left(2\sigma^2 \omega_j^2 \rho^2 + 2\sigma \omega_j \rho \sin(2\alpha_{j-1} + 2\theta) - 2\sigma^2 \omega_j^2 \rho^2 \cos(2\alpha_{j-1} + 2\theta) \middle| \mathcal{F}_{j-1} \right) \\ = 2\sigma^2 \rho^2 (1 - \cos(2\alpha_{j-1} + 2\theta)) \end{aligned} \tag{35}$$

and define

$$B_{i-1} = 2\rho^2 (1 - \cos(2\alpha_{i-1} + 2\theta)). \tag{36}$$

Clearly

$$|B_{i-1}| \leq 4\rho^2.$$

Moreover, the random variable in (35) is absolutely bounded above by

$$4\sigma^2 c_0^2 \rho^2 + 2\sigma c_0 \rho := c_1 \sigma \leq \frac{6}{25} < \frac{1}{4}, \quad \text{note: } c_1 \leq \frac{12}{5} c_0 \rho$$

where the inequality comes from (33). Write $\Sigma = \sum_{j=1}^k e^{2i\alpha_j}$, and sum (21) between 1 and $k - 1$ to get

$$\Sigma - e^{2i\alpha_1} = z^2(\Sigma - e^{2i\alpha_k}) + \sigma \sum_{j=1}^{k-1} \frac{\omega_{j+1} i \rho (z^2 e^{2i\alpha_j} - 1)^2}{1 + \sigma \omega_{j+1} i \rho (1 - z^2 e^{2i\alpha_j})}.$$

Call the sum on the right $\tilde{\Sigma}$. By (33), $\sigma \rho |\omega_j| \leq 1/10$, and the denominator is bounded below in absolute value by $4/5$. The terms in $\tilde{\Sigma}$ are thus bounded above in absolute value by $\frac{4c_0 \rho}{4/5} = 5c_0 \rho$. Rearranging gives

$$\Sigma = \frac{e^{2i\alpha_1} - z^2 e^{2i\alpha_k} + \sigma \tilde{\Sigma}}{1 - z^2}$$

and multiplying everything by $-2\rho^2 z^2 = -2\rho^2 e^{2i\theta}$ and taking the real part of both sides gives

$$-2\rho^2 \sum_{j=1}^k \cos(2\alpha_j + 2\theta) = -2\rho^2 \operatorname{Re} z^2 \frac{e^{2i\alpha_1} - z^2 e^{2i\alpha_k}}{1 - z^2} - 2\rho^2 \operatorname{Re} z^2 \frac{\sigma \tilde{\Sigma}}{1 - z^2}.$$

Call the first term on the right hand side F_k . We have

$$|\Delta F_k| \leq \frac{4\rho^2}{|1 - z^2|} = 4\rho^3 =: c_2, \quad |F_k| \leq \frac{4\rho^2}{|1 - z^2|} = 4\rho^3.$$

Moreover we have

$$|B_k - \Delta F_k - 2\rho^2| = |2\rho^2 \operatorname{Re} z^2 \frac{\sigma \Delta \tilde{\Sigma}_k}{1 - z^2}| \leq 10c_0 \rho^4 \sigma =: c_3 \sigma.$$

Now compute

$$\begin{aligned} & E \left((2\sigma^2 \omega_j^2 \rho^2 + 2\sigma \omega_j \rho \sin(2\alpha_{j-1} + 2\theta) - 2\sigma^2 \omega_j^2 \rho^2 \cos(2\alpha_{j-1} + 2\theta))^2 | \mathcal{F}_{j-1} \right) \\ &= E \left(4\sigma^4 \omega_j^4 \rho^4 (1 - 2\cos(2\alpha_{j-1} + 2\theta) + \cos^2(2\alpha_{j-1} + 2\theta)) \right. \\ &\quad \left. + 8\sigma^3 \omega_j^3 \rho^3 (\sin(2\alpha_{j-1} + 2\theta) - \sin(2\alpha_{j-1} + 2\theta) \cos(2\alpha_{j-1} + 2\theta)) \right. \\ &\quad \left. + 4\sigma^2 \omega_j^2 \sin^2(2\alpha_{j-1} + 2\theta) | \mathcal{F}_{j-1} \right) \\ &\leq E \left(16\sigma^4 \omega_j^2 c_0^2 \rho^4 + 16\sigma^3 \omega_j^2 c_0 \rho^3 + 4\sigma^2 \omega_j^2 \rho^2 \sin^2(2\alpha_{j-1} + 2\theta) | \mathcal{F}_{j-1} \right) \\ &= 16\sigma^4 c_0^2 \rho^4 + 16\sigma^3 c_0 \rho^3 + 2\sigma^2 \rho^2 (1 - \cos(4\alpha_{j-1} + 4\theta)) \\ &\leq 17.6\sigma^3 c_0 \rho^3 + 2\sigma^2 \rho^2 (1 - \cos(4\alpha_{j-1} + 4\theta)) \end{aligned} \tag{37}$$

where the first inequality comes from bounding the sine and cosine terms, and from $|\omega_j| \leq c_0$ (by definition), and the second inequality comes from (33). Now define

$$A_{j-1} = 17.6\sigma c_0 \rho^3 + 2\rho^2 - 2\rho^2 \cos(4\alpha_{j-1} + 4\theta) \tag{38}$$

which is upper bounded as

$$A_{i-1} \leq 17.6\sigma c_0 \rho^3 + 4\rho^2 \leq 1.76\rho^2 + 4\rho^2 \leq 6\rho^2$$

using (33). Write $\Sigma = \sum_{j=1}^k e^{4i\alpha_j}$, and square both sides of (21), then sum from 1 to $k - 1$ to get

$$\begin{aligned} \Sigma - e^{4i\alpha_1} &= z^4 (\Sigma - e^{4i\alpha_k}) + \sigma \sum_{j=1}^{k-1} \left(\sigma \frac{-\omega_{j+1}^2 \rho^2 (z^2 e^{2i\alpha_j} - 1)^4}{(1 + \sigma \omega_{j+1} i \rho (1 - z^2 e^{2i\alpha_j}))^2} \right. \\ &\quad \left. + 2e^{2i\alpha_k} z^2 \frac{\omega_{k+1} i \rho (z^2 e^{2i\alpha_k} - 1)}{1 - \sigma \omega_{k+1} i \rho (1 - z^2 e^{2i\alpha_k})} \right). \end{aligned}$$

Call the sum on the right $\tilde{\Sigma}$. The terms in $\tilde{\Sigma}$ are bounded by

$$\sigma c_0^2 \rho^2 2^4 / (4/5)^2 + 4c_0 \rho / (4/5) \leq 8c_0 \rho$$

again making use of (33) multiple times. Rearranging gives

$$\Sigma = \frac{e^{4i\alpha_1} - z^4 e^{4i\alpha_k} + \sigma \tilde{\Sigma}}{1 - z^4} \tag{39}$$

and multiplying everything by $-2\rho^2 z^4$ and taking the real part of both sides gives

$$-2\rho^2 \sum_{j=1}^k \cos(4\alpha_j + 4\theta) = -2\rho^2 \operatorname{Re} z^4 \frac{e^{4i\alpha_1} - z^4 e^{4i\alpha_k}}{1 - z^4} - 2\rho^2 \operatorname{Re} z^4 \frac{\sigma \tilde{\Sigma}}{1 - z^4}. \tag{40}$$

Call the first term on the right hand side G_k . We have

$$|\Delta G_k| \leq \frac{4\rho^2}{|1 - z^4|} = \frac{2\rho^2}{|\sin 2\theta|} =: c_4, \quad |G_k| \leq \frac{2\rho^2}{|\sin 2\theta|}$$

Moreover we have

$$\begin{aligned} |A_k - \Delta G_k - 2\rho^2| &= |2\rho^2 \operatorname{Re} z^4 \frac{\sigma \Delta \tilde{\Sigma}_k}{1 - z^4} + 17.6\sigma c_0 \rho^3| \\ &\leq 2\rho^2 \sigma |\operatorname{Re} z^4 \frac{\Delta \tilde{\Sigma}_k}{1 - z^4}| + 17.6\sigma c_0 \rho^3 \\ &\leq \frac{8\sigma c_0 \rho^3}{|\sin 2\theta|} + 17.6\sigma c_0 \rho^3 \leq \frac{26c_0 \rho^3}{|\sin 2\theta|} \sigma =: c_5 \sigma. \end{aligned}$$

We now collect constants:

$$\begin{aligned} c_1 &= \frac{12}{5} c_0 \rho, \quad c_2 = 4\rho^3, \quad c_3 = 10c_0 \rho^4, \quad c_4 = \frac{2\rho^2}{|\sin 2\theta|}, \quad c_5 = \frac{26c_0 \rho^3}{|\sin 2\theta|}, \\ \tilde{c} &= 2\rho^2, \quad c_6 := \frac{10c_0^3}{2 \sin \theta |\sin 2\theta|}. \end{aligned}$$

We now apply Lemma 16. The condition (24) on σ is easily satisfied by (32). For the condition (25), we use the inequality $1/2 \leq \rho \leq 1/|\sin 2\theta|$ and $1 \leq c_0$ to get the bound

$$\begin{aligned} &c_3 + c_5 + 6c_1^3 + 2(c_2 + c_4)^2/c_6 + 8c_1(c_2 + c_4) \\ &\leq \frac{c_0^3 \rho^3}{|\sin 2\theta|} \left(10 + 26 + 6\left(\frac{12}{5}\right)^3 + \frac{2}{10}(4+2)^2 + 2\frac{48}{5}(4+2) \right). \end{aligned}$$

The constant above is less than 242. The claim follows. \square

Lemma 18. For a positive supermartingale X_t

$$P = P(\exists t \text{ s.t. } X_t \geq B\mathbb{E}X_0) \leq 1/B.$$

Proof. Let τ be the first time that $X_t \geq B\mathbb{E}X_0$, and let $p_T = P(X_{(\tau \wedge T)} \geq B\mathbb{E}X_0)$. Then by optional stopping

$$\mathbb{E}X_0 \geq E(X_{\tau \wedge T}) \geq E(X_{\tau \wedge T}; X_{\tau \wedge T} \geq B\mathbb{E}X_0) \geq p_T B\mathbb{E}X_0.$$

But $p_T \uparrow P$. \square

4.3. Proof of Theorem 3.

Proof. We consider the functions F_k, G_k in Lemma 17. To simplify notation, let $f_{k,\delta} = (F_{k-1} - (2 - \delta/2) G_{k-1})$, and note that

$$|f_{k,\delta}| \leq \frac{6\rho^3}{|\sin 2\theta|} =: \bar{c}/2.$$

Lemma 15 tells us that $r_k Y_k = \sin \theta$, and under the conditions of Lemma 17 (which are satisfied by assumption), the process

$$e^{f_{k,\delta}\sigma^2(1-\delta)} \left(r_k e^{-\kappa k} \right)^{(\delta-1)} = \left(e^{f_{k,\delta}\sigma^2} \frac{Y_k}{\sin \theta} e^{\kappa k} \right)^{(1-\delta)}$$

is a positive supermartingale. Note also that due to the invariance of the supermartingale property, the $\sin \theta$ can be dropped from the right side of the above expression. Now choose

$$\delta = \frac{\kappa}{\sigma^2 \rho^2} + 242 \frac{c_0^3 \rho \sigma}{|\sin 2\theta|}.$$

Then by our bound on κ we have that

$$\delta \leq \frac{248c_0^3 \rho \sigma}{|\sin 2\theta|}$$

and by our bound on σ we have that

$$\delta \leq 1/2$$

so that the conditions of Lemma 18 are satisfied. Then Lemma 18 gives

$$P \left(\exists n : \left(e^{f_{n,\delta}\sigma^2} Y_n e^{\kappa n} \right)^{1-\delta} \geq \left(e^{f_{1,\delta}\sigma^2} Y_1 e^{\kappa} e^{B-\bar{c}\sigma^2} \right)^{1-\delta} \right) \leq e^{-(B-\bar{c}\sigma^2)(1-\delta)}.$$

Taking logs, the event above is equivalent to

$$\{\exists n : \log Y_n - \log Y_1 + \kappa (n - 1) \geq B - \bar{c}\sigma^2 + (f_{1,\delta} - f_{n,\delta}) \sigma^2\}$$

which is a subevent of

$$\{\exists n : \log Y_n - \log Y_1 + \kappa (n - 1) \geq B\}.$$

So the probability that the process $\log Y_n + \kappa n$ has a backtrack of size B starting from time 1 is at most $e^{-(B-\bar{c}\sigma^2)(1-\delta)}$. But

$$e^{-(B-\bar{c}\sigma^2)(1-\delta)} \leq e^{-B(1-\delta)} e^{\bar{c}\sigma^2}$$

and the bound on σ gives

$$\bar{c}\sigma^2 \leq \frac{12\rho^3\sigma^2}{|\sin 2\theta|} \leq \frac{12}{496^2}$$

so that

$$e^{\bar{c}\sigma^2} \leq 2.$$

Now by $\kappa \leq 6c_0^3\rho^3\sigma^3/|\sin 2\theta|$ and our choice of δ , we have

$$e^{-B(1-\delta)} \leq e^{-B \left(1 - \frac{248c_0^3}{2\sin\theta|\sin 2\theta|} \sigma \right)}$$

meaning the probability that the process $\log Y_n + \kappa n$ has a backtrack of size B starting from time 1 is at most $2e^{-B(1-248c_0^3\sigma/2\sin\theta|\sin 2\theta|)}$. \square

5. Proof of Theorem 1

Proof. Assume that \mathbb{P} has support bounded by c_0 . Recall that N_n is the number of eigenvalues of the operator $H_{\omega,n}$ in the interval $[\lambda_0, \lambda_0 + \lambda]$. Let $\lambda_0 \in (-2, 0) \cup (0, 2)$, $n \in \mathbb{N}$, $\lambda > 0$ and let $\sigma \leq \frac{2 \sin \theta |\sin 2\theta|}{496c_0^3}$, so it satisfies the conditions of Theorem 3.

Further, let $M = \frac{\sqrt{5}\sigma c_0^2}{2 \sin^2 \theta} \leq 1/2$, $\epsilon = 1$ and $\beta = \sigma^3$. We may assume that $\lambda \leq \sigma^3$, because otherwise the bound is trivial.

Choose $\kappa = \epsilon(\lambda + \beta) / \sin \theta + 2M\beta$. Then

$$\kappa \leq (\sigma^3 + \sigma^3) / \sin \theta + \sigma^3 \leq 3\sigma^3 / \sin \theta \leq 6c_0^3 \rho^3 \sigma^3 / |\sin 2\theta|.$$

By our choices above, and by Theorem 12, the conditions of Theorem 2 are satisfied. So by Theorem 2 we have that

$$\begin{aligned} N_n &\leq 1 + \text{the number of backtracks of size at least } \log(\epsilon\beta/\lambda) \text{ of } \log Y_n + \kappa n \\ &\leq 1 + \sum_{k=1}^n \mathbb{1}(\log Y_n + \kappa n \text{ has a backtrack of size } \log(\epsilon\beta/\lambda) \text{ starting at } k). \end{aligned}$$

Taking expectations and dividing both sides by n yields

$$\frac{1}{n} EN_n \leq \frac{1}{n} (1 + nP(\log Y_n + \kappa n \text{ has a backtrack of size } \log(\epsilon\beta/\lambda))).$$

Now set $B = \log(\epsilon\beta/\lambda)$. Applying Theorem 3 gives

$$\begin{aligned} \frac{1}{n} EN_n &\leq \frac{1}{n} + 2e^{-B(1-248c_0^3\sigma/2 \sin \theta |\sin 2\theta|)} \\ &= \frac{1}{n} + 2 \left(\frac{\lambda}{\epsilon\beta} \right)^{1-248c_0^3\sigma/2 \sin \theta |\sin 2\theta|} \\ &\leq \frac{1}{n} + \frac{2}{\sigma^3} \lambda^{1-248c_0^3\sigma/2 \sin \theta |\sin 2\theta|}. \end{aligned}$$

Taking the limit as $n \rightarrow \infty$ yields

$$\mu(\lambda_0, \lambda_0 + \lambda) \leq \frac{2}{\sigma^3} \lambda^{1-248c_0^3\sigma/2 \sin \theta |\sin 2\theta|}.$$

Now we use that

$$\begin{aligned} |2 \sin \theta \sin 2\theta| &= |2(\cos \theta)2 \sin^2 \theta| \\ &= |\lambda_0|(4 - \lambda_0^2)/2 \\ &= \frac{1}{2} |\lambda_0| |2 - |\lambda_0||2 + |\lambda_0|| \\ &\geq |\lambda_0| |2 - |\lambda_0|| \end{aligned}$$

so we have

$$\mu(\lambda_0, \lambda_0 + \lambda) \leq \frac{2}{\sigma^3} \lambda^{1-248c_0^3\sigma/|\lambda_0||2-|\lambda_0||}.$$

And note that

$$|\lambda_0||2 - |\lambda_0|| \geq \min(|\lambda_0|, 2 - |\lambda_0|)$$

so for λ_0 in $(-2 + \gamma, -\gamma) \cup (\gamma, 2 - \gamma)$,

$$|\lambda_0||2 - |\lambda_0|| \geq \gamma$$

giving

$$\mu(\lambda_0, \lambda_0 + \lambda) \leq \frac{2}{\sigma^3} \lambda^{1-248c_0^3\sigma/\gamma}.$$

Now the condition on σ gives

$$\begin{aligned} 496c_0^3\sigma &\leq 2 \sin \theta | \sin 2\theta | \\ &= 2 \sin^2 \theta | 2 \cos \theta | \\ &= |\lambda_0| 2 \sin^2 \theta \\ &= |\lambda_0| \frac{4 - \lambda_0^2}{2} \end{aligned}$$

so it is equivalent to

$$\lambda_0(4 - \lambda_0^2) \geq 992c_0^3\sigma. \tag{41}$$

If this condition is violated, we have that

$$\frac{992c_0^3\sigma}{\gamma} \geq \frac{992c_0^3\sigma}{|\lambda_0||2 - |\lambda_0||} \geq |2 + |\lambda_0|| \geq 2$$

meaning

$$1 - 496c_0^3\sigma/\gamma \leq 0.$$

This means that by allowing an extra factor of 2 in the constant of the exponent of λ , the bound on the IDS is trivially satisfied for λ_0 violating (41). In other words, if we loosen our bound on the IDS from

$$\mu(\lambda_0, \lambda_0 + \lambda) \leq \frac{2}{\sigma^3} \lambda^{1-248c_0^3\sigma/|\lambda_0||2-|\lambda_0||}$$

to

$$\mu(\lambda_0, \lambda_0 + \lambda) \leq \frac{2}{\sigma^3} \lambda^{1-496c_0^3\sigma/|\lambda_0||2-|\lambda_0||}$$

we may drop the condition on λ_0 . This completes the proof. \square

Acknowledgements. We thank Andrew Stewart and Jeffrey Schenker for careful reading and useful comments on earlier versions. This research was supported by the Canada Research Chair program and the NSERC Discovery Accelerator Grant of the second author.

References

1. Bourgain, J.: On localization for lattice Schrödinger operators involving Bernoulli variables. In: Milman, V.D., Schechtman, G. (eds.) Geometric Aspects of Functional Analysis. Lecture Notes in Mathematics, vol. 1850, pp. 77–99. Springer, Berlin (2004)
2. Bourgain, J.: On the Furstenberg measure and density of states for the Anderson–Bernoulli model at small disorder. J. d’Analyse Math. **117**(1), 273–295 (2012)

3. Bourgain, J., Schlag, W.: Anderson localization for Schrödinger operators on \mathbb{Z} with strongly mixing potentials. *Commun. Math. Phys.* **215**(1), 143–175 (2000)
4. Carmona, R., Klein, A., Martinelli, F.: Anderson localization for Bernoulli and other singular potentials. *Commun. Math. Phys.* **108**(1), 41–66 (1987)
5. Carmona, R., Lacroix, J.: *Spectral Theory of Random Schrödinger Operators*. Probability and Its Applications. Birkhäuser, Boston (1990)
6. Gol'dshtein, I., Molchanov, S., Pastur, L.: A pure point spectrum of the stochastic one-dimensional Schrödinger operator. *Funct. Anal. Appl.* **11**(1), 1–8 (1977)
7. Halperin, B.I.: *Properties of a Particle in a One-Dimensional Random Potential*, pp. 123–177. Wiley, London (1967)
8. Kirsch, W.: *An Invitation to Random Schroedinger operators*. Technical report. [arXiv:0709.3707](https://arxiv.org/abs/0709.3707) (2007)
9. Kotowski, M., Virág, B.: Dyson's spike for random Schroedinger operators and Novikov–Shubin invariants of groups. ArXiv e-prints. [arXiv:1602.06626](https://arxiv.org/abs/1602.06626) (2016)
10. Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux différences finies aléatoires. *Commun. Math. Phys.* **78**(2), 201–246 (1980)
11. Minami, N.: Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Commun. Math. Phys.* **177**(3), 709–725 (1996)
12. Shubin, C., Vakilian, R., Wolff, T.: Some harmonic analysis questions suggested by Anderson–Bernoulli models. *Geom. Funct. Anal. GAFA* **8**(5), 932–964 (1998)
13. 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)
14. Wegner, F.: Bounds on the density of states in disordered systems. *Z. Phys. B Condens. Matter* **44**(1), 9–15 (1981)

Communicated by L. Erdős