On discrete Schrödinger operators with stochastic potentials

W. Schlag Caltech

1 Introduction

Given an ergodic map $T : \mathbb{T}^d \to \mathbb{T}^d$, a potential $V_x(n) = v(T^n x)$, where $v : \mathbb{T}^d \to \mathbb{R}$ is analytic and nonconstant, define the Hamiltonian

$$H_x := -\triangle_{\mathbb{Z}} + \lambda V_x. \tag{1}$$

This is a version of the well-known *Anderson's model.* In the late 1950's, Phil Anderson predicted that random impurities could turn conductors into insulators. In mathematical terms, he predicted that random potentials should lead to pure point spectrum with rapidly decreasing eigenfunctions — at least for large disorders. Instead of random potentials, one considers here potentials given in terms of deterministic dynamics. The "randomness" is given by the parameter $x \in \mathbb{T}^d$.

Basic questions: Does (1) display pure point spectrum for large λ, and does Anderson localization take place (exponentially decaying eigenfunctions)? One can also ask about dynamical localization. This refers to the property $\sup_t \|(1+|n|)^A e^{itH}\psi_0\|_2 < \infty$ for all $A > 0$, provided ψ_0 decays very rapidly. Other topics of interest is the limiting distribution of the eigenvalues of H_x restricted to $[-N, N]$ as $N \to \infty$ (this is known as the integrated density of states or IDS). Other very interesting questions concern the presence of a.c. spectrum for small λ, as well as the structure of the spectrum (Cantor set).

We cannot review the long and rich history of this subject here, but rather refer the reader to the monographs by Carmona and Lacroix [9] as well as Figotin and Pastur [11] for this purpose. Another resource which is closely related to this note is the forthcoming book by J. Bourgain [3]. Let us merely mention the fundamental results by Dinaburg, Sinai [10] (a.c. spectrum for small λ), Fürstenberg [14] (positivity of the Lyapunov exponent), Goldsheid, Molchanov, Pastur [15] (A.L. for the one-dimensional model), Fröhlich, Spencer [12] ("multiscale analysis" in the random case), Aizenman, Molchanov [1] ("fractional moment method"), Fröhlich, Spencer, Wittwer [13] and Sinai [18] (A.L. for the quasiperiodic model), as well as the many deep and important contributions by Avron, Bellissard, Campanino, Carmona, Delyon, Eliasson, Figotin, Gordon, Jitomirskaya, Kirsch, Klein, Kotani, Last, Martinelli, Oseledec, Pastur, Ruelle, Simon, Simon-Wolff, Souillard, Thouless, Wegner.

We now consider the eigenvalue equation

$$(H_x\psi)_n = -\psi_{n+1} - \psi_{n-1} + \lambda v(T^n x)\psi_n = E\psi_n. \tag{2}$$

Examples of maps T are: the shift $Tx = x + \omega$ (mod$\mathbb{Z}^d$), the doubling map $Tx = 2x$ (mod1), and the skew-shift $T(x, y) = (x + y, y + \omega)$ (mod$\mathbb{Z}^2$). It follows from the covariance relation

$$H_{Tx} = UH_xU^* \text{ with } U\text{=left shift}$$

that $\Sigma_x := \mathrm{spec}(H_x)$ is constant for a.e. x. The same also holds for the spectral parts $\Sigma_x^{ac}, \Sigma_x^{pp}, \Sigma_x^{sc}$.

Recall the following basic objects in the study of (2). The monodromy matrices and Lyapunov exponent are defined as

$$\begin{aligned} M_n(x,\lambda,E) &:= \prod_{j=n}^{1} \begin{bmatrix} \lambda v(T^j x) - E & -1 \\ 1 & 0 \end{bmatrix} \\ L_n(\lambda,E) &:= \frac{1}{n}\int_{\mathbb{T}^d} \log \|M_n(x,\lambda,E)\|\, dx \\ L(\lambda,E) &:= \lim_{n\to\infty} L_n(\lambda,E) = \lim \frac{1}{n}\log\|M_n(x,\lambda,E)\| \end{aligned}$$

where the last relation holds for a.e. x by Kingman's subadditive ergodic theorem. Furthermore, by the Ishi-Pastur theorem $\Sigma_x^{ac} \subset \{E : L(\lambda,E) = 0\}$, whereas Kotani's converse statement is $\mathrm{meas}\big[\{E : L(\lambda,E)=0\} \setminus \Sigma_x^{ac}\big] = 0$, a.e. x.

Let $Tx = T_\omega x := x + \omega \pmod 1$ be the one-dimensional shift. We denote the Lyapunov exponent by $L(\omega,\lambda,E)$ to emphasize the dependence on ω. The following theorem is proved in [4].

Theorem 1 (Bourgain-Goldstein). *Suppose* $\inf_{\omega,E} L(\omega,\lambda,E) > 0$. *Then for a.e.* ω *and a.e.* x *the Hamiltonian* H_x *displays Anderson localization, as well as dynamical localization.*

Note: This *does not* explicitly require large λ. If $v(x) = \cos(2\pi x)$ (the almost Mathieu operator), then Theorem 1 applies for $\lambda > 2$ (by Herman's result $L(\omega,\lambda,E) \geq \log(\lambda/2)$ for $\lambda > 2$). Jitomirskaya [16] proved this result for the almost Mathieu operator and all *Diophantine* ω and a.e. x. The proofs of both Theorem 1 as well as [16] are nonperturbative, and are based on the transfer matrix formalism. The following theorem is from [17].

Theorem 2 (Goldstein-S.). *Suppose* ω *is Diophantine (which means here that* $\|n\omega\| \geq \frac{C_\omega}{n(\log n)^a}$ *for* $n > 1$, $a > 1$*). If* $L(\omega,E) > 0$ *for all* $E_1 < E < E_2$*, then* $L(\omega,\cdot)$ *as well as the integrated density of states* $N(\omega,\cdot)$ *are Hölder continuous in* $E \in [E_1,E_2]$.

Note: The integrated density of states (IDS) is the *deterministic* limit

$$N(\omega,E) = \lim_{N\to\infty} \frac{1}{2N+1} \#\{1 \leq j \leq 2N+1 \,:\, E_j^{(N)} < E\}$$

where $E_j^{(N)} = E_j^{(N)}(x,\omega)$ are the eigenvalues of H_x restricted to $[-N,N]$. Thouless's formula relates the Lyapunov exponent L and the IDS N:

$$L(\omega,E) = \int \log|E-E'|\, N(\omega,dE').$$

Sinai's work on large disorder [18] shows that the IDS can be no better than Hölder–$\frac{1}{2}$ continuous, and Bourgain [3] obtained the exponent $\frac{1}{2}-$ for the almost Mathieu operator and large λ. In [17], the Hölder exponent depended on the Lyapunov exponent. Bourgain refined [17] to obtain a Hölder exponent that is uniform in $L(E)$. Bourgain and Jitomirskaya [7] modified the method from [17] to show that $L(\omega,E)$ is jointly continuous in ω, E at every point (ω_0,E), $\omega_0 \in \mathbb{R}\setminus\mathbb{Q}$ as well as continuous in E for every ω (without assuming that L is positive).

The skew-shift on $\mathbb{T}^2$ is defined to be $T_\omega(x,y) = (x+y, y+\omega) \pmod{\mathbb{Z}^2}$. Note the quadratic dependence on n in the nth iterate

$$T_\omega^n(x,y) = (x+ny+n(n-1)\omega/2, y+n\omega) \pmod{\mathbb{Z}^2}.$$

It is conjectured to lead to $L(\omega,\lambda,E) > 0$ for all $\lambda > 0$. The following theorem is proved in [5].

Theorem 3 (BGS). *Fix $\varepsilon > 0$. Then there exist $\lambda_0(\varepsilon)$ and $\Omega(\lambda,\varepsilon) \subset \mathbb{T}^3$ such that $\mathrm{meas}[\mathbb{T}^3 \setminus \Omega(\lambda,\varepsilon)] < \varepsilon$ and for all $\lambda > \lambda_0(\varepsilon)$ and $(\omega,x,y) \in \Omega(\lambda,\varepsilon)$, $(H_{\omega,x,y}\psi)_n := -\psi_{n+1} - \psi_{n-1} + \lambda v(T_\omega^n(x,y))\psi_n$ displays Anderson localization.*

This theorem is related to the following "quantum kicked rotor" model:

$$i\partial_t \Psi(t,x) = a\partial_x^2 \Psi(t,x) + ib\partial_x \Psi(t,x) + V(t,x)\Psi(t,x)$$

where $V(t,x) = \kappa \cos(2\pi x)\sum_{n\in\mathbb{Z}} \delta(t-n)$. Let its unitary evolution operator be denoted by $S(t)$. Then the monodromy operator is defined as $W = S(1) = U_{a,b} \cdot W_{1,\kappa}$ where $U_{a,b} = e^{i(a\frac{d^2}{dx^2} + b\frac{d}{dx})}$, and $(W_{1,\kappa}f)(x) = e^{i\kappa\cos(2\pi x)}f(x) =: \rho(x)f(x)$ is a multiplication operator by the function ρ.

The action of these operators on Fourier modes is given by $U_{a,b} = e^{-i(4a\pi^2 n^2 + 2\pi b n)}\delta_{mn}$, $W_{1,\kappa}(m,n) = \hat{\rho}(m-n)$. Note that $\hat{\rho}$ decays exponentially, and $\hat{\rho}(0) = 1 + O(\kappa^2)$. Then the self-adjoint operator $H := \frac{1}{2}(W + W^*)$ is of the form $H_{nn} = v(T_\omega^n(0,y))$, $H_{mn} = \phi_{m-n}(T^m x) + \overline{\phi_{n-m}(T^n x)}$ for $m \neq n$, with ϕ exponentially decaying and small (for κ small). Therefore, H is a *long-range* version of the skew-shift Hamiltonian from Theorem 3. In analogy with Theorem 3, one can show that Anderson localization for H holds for small κ and most a, b. Hence, W also possesses an orthonormal basis of exponentially decaying eigenfunctions (where the exponential decay refers to the Fourier coefficients). More precisely, Bourgain proved the following result in [2].

Theorem 4 (B). *For small κ and (a,b) outside a set of small measure, one has: Let $\Psi(t,x) = (S(t)\Psi(0,\cdot))(x)$ be the solution of the kicked rotor equation. If $\Psi(0,\cdot)$ is a smooth function on $\mathbb{T}$, then $\Psi(t)$ is an almost-periodic $H^1(\mathbb{T})$-valued function and $\sup_t \|\Psi(t)\|_{H^1(\mathbb{T})} < \infty$.*

In recent work, Bourgain and Jitomirskaya have obtained a version of Theorem 4 that applies to a.e. choice of the parameters (a,b). Theorem 3 is proved by means of the transfer matrix formalism, whereas Theorem 4 cannot be obtained this way because of long-range interactions. One uses instead an approach that originated in the proof of the following theorem from [6].

Theorem 5 (BGS). *Let $v : \mathbb{T}^2 \to \mathbb{R}$ be analytic, nonconstant on any vertical and horizontal line segment. Let $(H_{xy}\psi)_{nm} = -\triangle_{\mathbb{Z}^2}\psi_{nm} + \lambda v(x+n\omega_1, y+m\omega_2)\psi_{nm}$ for $(n,m) \in \mathbb{Z}^2$. Then for all $\varepsilon > 0$, $\lambda > \lambda_0(\varepsilon)$ the operator H_{xy} displays Anderson localization for all $(x,y,\omega_1,\omega_2) \in \mathbb{T}^4$ up to a set of measure at most ε.*

2 Transfer matrix formalism

As in the case of truly random potentials, it is of basic importance to control the probability that a given energy E is close to an eigenvalue of the Hamiltonian H restricted to a box. In other words, one needs to estimate

$$\operatorname{meas}\big[x \in \mathbb{T}^d \,:\, \operatorname{dist}(H_x^\Lambda, E) < e^{-\rho}\big]$$

where $\Lambda = [-N, N]$ and ρ is relatively large (say, $\gg (\log N)^A$). By self-adjointness, this is the same as bounding the norm of the Green function $\big\|(H_x^\Lambda - (E + i0))^{-1}\big\| = \|G_\Lambda(x, E)\|$. By Cramer's rule, one has

$$G_\Lambda(x, E)(n, m) = \frac{f_{[-N,n-1]}(x, E) f_{[m+1,N]}(x, E)}{f_{[-N,N]}(x, E)}$$

where $f_{[p,q]}(x, E) = \det(H_x^{[p,q]} - E)$. Moreover, there is the well-known relation

$$M_n(x, E) = \begin{bmatrix} f_{[1,n]}(x, E) & -f_{[2,n]}(x, E) \\ f_{[1,n-1]}(x, E) & -f_{[3,n]}(x, E) \end{bmatrix}.$$

Recall that Kingman's theorem implies that $\frac{1}{N} \log \big\|M_N(x, E)\big\| \to L_N(E)$ for a.e. $x \in \mathbb{T}$. A quantitative version of this statement are the following estimates on the deviations: For large N

$$\operatorname{meas}\big[x \,:\, \big|\log \big\|M_N(x, E)\big\| - N\, L_N(E)\big| > N^\sigma\big] < e^{-N^\tau} \tag{3}$$

for some fixed $0 < \sigma, \tau < 1$. Such estimates are proved in [4], [17], [8], [5] for various different underlying ergodic transformations T. There they are referred to as large deviation theorems or LDT.

To motivate the proofs of these results, consider the following commutative model case:

$$u(x) = \sum_{n=1}^{N} \log |e(x + n\omega) - 1| = \int \log |z - \zeta| \, \mu(d\zeta)$$

where $z = e(x) = e^{2\pi i x}$, $\mu = \sum_{n=1}^{N} \delta_{e(-n\omega)}$. The left-hand side can be thought of as a Riemann sum. It is standard to estimate the error between such sums and their mean via *Koksma's inequality: Let* $\mathcal{S} := \{x_n\}_{n=1}^{N} \subset \mathbb{T}$

$$\Big|\sum_{n=1}^{N} f(x_n) - \int_0^1 f(x)\, dx\Big| \le C\, D_N(\mathcal{S}) \|f\|_{BV}$$

where $D_N(\mathcal{S}) = \sup_{J \subset \mathbb{T}} \big|\#\{n : x_n \in J\} - N|J|\big|$.

The problem here is, of course, that $\log |e(x) - 1| \notin BV$. A more direct approach is to observe that $u(x)$ is the Hilbert transform of a sum of saw-tooth functions. More precisely, let $f_0(x) = -\frac{1}{2} - x$ for $-\frac{1}{2} < x < 0$ and $f_0(x) = \frac{1}{2} - x$ for $0 < x < \frac{1}{2}$. Then, with $\mathcal{H}$ being the Hilbert transform,

$$\begin{aligned} u(x) &= \mathcal{H}\Big(\sum_{n=1}^{N} f_0(\cdot + n\omega)\Big)(x) \\ \|u\|_{BMO} &\le C\Big\|\sum_{n=1}^{N} f_0(\cdot + n\omega)\Big\|_\infty \sim D_N(\{n\omega\}_{n=1}^{N}). \end{aligned}$$

Since for a.e. ω and N large $D_N(\{n\omega\}_{n=1}^N) < N^\varepsilon$, the John-Nirenberg inequality implies

$$\operatorname{meas}\big[x\ :\ |u(x)-\langle u\rangle| > N^\sigma\big] \le \exp\big[-N^{\sigma-\varepsilon}\big].$$

Observe the analogy of this estimate with the large deviation theorem (3).

The transition to the noncommutative case is typically accomplished by means of two devices:

- subharmonic extensions of $\log\|M_N(x,E)|$
- the so-called "avalanche principle" (this requires positive exponents $L(E)>0$).

The former is well-known and straightforward: Since v is analytic, $u(x) = \log\|M_N(x,E)\|$ extends to a neighborhood of the unit circle as a subharmonic function. Such functions have a Riesz representation:

$$u(z) = \int \log|z-\zeta|\,\mu(d\zeta) + h(z)$$

with some measure $\mu\ge 0$, and a harmonic function h. Here $\|\mu\|\sim N$, but there is a lot of cancellation in the integral to ensure, for example, that $\|u\|_{BMO} \le (\log N)^A$. This latter property can only be captured by means of invoking the underlying dynamics. More precisely, one obtains the structure of u as a sum of shifts of another function by means of the following *Avalanche Principle (Goldstein-S.):*

Let $\{A_n\}_{n=1}^N \in S\ell(2,\mathbb{R})$. Suppose $\|A_n\| > \mu > N$, $\|A_{n+1}A_n\| \le \sqrt{\mu}\,\|A_{n+1}\|\,\|A_n\|$ for all n. Then

$$\Big|\log\|M_N\| + \sum_{n=2}^{N-1}\log\|A_n\| - \sum_{n=1}^{N-1}\log\|A_{n+1}A_n\|\Big| \le C\frac{N}{\mu}$$

where $M_N = A_N\cdots A_1$.

In the work by Bourgain, Goldstein, and the author, this device has been used to establish the following properties:

- positivity of the Lyapunov exponent for large disorder
- inductive proof of large deviation theorems
- Hölder continuity of the integrated density of states.

We now give a typical application of the avalanche principle. Write

$$M_{nN}(x,E) = \prod_{j=N-1}^{0} M_n(T^{jn}x,E), \quad \text{where} \ \ n\sim(\log N)^A.$$

Suppose, for $0<\sigma,\tau<1$ fixed,

$$\operatorname{meas}\big[x\ :\ \big|\log\big\|M_n(x,E)\| - n\,L_n(E)\big| > n^\sigma\big] < e^{-n^\tau} \tag{4}$$

and the same for $2n$, where $A\sigma\ge 100$. Furthermore, suppose $L_n(E)\ge L_{2n}(E)\ge 1$ for all E, and $L_n(E)-L_{2n}(E)\le 1/100$. Then up to a set of measure $< CNe^{-n^\sigma}\sim N^{-99}$, and for n large, one has for all j

$$\begin{aligned}
&\|M_n(T^{jn}x)\| > e^{n-n^\sigma} > e^{n/2} =: \mu\\
&\|M_n(T^{(j+1)n}x)M_n(T^{jn}x)\| > e^{2nL_{2n}(E)-n^\sigma}\\
&> e^{2nL_n(E)-2n^\sigma} > \mu^{-\frac12}\|M_n(T^{(j+1)n}x)\|\,\|M_n(T^{jn}x)\|.
\end{aligned}$$

From the avalanche principle, one concludes that up to a set of measure $< N^{-99}$,

$$\Big|\log\|M_{nN}(x,E)\| + \sum_{j=1}^{N-2}\log\|M_n(T^{jn}x,E)\| - \sum_{j=0}^{N-2}\log\|M_{2n}(T^{jn}x,E)\|\Big| < CNe^{-n/2}. \tag{5}$$

Typically, the sums are *uniformly* in x close to their means (since they represent averages over very *long orbits*). The conclusion then is that

$$u(x) := \frac{1}{Nn}\log\|M_{nN}(x,E)\| = u_0(x) + u_1(x)$$

where $\|u_0 - \langle u_0\rangle\|_\infty \le CN^{-1+\varepsilon} =: \varepsilon_0$, and $\|u_1\|_1 \le C\,N^{-90} =: \varepsilon_1$. Provided the Riesz mass of the subharmonic extension of $u(x)$ is $< N^{20}$, the *splitting lemma* (see below) yields $\|u\|_{BMO} \le C\big(\varepsilon_0 + \sqrt{N^{20}\varepsilon_1}\big) \le C\,N^{-1+\varepsilon}$. By means of the John-Nirenberg inequality, this implies that

$$\mathrm{meas}\big[x\ :\ \big|\log\big\|M_{nN}(x,E)\big\| - nN\,L_{nN}(E)\big| > (nN)^\sigma\big] < C\,e^{-c\,N^{\sigma-\varepsilon}},$$

which is (4) for N instead of n (provided $\tau < \sigma - \varepsilon$). The same argument applies to $2nN$. Moreover, averaging (5) over x yields that

$$|L_{nN}(E) + L_n(E) - 2L_{2n}(E)| < C\,N^{-1},$$

which implies $L_{nN}(E) \ge 1 - 2(L_n(E) - L_{2n}(E)) - C\,N^{-1}$. Continuing inductively leads to positivity of $L(E)$ as well (4) for all n.

The following result is the aforementioned splitting lemma from [5]:

Lemma 1. *Let u be subharmonic on a neighborhood of $\mathbb{T}$ with Riesz mass N. Suppose $u = u_0 + u_1$ on $\mathbb{T}$ with $\|u_0\|_{L^\infty(\mathbb{T})} = \varepsilon_0$ and $\|u_1\|_{L^1(\mathbb{T})} = \varepsilon_1$. Then $\|u\|_{BMO} \le C(\varepsilon_0 + \sqrt{N\varepsilon_1})$.*

To motivate this statement, consider N points $z_j = e(x_j)$ in $\mathbb{T}$. Suppose that $P(z) = \prod_{j=1}^N (z - z_j)$ satisfies $\sup_{|z|=1}|P(z)| < e^\tau$. We claim that $D_N(\{x_j\}_{j=1}^N) \le C\sqrt{N\tau}$ where D_N is the usual discrepancy (see above).
Proof: $u(x) = \log|P(e(x))| = \mathcal{H}\big(\sum_{j=1}^N f_0(\cdot - x_j)\big)$. Set $F := \sum_{j=1}^N f_0(\cdot - x_j)$. Let K_N be a smooth bump function $K_N(\theta) = K(N\theta)$, $K \ge 0, \int K = 1, \mathrm{supp}K \subset [-.01, .01]$. Then

$$(K_N * f_0)\big(\cdot - \frac{C}{N}\big) - \frac{C}{N} < f_0(\cdot) < (K_N * f_0)\big(\cdot + \frac{C}{N}\big) + \frac{C}{N}$$

Using $\int_{\mathbb{T}} u = 0$, $\|u\|_1 \le C\tau$, $F = \mathcal{H}^{-1}u = -\mathcal{H}u$, one now has

$$\begin{aligned}
&\|F\|_\infty \le 2\|\mathcal{H}^{-1}(u * K_M)\|_\infty + \frac{CN}{M}\\
&\le C\sqrt{M}\|\mathcal{H}^{-1}(u * K_M)\|_2 + \frac{CN}{M} \le C\,\sqrt{M}\|u * K_M\|_2 + \frac{CN}{M}\\
&\le C\,\sqrt{M}\|u\|_1\|K_M\|_2 + \frac{CN}{M} \le C\,M\tau + \frac{CN}{M}.
\end{aligned}$$

Setting $M = \sqrt{N/\tau}$ gives $\|F\|_\infty \le C\sqrt{N\tau}$ (and thus also $\|u\|_{BMO} \le C\sqrt{N\tau}$). Finally, it is easy to check that $\|F\|_\infty \sim D_N(\{x_j\}_{j=1}^N)$, as desired.

Note that this argument shows that if $u(z) = \int \log|z-\zeta|\, d\mu(\zeta)$ with $\mathrm{supp}(\mu) \subset \mathbb{T}$, $\sup_{\mathbb{T}} u \le \langle u\rangle + \tau$ (here $\langle u\rangle = 0$ and thus $\|u\|_1 \le \tau$), then $\|u\|_{BMO} \le C\sqrt{\|\mu\|\,\tau}$. A small variation of this argument proves Lemma 1 above.

Final remarks on the transfer matrix formalism:

- For the shift and skew-shift, large deviation theorems for $u(x) = \log\|M_N(x,E)\|$ are obtained via (a) subharmonicity (b) almost invariance $u(x) \approx u(Tx)$. The latter can be either be the simple invariance property $\sup_x |u(x) - u(Tx)| \le C$ (suffices for the shift) or the avalanche principle (needed for the skew-shift). The main difference between these cases is that the Riesz mass of u in the former is N, in the latter N^2 (or N^C for higher-dimensional versions of the skew-shift). The avalanche principle requires positive exponents, and the first step (but only that one) is perturbative. This means that both the initial positivity of the Lyapunov exponent, as well as the large deviation theorem at the initial scale are obtained by making the disorder very large. It is conceivable that the required information at the first step could also be provided by a numerical calculation. We would like to emphasize that for the plain shift, large deviation theorems do *not* require positive exponents and are non-perturbative. This allows Bourgain and Goldstein [4] to prove localization on the basis of $L(E) > 0$ alone.
- The avalanche principle gives positive Lyapunov exponents for large disorders under very general circumstances. But this either requires the Riesz mass of $\log\|M_N(x,E)\|$ to grow at most like N^C (so that the necessary large deviation theorem can be obtained simultaneously), or one needs to prove the LDT by other means at all scales. One case where the Riesz mass grows exponentially rather than polynomially is the doubling map $Tx = 2x$ (mod1). For this model, one proves the LDT by means of well-known subgaussian bounds for sums of martingale differences, see [8].
- In recent work, Goldstein and the author have obtained large deviation theorems for the entries of the monodromy matrix rather than the entire norm. Recall that the entries are the determinants $f_N(x,E)$.

3 Localization

The large deviation theorems from the previous section only control the probability of a single resonance at a fixed energy E, i.e.,

$$\mathrm{meas}\big[x \in \mathbb{T} \;:\; \mathrm{dist}(E, H_{x,\omega}^{[-N,N]}) < e^{-N^\sigma}\big].$$

It is well-known that this does not suffice in order to prove Anderson localization. Rather, one needs to control the probability of double resonances. More precisely, let $DC_{A,c} \subset \mathbb{T}$ denote the set of ω with $\|n\omega\| \ge \frac{c}{|n|^A}$ for all $n \ne 0$ and define

$$\begin{aligned}\delta_N := \mathrm{meas}\big[\omega \in DC_{A,c} \;:&\; \mathrm{dist}(E, \mathrm{spec}(H_{0,\omega}^{[-n,n]})) < e^{-n^\sigma},\\ &G_{[k,k+N]}(0,\omega,E) \text{ is bad for } \textit{some } E \text{ and } |k| \sim N^C\big].\end{aligned}$$

Here $n \sim (\log N)^A$, and $G_{[k,k+N]}(x,\omega,E)$ is *good* if both $\|G_{[k,k+N]}(x,\omega,E)\| < e^{N^\sigma}$ ($\sigma < 1$ fixed) and $|G_{[k,k+N]}(x,\omega,E)|(j,\ell) < e^{-\gamma N}$ if $k \le j, \ell \le N+k$, $|j-\ell| \sim N$

with $\gamma > 0$ a lower bound on the Lyapunov exponents. One then needs to show that $\sum_j \delta_{2^j} < \infty$. By means of standard methods (resolvent identity plus the polynomial Shnol-Simon bound on generalized eigenfunctions) this ensures that for a.e. $\omega \in DC_{A,c}$ the operators $H_{0,\omega}$ display Anderson localization. Finally, letting $c \to 0$ in the Diophantine condition allows one to extend this to a.e. $\omega \in \mathbb{T}$, thus proving Theorem 1.

To estimate δ_N, note that

$$\delta_N \le \sum_{E \in \mathrm{spec}(H_{0,\omega}^{[-n,n]})} \mathrm{meas}\big[\omega \in DC_{A,c} \,:\, \log \|M_N(k\omega, \omega, E)\| < NL_N(\omega, E) - N^\sigma, \text{ for some } |k| \sim N^C\big].$$

The set on the right-hand side is the projection onto the ω-axis of the intersections $\Omega_N := \bigcup_{|k|\sim N^C} \mathcal{S}_N \cap \ell_k$ of the lines $\ell_k := \{(\omega, k\omega) \,:\, \omega \in \mathbb{T}\}$ with

$$\mathcal{S}_N := \big\{(\omega, x) \in DC_{A,c} \times \mathbb{T} \,:\, \log \|M_N(x, \omega, E)\| < NL_N(\omega, E) - N^\sigma\big\}.$$

The measure of $\mathcal{S}_N$ is very small by the LDT, but this by itself does not preclude $\Omega_N = \mathbb{T}$ (consider the case where $\mathcal{S}_N$ is one of the lines ℓ_k). We need to know that the intersections of $\mathcal{S}_N$ with any horizontal line consist of a small number (say N^C many) connected components. This property is given by the Milnor-Thom bound of d^C on the number of connected components of semi-algebraic sets of degree d, at least if the potential v is a trigonometric polynomial (the case of general analytic v then follows by approximation). Indeed, if v is a trigonometric polynomial, then the horizontal sections of $\mathcal{S}_N$ are contained in semialgebraic sets of very small measure and degree N^C, as can be seen from the fact that the entries of M_N are polynomials in ω of degree $< N^C$. From this complexity bound and the LDT bound get $\delta_N \le Cn \,\mathrm{meas}[\Omega_N] < N^{-\tau}$ for some $\tau > 0$ by means of the following lemma from [4].

Lemma 2. *Let $\mathcal{S} \subset \mathbb{T}^2$ be such that $\{\omega \in \mathbb{T} : (\omega, x) \in \mathcal{S}\}$ consists of at most M intervals for every $x \in \mathbb{T}$. Then*

$$\begin{aligned} &\mathrm{meas}\big[\omega \in \mathbb{T} \,:\, (\omega, k\omega) \in \mathcal{S} \;\; \textit{for some } |k| \sim N\big] \\ &\le C\, N^C (\mathrm{meas}[\mathcal{S}])^{\frac{1}{2}} + C\, MN^{-1}. \end{aligned}$$

Concluding remarks:

• Localization can be obtained by this method for shifts of any dimension, as well as the skew-shift. In the latter case the main difficulty is the LDT (only known for large disorders).

• The long-range case as well as the Laplacian on $\mathbb{Z}^2$ cannot be treated by the transfer matrix formalism, so no LDT or avalanche principle are available. In those cases one needs to develop estimates on the probability that a given Green's function is bad (in the spirit of Fröhlich-Spencer's multiscale method), which again relies on subharmonic function arguments and reductions to smaller scales.

• Establishing Hölder regularity of the IDS requires a sharp LDT

$$\mathrm{meas}\big[x \in \mathbb{T} : |\log \|M_N(x, E) - NL_N(E)| > \delta N\big] < e^{-c_\delta N}.$$

This LDT is known, but *only* for the shift on $\mathbb{T}$. Averaging the avalanche principle over x by means of this LDT yields

$$|L_N(E) + L_n(E) - 2L_{2n}(E)| < N^{-1+\varepsilon} \text{ where } n \sim \log N.$$

Thus $|L(E) - L_N(E)| \le CN^{-1+\varepsilon}$ and

$$|L_N(E) - L_N(E')| \le CN^{-1+\varepsilon} + e^{C_2 n}|E - E'|.$$

Hence $|L(E) - L(E')| \le C|E - E'|^{\alpha}$, for some $\alpha > 0$, as desired.

Open problems:

• Hölder regularity of IDS for shifts on $\mathbb{T}^d$ with $d \ge 2$. Does the IDS become more regular as the number of frequencies increases? Currently, the results deteriorate with the number of frequencies. Related question: Is there a LDT of the form

$$\mathrm{meas}\big[x \in \mathbb{T} : |\log \|M_N(x,E) - NL_N(E)| > \delta N\big] < e^{-c_\delta N}$$

for shifts in higher dimensions?

• Determine the Hölder exponent for the IDS. More precisely, can one get Hölder $\frac{1}{2}-$ provided the potential has only non-degenerate critical points? This is known (Bourgain) for the almost Mathieu model and large disorders.

• Prove a version of Theorem 5 on $\mathbb{Z}^d$, $d \ge 3$.

• Positivity of the Lyapunov exponent for small disorders for the skew-shift.

• Positivity of the Lypunov exponent and Anderson localization for all positive disorders with $Tx = 2x$ (mod1), see [8].

References

[1] Aizenman, M., Molchanov, S. *Localization at large disorder and at extreme energies: an elementary derivation.* Comm. Math. Phys. 157 (1993), 245–278.

[2] Bourgain, J. *Estimates on Green's functions, localization and the quantum kicked rotor model.* Ann. of Math. (2) 156 (2002), no. 1, 249–294.

[3] Bourgain, J. *Green's function estimates for lattice Schrödinger operators and applications.* To appear in Annals of Mathematical Studies, Princeton.

[4] Bourgain, J., Goldstein, M. *On nonperturbative localization with quasiperiodic potential.* Annals of Math. 152 (2000), 835-879..

[5] Bourgain, J., Goldstein, M., Schlag, W. *Anderson localization for Schrödinger operators on $\mathbb{Z}$ with potentials given by the skew-shift.* Comm. Math. Phys. 220 (2001), no. 3, 583–621.

[6] Bourgain, J., Goldstein, M., Schlag, W. *Anderson localization for Schrödinger operators on $\mathbf{Z}^2$ with quasi-periodic potential.* Acta Math. 188 (2002), no. 1, 41–86.

[7] Bourgain, J., Jitomirskaya, S. *Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential.* Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays. J. Statist. Phys. 108 (2002), no. 5-6, 1203–1218.

[8] Bourgain, J., Schlag, W. *Anderson localization for Schrödinger operators on Z with strongly mixing potentials.* Comm. Math. Phys. 215 (2000), no. 1, 143–175.

[9] Carmona, R., Lacroix, J. *Spectral theory of random Schrödinger operators.* Birkhäuser, Boston 1990.

[10] Dinaburg, E. I., Sinai, Y. G. *The one-dimensional Schrödinger equation with quasiperiodic potential.* Funkt. Anal. i. Priloz. **9** (1975), 8–21.

[11] Figotin, A., Pastur, L. *Spectra of random and almost–periodic operators.* Grundlehren der mathematischen Wissenschaften 297, Springer 1992.

[12] Fröhlich, J., Spencer, T. *Absence of diffusion in the Anderson tight binding model for large disorder or low energy.* Commun. Math. Phys. **88** (1983), 151–189.

[13] Fröhlich, J., Spencer, T., Wittwer, P. *Localization for a class of one-dimensional quasi-periodic Schrödinger operators.* Commun. Math. Phys. **132** (1990), 5–25.

[14] Fürstenberg, H. *Noncommuting random products.* Trans. AMS 108 (1963), 377–428.

[15] Goldsheid, I. Ya., Molchanov, S. A., Pastur, L. A. *A pure point spectrum of the stochastic one-dimensional Schrödinger equation.* Funkt. Anal. Appl. **11** (1977), 1–10.

[16] Jitomirskaya, S. *Metal–insulator transition for the almost Mathieu operator.* Annals of Math. 150, no. 3 (1999).

[17] Goldstein, M., Schlag, W. *Hölder continuity of the integrated density of states for quasi-periodic Schröinger equations and averages of shifts of subharmonic functions.* Ann. of Math. (2) 154 (2001), no. 1, 155–203.

[18] Sinai, Y. G. *Anderson localization for one-dimensional difference Schrödinger operator with quasi-periodic potential.* J. Stat. Phys. **46** (1987), 861–909.