



An introduction to multiscale techniques in the theory of Anderson localization, Part I[☆]

Wilhelm Schlag

Department of Mathematics, Yale University, New Haven, CT 06511, USA

ARTICLE INFO

Article history:

Received 29 April 2021

Accepted 28 February 2022

Communicated by Yvan Martel

Keywords:

Disordered models in quantum theory

Anderson localization

Multiscale techniques

ABSTRACT

These lectures present some basic ideas and techniques in the spectral analysis of lattice Schrödinger operators with disordered potentials. In contrast to the classical Anderson tight binding model, the randomness is also allowed to possess only finitely many degrees of freedom. This refers to dynamically defined potentials, i.e., those given by evaluating a function along an orbit of some ergodic transformation (or of several commuting such transformations on higher-dimensional lattices). Classical localization theorems by Fröhlich–Spencer for large disorders are presented, both for random potentials in all dimensions, as well as even quasiperiodic ones on the line. After providing the needed background on subharmonic functions, we then discuss the Bourgain–Goldstein theorem on localization for quasiperiodic Schrödinger cocycles assuming positive Lyapunov exponents.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

In the 1950s Phil Anderson studied random operators of the form

$$H = \Delta_{\mathbb{Z}^d} + \lambda V$$

where $\Delta_{\mathbb{Z}^d}$ is the discrete Laplacian on the d -dimensional lattice and $V : \mathbb{Z}^d \rightarrow \mathbb{R}$ a random field with i.i.d. components, and a real parameter λ . His pioneering work suggested by physical arguments that for large λ , with probability 1, a typical realization of the random operator H exhibits exponentially decaying eigenfunctions which form a basis of $\ell^2(\mathbb{Z}^d)$. This is referred to as *Anderson localization* (AL). It is in stark contrast to periodic V for which the spectrum is absolutely continuous (a.c.) with a distorted Fourier basis of

[☆] The author was partially supported by US National Science Foundation grant DMS-1902691. He thanks Adam Black, Yakir Forman, and Tom VandenBoom for their assistance during the preparation of these notes, and Gong Chen and David Damanik for their comments on a preliminary version. The author is indebted to Massimiliano Berti at SISSA, Trieste, Italy, for inviting him to teach a mini-course at SISSA in March of 2020 on Anderson localization. Due to the COVID-19 pandemic these lectures did not materialize, but they still provided the impetus for the writing of these notes. Many thanks to two anonymous referees for their careful reading and their remarks, which greatly improved the presentation.

E-mail address: wilhelm.schlag@yale.edu.

Bloch–Floquet waves, see [40,42]. Furthermore, and most importantly, Anderson found a phase transition in dimensions three and higher, leading to the a.s. presence of a.c. spectrum for small λ . This famous *extended states problem* is still not understood.

On the other hand, a large mathematical literature now exists dealing with Anderson localization and its ramifications (density of states, Poisson behavior of eigenvalues). This introduction is not meant as a broad introduction to this field, for which we refer the reader to the recent textbook [2], as well as the more classical treatises [7,14,20] and the forthcoming texts [16,17]. Our focus here is with the body of techniques commonly referred to as *multiscale*. They are all based on some form of induction on scales, and are reminiscent of KAM arguments.

This approach is effective both in random models, as well as those with *deterministic* potentials, which refers to $V(n)$ being fixed by a finite number of parameters. For example, Harper’s model on $\mathbb{Z}$ is given by $V(n) = \cos(2\pi(n\omega + x))$ with irrational ω and $x \in \mathbb{R}/\mathbb{Z}$. The only stochastic parameter is this choice of x . The Harper operator, which is also known as almost Mathieu operator, as well as more general quasi-periodic operators, exhibit a rich and subtle spectral theory, see for example the survey [35].

Bourgain’s book [8] contains a wealth of material on a wide class of stochastic Schrödinger operators with deterministic potentials. An important basic assumption in that book is the analyticity of the generating function, i.e., if $V(n) = F(T^n x)$ for some ergodic transformation T on a torus, then F is assumed to be analytic or a trigonometric polynomial. The analyticity allows for the use of subharmonic functions. These are relevant for *large deviation theorems*, which in turn hinge on some Cartan type lower bound for subharmonic functions. This first part of the notes can be seen as a companion to Bourgain’s book [8] but only up to Chapter 12. The plan for the second part of this introduction is to focus on the matrix-valued Cartan theorem of [11], and the higher-dimensional theory as in [9], with applications. This will then hopefully serve to make Chapters 14 through 19 of [8] more accessible.

2. Polynomially bounded Fourier basis

In this section we establish the following widely known fact concerning the Fourier transform associated with a Schrödinger operator. It is a particular case of a more general theory, see the text [6] and the survey [44]. Results of this type go by the name of *Shnol theorem*. We follow the argument in [19]. Throughout, the discrete Laplacian on $\mathbb{Z}^d$ is defined as the sum over nearest neighbors, i.e.,

$$(\Delta f)(x) = \sum_{\pm} \sum_{j=1}^d f(x \pm e_j) \quad \forall x \in \mathbb{Z}^d \quad (2.1)$$

where e_j are the standard coordinate vectors. If $\mathcal{F} : \ell^2(\mathbb{Z}^d) \rightarrow L^2(\mathbb{T}^d)$ denotes the Fourier transform, then

$$(\mathcal{F} \circ \Delta \circ \mathcal{F}^{-1} f)(\theta) = m(\theta) f(\theta) = 2 \sum_{j=1}^d \cos(2\pi\theta_j) f(\theta) \quad (2.2)$$

and the spectrum satisfies $\text{spec}(\Delta) = [-2d, 2d]$. The Laplacian (2.1) differs from the more customary $-\tilde{\Delta} = \nabla^* \nabla$ where $(\nabla f)(x) = \{f(x + e_j) - f(x)\}_{j=1}^d$, by a diagonal term: $-\tilde{\Delta} = -\Delta + 2d$. For the original reference, see [47].

Theorem 2.1. *Consider $H = \Delta + V$ as a bounded operator on $\ell^2(\mathbb{Z}^d)$, with $V \in \ell^\infty(\mathbb{Z}^d)$ real-valued and acting by multiplication. Fix $\sigma > \frac{d}{2}$. Then for almost every $E \in \mathbb{R}$ with respect to the spectral measure¹ of H there exists $\psi : \mathbb{Z}^d \rightarrow \mathbb{R}$ not identically vanishing with $H\psi = E\psi$ and $|\psi(n)| \leq C(d, \sigma, E) \langle n \rangle^\sigma$ for all $n \in \mathbb{Z}^d$.*

¹ I.e., up to a set of measure 0 relative to the spectral measure of H .

Proof. Take any $z \in \mathbb{R} \setminus \Sigma$, where Σ is the spectrum of H (we take $z \in \mathbb{R}$ for simplicity). By the Combes–Thomas estimate the Green function $(H - z)^{-1}(x, y)$ has exponential decay: there exist positive constants C, β so that

$$|(H - z)^{-1}(x, y)| \leq C \exp(-\beta|x - y|) \quad \forall x, y \in \mathbb{Z}^d. \quad (2.3)$$

We follow the convention that an operator T has kernel $T(x, y)$. To see this, let $(M_{\beta,j}f)(n) = e^{\beta n_j} f(n)$ and compute

$$M_{\beta,j}^{-1} \circ (H - z) \circ M_{\beta,j} = H - z + S_{\beta,j}$$

with $\|S_{\beta,j}\|_{2 \rightarrow 2} \leq C|\beta|$ uniformly in $|\beta| \leq 1$. Hence,

$$M_{\beta,j}^{-1} \circ (H - z)^{-1} \circ M_{\beta,j} = (H - z)^{-1} \circ (I + S_{\beta,j}(H - z)^{-1})^{-1}$$

where the inverse on the right exists by a Neumann series as long as in the operator norm

$$\|S_{\beta,j}(H - z)^{-1}\| < 1$$

which holds if $|\beta|\text{dist}(z, \Sigma) \leq c$, some small constant. In particular,

$$|\langle \delta_y, (H - z)^{-1} \delta_x \rangle| = |(H - z)^{-1}(x, y)| \leq C e^{-\beta(x_j - y_j)}$$

Since the sign of β and the choice of j are arbitrary, (2.3) follows. Throughout, $\langle \cdot, \cdot \rangle$ denotes the complex inner product which is linear in the second variable.

Let $w_\sigma(x) := \langle x \rangle^{-\sigma}$ on $\mathbb{Z}^d$. Fixing any $\sigma > \frac{d}{2}$ so that $\langle x \rangle^{-\sigma} \in \ell^2(\mathbb{Z}^d)$, the Combes–Thomas bound (2.3) implies that $(H - z)^{-1}(x, y)w_\sigma(y) \in \ell^2(\mathbb{Z}^d \times \mathbb{Z}^d)$ whence

$$(H - z)^{-1} : \ell_\sigma^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)$$

is a Hilbert–Schmidt operator. Here $\ell_\sigma^2(\mathbb{Z}^d) := w_\sigma \ell^2(\mathbb{Z}^d)$. By the spectral theorem there exists a unitary $U : \ell^2(\mathbb{Z}^d) \rightarrow L^2(X, \mu)$ where μ is a σ -finite measure, and $\phi \in L^\infty(X)$ real-valued, with $UHf = \phi Uf$ for all $f \in \ell^2(\mathbb{Z}^d)$. The μ -essential range of ϕ equals Σ . The composition

$$T := U(H - z)^{-1} = (\phi - z)^{-1}U : \ell_\sigma^2(\mathbb{Z}^d) \rightarrow L^2(X, \mu)$$

is Hilbert–Schmidt, whence by the standard kernel representation of such operators, for every $n \in \mathbb{Z}^d$ there exists $K(\cdot, n) \in L^2(X, \mu)$ with

$$\int_X \sum_{n \in \mathbb{Z}^d} |K(x, n)|^2 w_\sigma^2(n) \mu(dx) < \infty$$

and

$$Tf(x) = \sum_{n \in \mathbb{Z}^d} K(x, n)f(n) \quad \forall f \in \ell_\sigma^2(\mathbb{Z}^d)$$

The series converges in $L^2(X, \mu)$. Define $\psi_x(n) := (\phi(x) - z)\overline{K(x, n)}$. Then for all $f \in \ell_\sigma^2(\mathbb{Z}^d)$, and μ -a.e. x ,

$$(Uf)(x) = (\phi(x) - z)(Tf)(x) = \sum_{n \in \mathbb{Z}^d} (\phi(x) - z)K(x, n)f(n) = \langle \psi_x, f \rangle_{\ell^2(\mathbb{Z}^d)} \quad (2.4)$$

By the preceding $\psi_x \in (\ell_\sigma^2(\mathbb{Z}^d))^* = \ell_{-\sigma}^2(\mathbb{Z}^d)$ for μ -a.e. x . Next, we claim that a.e. in x and in the point-wise sense on $\mathbb{Z}^d$

$$H\psi_x = \phi(x)\psi_x \quad (2.5)$$

as well as $\psi_x \neq 0$. Take f on the lattice with finite support. Then Hf has finite support and by (2.4)

$$\langle H\psi_x, f \rangle = \langle \psi_x, Hf \rangle = (UHF)(x) = \phi(x)(Uf)(x) = \langle \phi(x)\psi_x, f \rangle$$

all scalar products in $\ell^2(\mathbb{Z}^d)$, and for μ -a.e. x . It follows that $H\psi_x = \phi(x)\psi_x$ whence (2.5). Now suppose $\psi_x \equiv 0$ for all $x \in \mathcal{S} \subset X$, $\mu(\mathcal{S}) > 0$. Then for all $f \in \ell^2(\mathbb{Z}^d)$, this implies that $Uf = 0$ μ -a.e. on $\mathcal{S}$, and

$$0 = \langle Uf, \chi_{\mathcal{S}} \rangle_{L^2(\mu)} = \langle f, U^* \chi_{\mathcal{S}} \rangle_{\ell^2(\mathbb{Z}^d)}$$

But this means that $U^* \chi_{\mathcal{S}} = 0$ which contradicts $\|U^* \chi_{\mathcal{S}}\|_2^2 = \mu(\mathcal{S}) > 0$. To summarize, there is $\mathcal{N} \subset X$ with $\mu(\mathcal{N}) = 0$ so that for all $E \in \mathcal{G} := \phi(X \setminus \mathcal{N})$ the equation $H\psi = E\psi$ has a nonzero solution $\psi \in \ell_{-\sigma}^2(\mathbb{Z}^d)$. We claim that $\mathcal{E}(\mathbb{R} \setminus \mathcal{G}) = 0$ where $\mathcal{E}$ is the spectral resolution of H , i.e., a projection-valued Borel measure with $H = \int \lambda \mathcal{E}(d\lambda)$. But for any Borel set $B \subset \mathbb{R}$, $U\mathcal{E}(B)U^*g = \chi_{\phi^{-1}(B)}g$ on all $g \in L^2(X, \mu)$. Since $\phi^{-1}(\mathbb{R} \setminus \mathcal{G}) \subset \mathcal{N}$ is a μ -nullset, we conclude that $U\mathcal{E}(\mathbb{R} \setminus \mathcal{G})U^* = 0$ whence $\mathcal{E}(\mathbb{R} \setminus \mathcal{G}) = 0$. So $H\psi = E\psi$ has nonzero solution $\psi \in \ell_{-\sigma}^2(\mathbb{Z}^d)$ spectrally a.e. $\square$

This proof applies much more generally and the discrete Laplacian is used only sparingly. For example, it can be replaced with a self-adjoint Töplitz operator with exponential off-diagonal decay.

3. The Anderson model and localization

Let

$$H_\omega = \Delta + V_\omega \quad (3.1)$$

on $\ell^2(\mathbb{Z}^d)$ where V_ω is a diagonal operator given by i.i.d. random variables at each lattice site $n \in \mathbb{Z}^d$. The single site distribution refers to the law of $V_\omega(0)$ which we assume to be a.s. bounded. Then H_ω is a.s. a bounded operator. The notation is based on an underlying probability space $(\Omega, \mathbb{P})$, with $\omega \in \Omega$. If $d = 1$ we may consider a more general model with a potential generated by any ergodic dynamical system. Thus, let $T : X \rightarrow X$ be measure preserving, invertible, and ergodic on X relative to the probability measure ν . Then set $V_x(n) := f(T^n x)$ where $f : X \rightarrow \mathbb{R}$ is measurable and ν -essentially bounded, and define $H_x := \Delta + V_x$ in $\ell^2(\mathbb{Z})$. The one-dimensional random Anderson model is included in this formalism. Indeed, let T be the Bernoulli shift and f the projection on the 0-coordinate with the ergodic measure being the infinite product measure generated by the single-site distribution. For the general 1-dimensional model we shall now demonstrate that the spectrum of H_x is deterministic by ergodicity.

Proposition 3.1. *There exists a compact set $\Sigma \subset \mathbb{R}$ so that $\text{spec}(H_x) = \Sigma$ for ν -almost every $x \in X$.*

Proof. One has the conjugation

$$H_{Tx} = U^{-1} \circ H_x \circ U \quad (3.2)$$

with the right-shift $U : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$ and so also $\mathcal{E}_{Tx}(B) = U^{-1} \circ \mathcal{E}_x(B) \circ U$ with the spectral resolution $\mathcal{E}_x$ of H_x and any Borel set $B \subset \mathbb{R}$. Recall that $H_x = \int \lambda \mathcal{E}_x(d\lambda)$ in a suitable sense.

$$E(B) := \{x \in X \mid \text{rank}(\mathcal{E}_x(B)) > 0\}$$

is an invariant set, i.e., $E(B) = T^{-1}(E(B))$ whence $\nu(E(B)) = 0$ or $\nu(E(B)) = 1$. Define

$$\Sigma = \mathbb{R} \setminus \bigcup_{a < b} \{(a, b) \mid \nu(E((a, b))) = 0\}$$

where the union is over rational a, b . If any interval $I \subset \mathbb{R}$ intersects Σ , then $\text{rank}(\mathcal{E}_x(I)) > 0$ for ν -a.e. $x \in X$. Hence also $I \cap \text{spec}(H_x) \neq \emptyset$ for a.e. x . $\square$

A finer description of the set Σ can be obtained by a similar argument. The subscripts ac, sc, pp stand for, respectively, absolutely continuous, singular continuous, and pure point.

Proposition 3.2. *There exist compact subsets Σ_{ac} , Σ_{sc} , and Σ_{pp} of Σ such that $\Sigma = \Sigma_{ac} \cup \Sigma_{sc} \cup \Sigma_{pp}$ (not necessarily disjoint) such that for any Borel set B with $B \cap \Sigma_{ac} \neq \emptyset$ the following holds: for a.e. $x \in X$ there exists $f \in \ell^2(\mathbb{Z})$ so that $\mu(A) := \langle \mathcal{E}_x(A \cap B)f, f \rangle$ defined on Borel sets A is an absolutely continuous probability measure. Analogous statements hold for the singular continuous, and pure point (atomic) parts.*

Proof. We define, with the union being over rationals,

$$\Sigma_{ac} = \mathbb{R} \setminus \bigcup_{a < b} \{(a, b) \mid \forall f \in \ell^2(\mathbb{Z}), A \mapsto \langle \mathcal{E}_x((a, b) \cap A)f, f \rangle \\ \nu - a.s. \text{ has no absolutely continuous component}\}$$

where the latter property refers to the Lebesgue decomposition. We adopt the convention that the 0 measure has no absolutely continuous component (as well as no singular component). By ergodicity and the conjugacy of H_x and H_{Tx} , respectively, by the shift, the set

$$Y(a, b) := \{x \in X \mid \exists f \in \ell^2(\mathbb{Z}), A \mapsto \langle \mathcal{E}_x((a, b) \cap A)f, f \rangle \\ \text{has an absolutely continuous component}\}$$

is T -invariant and thus $\nu(Y(a, b)) = 0$ or $\nu(Y(a, b)) = 1$. Hence

$$\Sigma_{ac} = \mathbb{R} \setminus \bigcup_{a < b} \{(a, b) \mid a, b \in \mathbb{Q}, \nu(Y(a, b)) = 0\}$$

Now suppose $B \cap \Sigma_{ac} \neq \emptyset$. Without loss of generality, $B \subset \Sigma_{ac}$. If $B \cap (a, b) \neq \emptyset$ with $a, b \in \mathbb{Q}$, then $\nu(Y(a, b)) = 1$. Thus ν -a.s., $A \mapsto \langle \mathcal{E}_x((a, b) \cap A)f, f \rangle$ is absolutely continuous for some f . We used here that we may pass from the existence of an absolutely continuous component to purely absolutely continuous by projecting f on the a.c. subspace of H_x . The claim of having a probability measure is obtained by normalization. The proofs for the singular parts is identical. $\square$

These arguments make no use of the Laplacian and therefore apply to the diagonal operator given by multiplication by the potential V_x . In that case the eigenvalues are $\{V_x(n) = f(T^n x) \mid n \in \mathbb{Z}\}$ and the closure of this set is deterministic and equals Σ_{pp} . Moreover, $\Sigma_{ac} = \Sigma_{sc} = \emptyset$.

Propositions 3.1 and **3.2** apply as stated to the random model H_ω from above, as the reader is invited to explore. In fact, on $\ell^2(\mathbb{Z}^d)$ we may consider d measure preserving, invertible, commuting transformations $T_j : X \rightarrow X$ with the following ergodicity property: if $A \subset X$ is invariant under *all* T_j , then $\nu(A) = 0$ or $\nu(A) = 1$. Then the previous two propositions apply to the operator $H_x := \Delta_{\mathbb{Z}^d} + V_x$ with $V_x(\bar{n}) = f(T_1^{n_1} \circ T_2^{n_2} \circ \dots \circ T_d^{n_d} x)$ for any $\bar{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ with essentially the same proofs. See [20,39] for a systematic development of the spectral theory of ergodic families of operators.

For the random model, which is the original Anderson model, we can now explicitly compute the almost sure spectrum Σ in **Proposition 3.1**. Recall that we are assuming bounded support of the single site distribution.

Proposition 3.3. For H_ω as defined in (3.1) satisfies

$$\Sigma = [-2d, 2d] + K$$

where K is the essential support of the single site distribution $V_\omega(0)$.

Proof. By definition, $K = \mathbb{R} \setminus \bigcup\{I \mid \mu(I) = 0\}$ where I is an interval with rational endpoints. If $\lambda_0 \in [-2d, 2d]$, then by (2.2) there exists $\alpha \in \mathbb{T}^d$ with $m(\alpha) = \lambda_0$. Thus, $\Delta e_\alpha = \lambda_0 e_\alpha$ where $e_\alpha(n) = e^{2\pi i \alpha \cdot n}$ for all $n \in \mathbb{Z}^d$. The following holds almost surely: given $L \geq 1$, $\varepsilon > 0$, and $\lambda_1 \in K$, there exists a cube $A \subset \mathbb{Z}^d$ of side length L such that $\|V - \lambda_1\|_{\ell^\infty(A)} \leq \varepsilon$. Then with $\lambda = \lambda_0 + \lambda_1$,

$$(H - \lambda)\chi_A e_\alpha = (V - \lambda_1)\chi_A e_\alpha + g$$

with $\|g\|_2^2 \lesssim |\partial A| \lesssim L^{d-1}$. Here ∂A is defined as those $x \in A$ which have a nearest neighbor in $\mathbb{Z}^d \setminus A$, and $|\cdot|$ is the cardinality (or volume). Hence, with the normalized function $\varphi = \chi_A e_\alpha |A|^{-\frac{1}{2}}$

$$\|(H - \lambda)\varphi\|_2 \lesssim \varepsilon + L^{-\frac{1}{2}}$$

which implies that almost surely,

$$\sup_{\varepsilon > 0} \|(H - \lambda - i\varepsilon)^{-1}\| = \infty$$

and thus $\lambda \in \text{spec}(H)$. This shows that $[-2d, 2d] + K \subset \Sigma$.

Conversely, suppose $\lambda \in \mathbb{R} \setminus ([-2d, 2d] + K)$. By compactness of the sum set there exists $\delta > 0$ so that almost surely

$$\inf_{n \in \mathbb{Z}^d} |V_\omega(n) - \lambda| \geq 2d + \delta$$

Thus, a.s. the resolvent

$$(H - \lambda)^{-1} = (I + (V_\omega - \lambda)^{-1} \Delta)^{-1} (V_\omega - \lambda)^{-1}$$

exists as a bounded operator on $\ell^2(\mathbb{Z}^d)$. $\square$

For any cube $A \subset \mathbb{Z}^d$ we denote by P_A the projection onto all states, i.e., $f \in \ell^2(\mathbb{Z}^d)$ supported in A . Thus, $P_A f = \mathbb{1}_A f$ for any $f \in \ell^2(\mathbb{Z}^d)$. By $H_A := P_A H P_A$ we denote the restriction of H as in (3.1) to the cube A with Dirichlet boundary conditions. Note that the randomness of H is understood and not indicated in the notation, say by an index ω .

It is natural to ask about the probability that any given number $E \in \mathbb{R}$ comes close to the spectrum of H_A . In other words, for any $\rho > 0$ what is

$$\mathbb{P}(\{\text{dist}(E, \text{spec}(H_A)) < \rho\}) = \mathbb{P}(\{\|(H_A - E)^{-1}\| > \rho^{-1}\})? \quad (3.3)$$

The diagonal operator given by the random potential V alone satisfies

$$\begin{aligned} \mathbb{P}(\{\text{dist}(E, \text{spec}(P_A V P_A)) < \rho\}) &\leq \mathbb{E} \#\{n \in A \mid V(n) \in (E - \rho, E + \rho)\} \\ &\leq |A| \mu((E - \rho, E + \rho)) \leq 2\rho |A| \left\| \frac{d\mu}{dx} \right\|_\infty \end{aligned} \quad (3.4)$$

where μ is the law of $V(0)$. A classical fact concerning the random Schrödinger operator is that (3.3) permits essentially the same bound as (3.4). This is known as *Wegner's estimate*, see [52].

Proposition 3.4. Assume the single site distribution of the random operator (3.1) satisfies $\mu' = \frac{d\mu}{dy} \in L^\infty(\mathbb{R})$. Then for all $E \in \mathbb{R}$,

$$\mathbb{P}(\{\text{dist}(E, \text{spec}(H_\Lambda)) < \rho\}) \leq 4\rho \|\mu'\|_\infty |\Lambda| \quad (3.5)$$

for all cubes $\Lambda \subset \mathbb{Z}^d$ and $\rho > 0$.

Proof. We will present two proofs. For the first we follow Wegner's original argument [52]. Denote by $N_\Lambda(x)$ the integrated density of states for the random operator H_Λ . To wit, if $E_\Lambda^1 \leq E_\Lambda^2 \leq \dots \leq E_\Lambda^m$, $m = |\Lambda|$, denote the eigenvalues of H_Λ with multiplicity, then

$$N_\Lambda(x) = \# \{1 \leq j \leq m \mid E_\Lambda^j \leq x\}$$

Let $\varphi \geq 0$ be a smooth bump function on $\mathbb{R}$ supported in $[-1, 1]$, and set $\varphi_\rho(x) = \rho^{-1}\varphi(x/\rho)$. Normalize so that $\int_{\mathbb{R}} \varphi(x) dx = 1$. Then with $F_{\Lambda, \rho} = N_\Lambda * \varphi_\rho$ one has

$$N_\Lambda(E + \rho) - N_\Lambda(E - \rho) \leq F_{\Lambda, \rho}(E + 2\rho) - F_{\Lambda, \rho}(E - 2\rho) = \int_{E-2\rho}^{E+2\rho} F'_{\Lambda, \rho}(x) dx$$

Since N_Λ is a monotone increasing step-function, we have $F'_{\Lambda, \rho} \geq 0$. We may interpret $N_\Lambda(x) = N_\Lambda(V_\Lambda, x)$, indicating the dependence of N_Λ on all the potential values in Λ . Then $N_\Lambda(x + h) = N_\Lambda(V_\Lambda - h, x)$ whence

$$N'_\Lambda(x) = - \sum_{j \in \Lambda} \frac{\partial N_\Lambda}{\partial v_j}(x), \quad F'_{\Lambda, \rho}(x) = - \sum_{j \in \Lambda} \frac{\partial F_{\Lambda, \rho}}{\partial v_j}(x)$$

as identities between distributional derivatives, respectively smooth functions. Note that $\frac{\partial F_{\Lambda, \rho}}{\partial v_j} \leq 0$ for each j . Indeed, N_Λ is decreasing in each v_j separately by the min-max characterization of the eigenvalues of a symmetric matrix. More generally, min-max shows that if $A \geq B$ for any two symmetric matrices, then the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots$ of A dominate those of B , denoted by $\mu_1 \geq \mu_2 \geq \dots$ which means that $\lambda_k \geq \mu_k$ for all k .

Thus, with $[-L, L]$ containing the support of μ ,

$$\begin{aligned} \mathbb{P}(\{\text{dist}(E, \text{spec}(H_\Lambda)) < \rho\}) &\leq - \int_{E-2\rho}^{E+2\rho} \sum_{j \in \Lambda} \mathbb{E} \frac{\partial F_{\Lambda, \rho}}{\partial v_j}(x) dx \\ &\leq \int_{E-2\rho}^{E+2\rho} \sum_{j \in \Lambda} \mathbb{E}'_j \int_{-L}^L - \frac{\partial F_{\Lambda, \rho}}{\partial v_j}(x) \mu'(v_j) dv_j dx \end{aligned} \quad (3.6)$$

where $\mathbb{E}'_j$ refers to the expectation relative to $\{v_k\}_{k \in \Lambda \setminus \{j\}}$. Further, using the positivity of the integrand,

$$\begin{aligned} \int_{-L}^L - \frac{\partial F_{\Lambda, \rho}}{\partial v_j}(x) \mu'(v_j) dv_j &\leq \|\mu'\|_\infty \int_{-L}^L - \frac{\partial F_{\Lambda, \rho}}{\partial v_j}(x) dv_j \\ &= \|\mu'\|_\infty (F_{\Lambda, \rho}(v_j = -L) - F_{\Lambda, \rho}(v_j = L)) \\ &= \|\mu'\|_\infty \int_{\mathbb{R}} (N_\Lambda(v_j = -L, x) - N_\Lambda(v_j = L, x)) \varphi_\rho(x) dx \leq \|\mu'\|_\infty \end{aligned} \quad (3.7)$$

For the final estimate we use that passing from $v_j = -L$ to $v_j = L$ in H_Λ constitutes a rank-1 perturbation which implies by min-max that the eigenvalues of $H_\Lambda(v_j = -L)$ and $H_\Lambda(v_j = L)$ interlace. This in turn guarantees that

$$N_\Lambda(v_j = -L, x) - N_\Lambda(v_j = L, x) \leq 1 \quad \forall x \in \mathbb{R}$$

and thus (3.7). Combining (3.6) with (3.7) implies (3.5).

For the second proof, we follow [45] and estimate

$$\begin{aligned} \mathbb{P}(\{\text{dist}(E, \text{spec}(H_\Lambda)) < \rho\}) &\leq \mathbb{E} \text{trace} \mathbb{1}_{[E-\rho, E+\rho]}(H_\Lambda) \leq 2\rho \mathbb{E} \text{trace} \text{Im} (H_\Lambda - (E + i\rho))^{-1} \\ &\leq 2\rho \mathbb{E} \sum_{n \in \Lambda} \text{Im} \langle (H_\Lambda - (E + i\rho))^{-1} \delta_n, \delta_n \rangle \end{aligned} \quad (3.8)$$

where we used that

$$\mathbb{1}_{[E-\rho, E+\rho]}(x) \leq \frac{2\rho^2}{\rho^2 + (x - E)^2} = \text{Im} \frac{2\rho}{x - (E + i\rho)}$$

Next, we establish a fundamental relation on the resolvents of rank-1 perturbations. Let A be any self-adjoint operator on a Hilbert space and φ a unit vector, λ a real scalar. From the resolvent identity, for any complex z with $\text{Im } z > 0$,

$$\begin{aligned} (A + \lambda\varphi \otimes \varphi - z)^{-1} - (A - z)^{-1} &= -\lambda(A - z)^{-1}(\varphi \otimes \varphi)(A + \lambda\varphi \otimes \varphi - z)^{-1} \\ \langle (A + \lambda\varphi \otimes \varphi - z)^{-1} \varphi, \varphi \rangle - \langle (A - z)^{-1} \varphi, \varphi \rangle &= -\lambda \langle (A + \lambda\varphi \otimes \varphi - z)^{-1} \varphi, \varphi \rangle \langle (A - z)^{-1} \varphi, \varphi \rangle \\ \langle (A + \lambda\varphi \otimes \varphi - z)^{-1} \varphi, \varphi \rangle &= [\lambda + \langle (A - z)^{-1} \varphi, \varphi \rangle^{-1}]^{-1} \end{aligned}$$

Note that $\text{Im} \langle (A - z)^{-1} \varphi, \varphi \rangle^{-1} \neq 0$ by $\text{Im } z \neq 0$. Applying this to

$$H_\Lambda = H_\Lambda^{(n)} + V_\omega(n) \delta_n \otimes \delta_n, \quad n \in \Lambda$$

where $H_\Lambda^{(n)}$ is the operator with the potential at lattice site n set equal to 0, yields

$$\langle (H_\Lambda - (E + i\rho))^{-1} \delta_n, \delta_n \rangle = [V_\omega(n) + \langle (H_\Lambda^{(n)} - (E + i\rho))^{-1} \delta_n, \delta_n \rangle^{-1}]^{-1}$$

Writing

$$\langle (H_\Lambda^{(n)} - (E + i\rho))^{-1} \delta_n, \delta_n \rangle^{-1} = -t - is, \quad s > 0$$

the random variables t, s only depend on the random lattice sites in $\Lambda \setminus \{n\}$. Consequently, the inner product in the final expression of (3.8) is bounded by

$$\mathbb{E} \text{Im} \langle (H_\Lambda - (E + i\rho))^{-1} \delta_n, \delta_n \rangle \leq \mathbb{E}'_n \int \frac{s \mu(dx)}{(x - t)^2 + s^2} \leq \pi \|\mu'\|_\infty$$

which in combination with (3.8) yields

$$\mathbb{P}(\{\text{dist}(E, \text{spec}(H_\Lambda)) < \rho\}) \leq 2\pi\rho \|\mu'\|_\infty |\Lambda|$$

This is slightly worse than the previous proof but the precise constant is irrelevant. $\square$

The assumption of bounded density μ' can be relaxed, but some amount of regularity of the single-site distribution is needed. Indeed, the *mobility of the eigenvalues* under the randomness expressed by Wegner's estimate is reduced to the mobility of the potential at each site. The heuristic notion of “mobility” refers to the movement of the eigenvalues as a result of the movement of the potential. Both arguments presented above hinge on this step. See, however, an alternative approach by Stollmann [49].

Anderson localization refers to the following statement.

Theorem 3.5. *Let $H = \varepsilon^2 \Delta_{\mathbb{Z}^d} + V_\omega$ where V_ω is random i.i.d. potential with single site distribution μ of compact support and of bounded density with $\|\mu'\|_\infty \leq 1$. Then there exists $\varepsilon_0 = \varepsilon_0(d) > 0$ so that for all $0 < \varepsilon \leq \varepsilon_0$, almost surely $\ell^2(\mathbb{Z}^d)$ has an orthonormal basis of exponentially decaying eigenfunctions of the random operator H . In particular, the spectrum is a.s. pure point and thus $\Sigma_{sc} = \Sigma_{ac} = \emptyset$.*

In stark contrast to this result, periodic potentials exhibit a Fourier basis of Bloch–Floquet solutions with $\Sigma_{pp} = \Sigma_{sc} = \emptyset$. Thus their spectral measures are absolutely continuous. This, as well as $V_\omega = \text{const.}$ shows that [Theorem 3.5](#) requires the removal of a zero probability event. A wide open problem is to prove $\Sigma_{ac} \neq \emptyset$ for large ε in dimensions $d \geq 3$. This is known as Anderson’s *extended states* conjecture.

There are two main techniques known to prove [Theorem 3.5](#): Fröhlich–Spencer [22] *multiscale analysis* on the one hand, and the Aizenman–Molchanov [1] *fractional moment method* on the other hand. We will sketch the former and refer to [33] for an introduction to the latter. A streamlined rendition of the induction-on-scales method of [22] can be found in [51], which also does not require the use of the Simon–Wolff criterion [45], as earlier multi-scale proofs of [Theorem 3.5](#) had done. Germinet and Klein have obtained significant refinements of the multi-scale argument in a series of papers, see for example [25].

Returning briefly to the Wegner estimate, we remark that the physically important example of Bernoulli potentials taking discrete values completely falls outside the range of [Proposition 3.4](#). See [18] for a recent advance on this case in two dimensions and on localization for Anderson Bernoulli. The mobility of the eigenvalues of H_Λ if $V = \pm 1$ derives from the interaction between eigenfunctions and is more delicate. On the other hand, localization in the one-dimensional Bernoulli model is a classical result by Carmona, Klein and Martinelli [13]. While these authors rely on the original multi-scale methods of Fröhlich and Spencer, this is avoided in the recent papers [12,30], and [37]. The arguments there use the large deviation theorems and the methods of Bourgain, Goldstein [10], see the final section of these notes.

Before getting in to the details, some basic ideas and motivation. Suppose H has an ℓ^2 -complete sequence of exponentially decaying normalized eigenfunctions $\{\phi_j\}_{j \in \mathbb{Z}}$ with eigenvalues E_j , both random. Restrict H to a large cube Λ and write (heuristically)

$$(H_\Lambda - (E + i\varepsilon))^{-1} \approx \sum_j \frac{\phi_j \otimes \phi_j}{E_j - (E + i\varepsilon)}$$

where the sum extends over all eigenfunctions “supported” in the box Λ . It should be intuitively clear what this means. Then $|(\phi_j \otimes \phi_j)(x, y)| \lesssim \exp(-\gamma|x - y|)$ with $\gamma > 0$ for those j , for which either x or y are in the support of ϕ_j . All the others make much smaller contributions which we can essentially ignore. In conclusion, if

$$\|(H_\Lambda - (E + i0))^{-1}\| \leq K \quad \text{then} \quad |(H_\Lambda - (E + i0))^{-1}(x, y)| \leq K \exp(-\gamma|x - y|) \quad \forall x, y \in \Lambda$$

The condition here is precisely what Wegner’s estimate controls, and a cube which exhibits both the separation from the spectrum and the exponential off-diagonal decay will be called *regular* for energy E . A substantial effort below is to show that cubes are regular for a given energy with high probability. However, this is insufficient to prove localization and one needs to consider two disjoint cubes and understand the probability that they are both singular for *any energy*. The essential feature of this idea is to control the probability of an event uniformly in all energies, rather than for a fixed energy. The latter can never imply an a.s. statement about the spectrum since we cannot take the union of a bad event over an uncountable family of energies. More importantly, excluding the event that two boxes are in resonance simultaneously (which refers that they are both singular at the same E) will precisely allow us to show that a.s. tunneling cannot occur over long distances leading to exponentially localized eigenfunctions.

We shall now prove [Theorem 3.5](#) by induction on scales. We will need to allow rectangles as regions of finite volume rather than just cubes. Thus, define a *box* centered at x of scale L to be any rectangle of the form

$$A_L(x) = \{y \in \mathbb{Z}^d \mid -m_j \leq y_j - x_j \leq M_j \quad \forall 1 \leq j \leq d\} \quad (3.9)$$

where $m_j \geq 0$, $M_j \geq 0$ and $\max(m_j, M_j) = L$ for each j . If $m_j = M_j = L$ for each j , then we have standard cube which we denote by $Q_L(x)$. These rectangles arise as intersections of cubes $Q_L(x) \cap Q_L(y)$ if $x \in Q_L(y)$

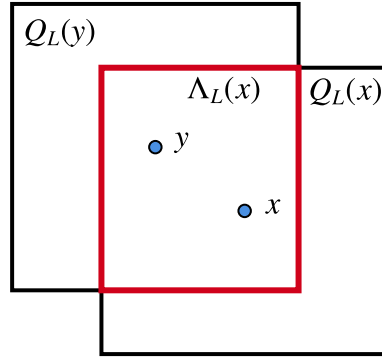


Fig. 1. A box arising as intersection of cubes.

with $\tilde{L} \geq L$, see Fig. 1. Wegner's estimate applies unchanged to boxes. Note that for a given integer $L \geq 1$ and $x \in \mathbb{Z}^d$ there are $B(L) = (2L+1)^d$ boxes $\Lambda_L(x)$. The following *deterministic lemma* allows us to bound the Green function at an initial scale which will be specified later.

Lemma 3.6. Suppose $4d\varepsilon \leq \delta$ and $0 < \varepsilon \leq \frac{1}{2}$. Let Λ be any box as in (3.9) and assume

$$\text{dist}(\text{spec}(H_\Lambda), E) \geq \delta.$$

Then

$$|G_\Lambda(E)(x, y)| \leq 4\delta^{-1}\varepsilon^{|x-y|} \quad \forall x, y \in \Lambda. \quad (3.10)$$

Here $|x| = \max_j |x_j|$ and $G_\Lambda(E) = (H_\Lambda - E)^{-1}$ is the Green function on Λ with energy E .

Proof. By min-max, $|V(x) - E| \geq \delta - 2d\varepsilon^2 \geq \delta/2$ for all $x \in \Lambda$. Then

$$\begin{aligned} G_\Lambda(E)(x, y) &= (I + \varepsilon^2(V_\Lambda - E)^{-1}\Delta)^{-1}(V_\Lambda - E)^{-1}(x, y) \\ &= \sum_{\ell=0}^{\infty} (-1)^\ell \varepsilon^{2\ell} [(V_\Lambda - E)^{-1}\Delta]^\ell(x, y)(V_\Lambda - E)^{-1}(x) \end{aligned}$$

Recall that $(V_\Lambda - E)^{-1}$ is a diagonal operator. Using that $\|(V_\Lambda - E)^{-1}\Delta\| \leq 2\delta^{-1}\|\Delta\| \leq 4d\delta^{-1}$, we have $|\varepsilon^{2\ell}[(V_\Lambda - E)^{-1}\Delta]^\ell(x, y)| \leq (4d\varepsilon^2\delta^{-1})^\ell \leq \varepsilon^\ell$ where it suffices to consider $\ell \geq |x-y|_1 \geq |x-y|$ with $|x|_1 = \sum_{j=1}^d |x_j|$ (otherwise this term vanishes). Summing up the geometric series using that $\varepsilon \leq \frac{1}{2}$ proves (3.10). $\square$

In terms of random operators one has (3.10) with high probability.

Corollary 3.7. Suppose $4d\varepsilon \leq \delta$ and $0 < \varepsilon \leq \frac{1}{2}$. Then (3.10) holds up to probability at most $4\delta|\Lambda|$.

Proof. Apply Wegner. $\square$

The following lemma demonstrates how to obtain exponential decay of the Green function at a large scale box if all boxes contained in it of a much smaller scale have this property, *with possibly one exception*. The latter is needed in order to be able to *square the probabilities* of a having a bad small box inside a bigger one as we pass to the next scale. We will use a resolvent expansion, obtained by iterating the resolvent identity:

let $\Lambda' \subset \Lambda$ be boxes, and let $A = H_\Lambda$, $B = H_{\Lambda'} \oplus H_{\Lambda \setminus \Lambda'}$ viewed as operators on $\ell^2(\Lambda)$. Then, observing that $A - B = \varepsilon^2 \Gamma_{\Lambda, \Lambda'}$,

$$\begin{aligned} (H_\Lambda - E)^{-1} &= (H_{\Lambda'} \oplus H_{\Lambda \setminus \Lambda'} - E)^{-1} - \varepsilon^2 (H_\Lambda - E)^{-1} \Gamma_{\Lambda, \Lambda'} (H_{\Lambda'} \oplus H_{\Lambda \setminus \Lambda'} - E)^{-1} \\ G_\Lambda(E)(x, y) &= G_{\Lambda'}(E)(x, y) \mathbb{1}_{[y \in \Lambda']} - \varepsilon^2 \sum_{(w', w) \in \partial \Lambda'} G_\Lambda(E)(w, y) G_{\Lambda'}(E)(x, w') \end{aligned} \quad (3.11)$$

for all $x \in \Lambda'$ and $y \in \Lambda$. Here $\partial \Lambda' = \{(w', w) \mid w' \in \Lambda', w \in \Lambda \setminus \Lambda', |w - w'| = 1\}$ is the *relative boundary* of Λ' inside of Λ , and $\Gamma_{\Lambda, \Lambda'} = \mathbb{1}_{\partial \Lambda'}$.

Lemma 3.8. *Let Λ be a box at scale $L_1 \geq 100L_0$ and assume $\text{dist}(\text{spec}(H_\Lambda), E) \geq \delta_1$ with $0 < \delta_1 \leq 1$. Let $\Lambda'_* \subset \Lambda$ be some box at scale $L_0 \geq 1$ and assume that all boxes $\Lambda' \subset \Lambda \setminus \Lambda'_*$ at scale L_0 satisfy*

$$|G_{\Lambda'}(E)(x, y)| \leq 4\delta_0^{-1} \varepsilon^{\gamma_0 |x-y|} \quad \forall x, y \in \Lambda', |x - y| \geq L_0/2 \quad (3.12)$$

Suppose

$$8d(2L_0 + 1)^{d-1} \varepsilon^{2+\gamma_0 L_0} \leq \delta_0 \leq 1 \quad (3.13)$$

Then

$$|G_\Lambda(E)(x, y)| \leq \delta_1^{-1} \varepsilon^{\gamma_1 |x-y|} \quad \forall x, y \in \Lambda, |x - y| \geq L_1/2 \quad (3.14)$$

provided

$$\left[\gamma_0 - \gamma_1 \left(1 - \frac{1}{L_0 + 1} - \frac{8L_0}{L_1} \right)^{-1} + \frac{2}{L_0} \right] \log \frac{1}{\varepsilon} \geq L_0^{-1} \log(8d(2L_0 + 1)^{d-1} \delta_0^{-1}) \quad (3.15)$$

Proof. Pick $x, y \in \Lambda$ with $|x - y| \geq L_1/2$ and set $\Lambda_x = Q_{L_0}(x) \cap \Lambda$. If $\Lambda_x \cap \Lambda'_* \neq \emptyset$, then we do not expand around x and instead expand around y since $L_1 \geq 100L_0$ implies that $\Lambda_y \cap \Lambda'_* = \emptyset$.

Iterating (3.11) leads to an expression of the form, with $w_0 = x$, $\tilde{w}_0 = y$,

$$\begin{aligned} G_\Lambda(E)(x, y) &= (-\varepsilon^2)^{s+t} \sum_{(w'_1, w_1) \in \partial \Lambda'(x)} \sum_{(w'_2, w_2) \in \partial \Lambda'(w_1)} \dots \sum_{(w'_s, w_s) \in \partial \Lambda'(w_{s-1})} \prod_{j=1}^s G_{\Lambda'(w_{j-1})}(w_{j-1}, w'_j) \\ &\quad \sum_{(\tilde{w}'_1, \tilde{w}_1) \in \partial \Lambda'(y)} \sum_{(\tilde{w}'_2, \tilde{w}_2) \in \partial \Lambda'(\tilde{w}_1)} \dots \sum_{(\tilde{w}'_t, \tilde{w}_t) \in \partial \Lambda'(\tilde{w}_{t-1})} G_\Lambda(w_s, \tilde{w}_t) \prod_{k=1}^t G_{\Lambda'(\tilde{w}_{k-1})}(\tilde{w}_{k-1}, \tilde{w}'_k) \end{aligned} \quad (3.16)$$

with all Green function on the right-hand side being at energy E . Here $s \geq 0$ and $t \geq 0$ are the maximal number of steps we can take from x , respectively, y with any boxes of size L_0 centered at points distance 1 from the boundary of a previous box, before they might intersect Λ'_* . All boxes here are of the form $Q_{L_0}(w_j) \cap \Lambda = \Lambda'(w_j)$. In particular, if $(y', y) \in \partial \Lambda'(w_j)$, then $|y' - w_j| = L_0$. We claim that $s + t \geq 1$ satisfies

$$|x - y| < (t + s)(L_0 + 1) + 4L_0 + 1$$

Indeed, if $|x - y| \geq (t + s)(L_0 + 1) + 4L_0 + 1$, then

$$|x - y| + 1 - [(t + s)(L_0 + 1) + 2L_0 + 1] \geq 2L_0 + 1$$

which implies that we could go either one more step in the x , resp. y , expansion without intersecting Λ'_* . Thus,

$$\xi - 1 < s + t \leq \xi, \quad \xi = \frac{|x - y| - 3L_0}{L_0 + 1} \quad (3.17)$$

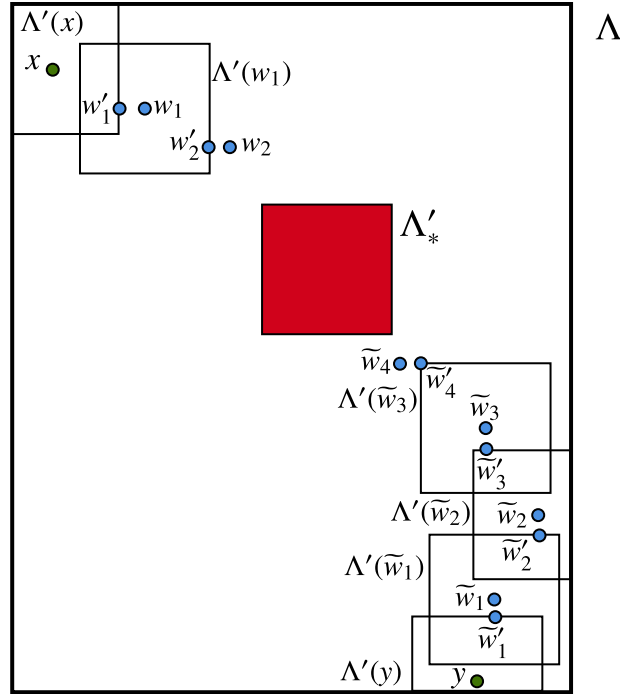


Fig. 2. One term in the expansion (3.16) with $s = 2, t = 4$.

To estimate (3.16), use $|G_\Lambda(w_s, \tilde{w}_t)| \leq \|G_\Lambda\| \leq \delta_1^{-1}$ and $|G_{\Lambda'(w_{j-1})}(w_{j-1}, w'_j)| \leq 4\delta_0^{-1}\varepsilon^{\gamma_0 L_0}$ and the same for all of the Green functions over the smaller boxes. The number of pairs in the boundary satisfy $|\partial\Lambda'| \leq 2d(2L_0 + 1)^{d-1}$ whence

$$\begin{aligned} |G_\Lambda(E)(x, y)| &\leq \delta_1^{-1} (8d\varepsilon^2(2L_0 + 1)^{d-1}\delta_0^{-1}\varepsilon^{\gamma_0 L_0})^{s+t} \\ &\leq \delta_1^{-1} (8d\varepsilon^2(2L_0 + 1)^{d-1}\delta_0^{-1}\varepsilon^{\gamma_0 L_0})^{\frac{|x-y|}{L_0+1}-4} \end{aligned} \quad (3.18)$$

using that the parenthesis is a number in $(0, 1]$. Note that $\frac{|x-y|}{L_0+1} - 4 \geq \frac{46L_0-4}{L_0+1} \geq 21$. We need to ensure that for all $x, y \in \Lambda$, $|x-y| \geq L_1/2$ we have

$$(8d\varepsilon^2(2L_0 + 1)^{d-1}\delta_0^{-1}\varepsilon^{\gamma_0 L_0})^{\frac{|x-y|}{L_0+1}-4} \leq \varepsilon^{\gamma_1|x-y|}$$

which then implies (3.14) via (3.18). Taking logarithms, this reduces to

$$\left[\gamma_0 - \gamma_1 \left(1 - \frac{1}{L_0 + 1} - \frac{4L_0}{|x-y|} \right)^{-1} + \frac{2}{L_0} \right] \log \frac{1}{\varepsilon} \geq L_0^{-1} \log(8d(2L_0 + 1)^{d-1}\delta_0^{-1})$$

The worst case here is $|x-y| = L_1/2$ which gives (3.15) (see Fig. 2). $\square$

Definition 3.9. Fix any $x_0 \in \mathbb{Z}^d$. Then we define an L -box $\Lambda_L(x_0)$ to be (γ, E) -regular if it exhibits

- *non-resonance at energy E* : $\text{dist}(\text{spec}(H_{\Lambda_L(x_0)}), E) \geq \delta(L) = \exp(-L^\beta)$
- *exponential Green function decay*: $|G_{\Lambda_L(x_0)}(E)(x, y)| \leq 4\delta(L)^{-1}\varepsilon^{\gamma|x-y|} \quad \forall x, y \in \Lambda_L(x_0) \text{ with } |x-y| \geq L/2$

Here $E \in \mathbb{R}$ is arbitrary, $\gamma > 0$ will be specified below, depending on the scale, and $\beta \in (0, 1)$ will be a fixed constant. A box is (γ, E) -singular if it is not (γ, E) -regular.

At the initial scale of the induction, by [Corollary 3.7](#)

$$\sup_{E \in \mathbb{R}} \mathbb{P}(\{\exists \Lambda_{L_0}(x_0) \text{ which is } (1, E)\text{-singular}\}) \leq 4B(L_0)|Q_{L_0}(x_0)|\delta(L_0) =: p_0 \quad (3.19)$$

The existence inside the set refers to all possible boxes of the initial scale $L_0 = L_0(d, \beta) \geq 100$ centered at x_0 of which there are $B(L_0) = (2(L_0 + 1))^d$, while $|Q_{L_0}(x_0)| = (2L_0 + 1)^d$ is the volume of the largest L_0 -box. Thus we have

$$p_0 = 4(2(L_0 + 1))^d(2L_0 + 1)^d \exp(-L_0^\beta)$$

where β is just chosen here to so that $\exp(-L_0^\beta) = \delta(L_0)$ and will in fact be in $(0, 1)$. [Corollary 3.7](#) requires that, where $\delta_0 := \delta(L_0)$,

$$4d\varepsilon \leq \delta_0. \quad (3.20)$$

Set $L_1 = \lceil L_0^\alpha \rceil$ where $\alpha > 1$ will also be specified later. By [Lemma 3.8](#), whose condition [\(3.13\)](#) we leave to the reader to verify,

$$\begin{cases} \sup_{E \in \mathbb{R}} \mathbb{P}(\{\exists \Lambda_{L_1}(x_0) \text{ which is } (\gamma_1, E)\text{-singular}\}) \leq p_1 \\ p_1 := 4B(L_1)|Q_{L_1}(x_0)|\delta_1 + |Q_{L_1}(x_0)|^2 p_0^2 \\ \gamma_1 := (1 - \frac{1}{L_0+1} - \frac{8L_0}{L_1})(1 - L_0^{\beta-1}) \end{cases} \quad (3.21)$$

In fact, p_1 is the sum of two contributions. On the one hand, Wegner's estimate gives, with $\delta_1 = \exp(-L_1^\beta)$,

$$\mathbb{P}(\{\exists \Lambda_{L_1}(x_0) \text{ with } \text{dist}(\text{spec}(H_{\Lambda_{L_1}(x_0)}), E) \leq \delta_1\}) \leq 4B(L_1)|Q_{L_1}(x_0)|\delta_1$$

which is the first term on the right-hand side of p_1 . It controls the probability that one of the boxes $\Lambda_{L_1}(x_0)$ is resonant at energy E with resonance width δ_1 . The other term bounds

$$\mathbb{P}(\{\exists \text{ two disjoint } (1, E)\text{-singular } L_0 \text{ boxes in } Q_{L_1}(x_0)\}) \leq |Q_{L_1}(x_0)|^2 p_0^2$$

where the factor $|Q_{L_1}(x_0)|^2 = (2L_1 + 1)^{2d}$ is a result of selecting the centers of the L_0 boxes in $Q_{L_1}(x_0)$. Assuming $L_0^{\beta(\alpha-1)} \geq 2$, we have $\delta_1 \leq \delta_0^2 \leq p_0^2$ and thus $p_1 \leq 5B(L_1)^2 p_0^2$. Finally, setting $\gamma_0 = 1$ and γ_1 as above in [\(3.15\)](#) yields

$$(1 + 2L_0^{-\beta}) \log \frac{1}{\varepsilon} \geq 1 + L_0^{-\beta} (\log(8d) + (d-1) \log(2L_0 + 1))$$

In view of [\(3.20\)](#) this holds for $L_0(d, \beta)$ sufficiently large, proving [\(3.21\)](#).

Inductively, define $L_{k+1} = \lceil L_k^\alpha \rceil$ with $1 < \alpha < 2$ fixed. In analogy with [\(3.21\)](#) one has with $\delta_k = \exp(-L_k^\beta)$,

$$\begin{cases} \sup_{E \in \mathbb{R}} \mathbb{P}(\{\exists \Lambda_{L_k}(x_0) \text{ which is } (\gamma_k, E)\text{-singular}\}) \leq p_k \\ p_k := 4B(L_k)|Q_{L_k}(x_0)|\delta_k + |Q_{L_k}(x_0)|^2 p_{k-1}^2 \\ \gamma_k := \gamma_{k-1} (1 - \frac{1}{L_{k-1}+1} - \frac{8L_{k-1}}{L_k})(1 - L_{k-1}^{\beta-1}) \end{cases} \quad (3.22)$$

We leave the condition [\(3.13\)](#) to the reader to verify. Note that

$$p_k \leq B(L_k)^2 (4\delta_k + p_{k-1}^2) \leq (2(L_k + 1))^{2d} (4\delta_k + p_{k-1}^2)$$

On the one hand, if L_0 is large enough, then

$$\prod_{k=1}^{\infty} \left(1 - \frac{1}{L_{k-1}+1} - \frac{8L_{k-1}}{L_k}\right) (1 - L_{k-1}^{\beta-1}) \geq \prod_{k=1}^{\infty} (1 - 9L_{k-1}^{1-\alpha})(1 - L_{k-1}^{\beta-1}) > 0$$

since $L_k \geq (L_0)^{\alpha^k}$ and $\sum_{j \geq 0} (L_j^{1-\alpha} + L_j^{\beta-1}) = o(1)$ as $L_0 \rightarrow \infty$, whence $\inf_{k \geq 1} \gamma_k > 0$ (approaches 1 for large L_0). On the other hand, we claim that $\sum_{k=0}^{\infty} p_k < \infty$. Indeed, from (3.22),

$$\begin{aligned} L_{k+1}^m p_k &\leq 4L_{k+1}^m B(L_k)^2 \delta_k + [(L_k^\alpha + 1)^{\frac{m}{2}} (2L_k + 1)^d p_{k-1}]^2 \\ &\leq 4L_{k+1}^m (2(L_k + 1))^{2d} \delta_k + (L_k^m p_{k-1})^2 \end{aligned} \quad (3.23)$$

where the second line holds provided $\alpha m/2 + d < m$ which requires $\alpha < 2$, and for L_0 large enough. We conclude from (3.23) that $L_{k+1}^m p_k \leq 1$ if L_0 is large. Moreover, due to

$$\sum_k 4L_{k+1}^m (2(L_k + 1))^{2d} \delta_k < \infty \quad \text{also} \quad \sum_k L_{k+1}^m p_k < \infty$$

which is stronger than the claim. From the preceding analysis, the parameters need to be in the ranges $1 < \alpha < 2$ and $0 < \beta < 1$. To summarize, we have obtained this result.

Proposition 3.10. *Fix $1 < \alpha < 2$ and $0 < \beta < 1$. For $L_0 = L_0(d, \alpha, \beta)$ large enough, define scales $L_{k+1} = \lceil L_k^\alpha \rceil$ for $k \geq 0$. Then for arbitrary $x_0 \in \mathbb{Z}^d$ and $E \in \mathbb{R}$,*

$$\mathbb{P}(\{\text{all boxes } \Lambda_{L_k}(x_0) \text{ are } (\gamma_k, E)\text{-regular}\}) \geq 1 - p_k \quad (3.24)$$

with $0 < p_k \leq L_{k+1}^{-m} \leq L_0^{-m\alpha^{k+1}}$ for all $k \geq 0$. Here $m > \frac{2d}{2-\alpha}$ and $L_0(d, \alpha, \beta, m)$ is sufficiently large. The p_k depend neither on x_0 nor on E , and $\gamma_k \geq \frac{1}{2}$ for all k .

Remark 3.11. We shall use below that (3.24) holds as stated for $k \geq 1$ if we weaken the non-resonance condition in Definition 3.9 to the following one: $\text{dist}(\text{spec}(H_{\Lambda_{L_k}(x_0)}), E) \geq \delta(L_k)/4$. This is due to some room built into Lemma 3.8, cf. the factor $4\delta_0^{-1}$ in (3.12) which improves to δ_1^{-1} in (3.14). This allows us to replace $\delta(L_k)$ in the resonance width with $\delta(L_k)/4$.

An essential feature in the derivation of this result is *stability in the energy*. This means that we can obtain (3.24) *uniformly in an energy interval* of length half of the resonance width.

Corollary 3.12. *Under the assumptions of the previous proposition the following holds: for arbitrary $x_0 \in \mathbb{Z}^d$ and $E_* \in \mathbb{R}$,*

$$\mathbb{P}(\{\text{all boxes } \Lambda_{L_k}(x_0) \text{ are } (\gamma_k, E)\text{-regular for all } E \in [E_* - \delta_k/2, E_* + \delta_k/2]\}) \geq 1 - p_k \quad (3.25)$$

for all $k \geq 1$ and the same p_k as above.

Proof. We leave the base case $k = 1$ to the reader. The inductive step $k - 1 \rightarrow k$ with $k \geq 2$, consists of the inequality (dropping x_0 for simplicity)

$$\begin{aligned} &\mathbb{P}(\{\exists \Lambda_{L_k} \text{ which is } (\gamma_k, E)\text{-singular for some } E \in [E_* - \delta_k/2, E_* + \delta_k/2]\}) \\ &\leq \mathbb{P}(\{\exists \Lambda_{L_k} \text{ with } \text{dist}(H_{\Lambda_{L_k}}, E) \leq \delta_k/2 \text{ for some } E \in [E_* - \delta_k/2, E_* + \delta_k/2]\}) \\ &\quad + \mathbb{P}(\{\exists \Lambda_{L_k} \text{ with } \text{dist}(H_{\Lambda_{L_k}}, E) \geq \delta_k/2 \text{ for some } E \in [E_* - \delta_k/2, E_* + \delta_k/2]\}) \\ &\quad \text{which is } (\gamma_k, E)\text{-singular for the same } E \\ &\leq \mathbb{P}(\{\exists \Lambda_{L_k} \text{ with } \text{dist}(H_{\Lambda_{L_k}}, E_*) \leq \delta_k\}) \\ &\quad + \mathbb{P}(\{\exists \Lambda_{L_k} \text{ which contains two disjoint } L_{k-1}\text{-boxes which are both} \\ &\quad (\gamma_{k-1}, E)\text{-singular for the same } E \in [E_* - \delta_{k-1}/2, E_* + \delta_{k-1}/2]\}) \end{aligned}$$

The final two lines here follow from Lemma 3.8, see also Remark 3.11. Note how we widened the E -interval in the last line, which makes it clear how to use the inductive assumption. The proof proceeds exactly as before. $\square$

This result cannot by itself establish localization, since it only controls the resonance of H_A with a given energy E on a *single box* A . Localization requires excluding simultaneous resonances on several disjoint boxes. This in turn allows us to *eliminate the energy* E from these events, and thus estimate them uniformly over all energies. It suffices to carry out this process on two disjoint boxes, in other words, to show that *double resonances* are highly unlikely. The following natural result contains the elimination of energies and absence of double resonances in its proof, but not in the statement. Note, however, that the event of low probability described in the following proposition is uniform in all energies.

Proposition 3.13. *Under the assumptions of the previous proposition, for all $k \geq 1$,*

$$\mathbb{P}(\{\text{for some } E \text{ a box } A_{L_k}(x_0) \text{ is nonresonant at } E \text{ but } (\gamma_k, E)\text{-singular}\}) \leq q_k \quad (3.26)$$

where for any $b > 1$ and all k , $q_k \leq L_{k+1}^{-b}$ provided L_0 is large (and thus ε is small) enough. Here nonresonant is as in Definition 3.9 but with $\delta_k/2$.

Proof. E -nres stands for nonresonant at energy E , E -res for resonant at E , and $sing$ for singular. Let $A_1 = A_{L_1}(x_0)$ be E -nres, i.e., $\text{dist}(\text{spec}(H_{A_1}), E) \geq \delta_1/2$ but (γ_1, E) -sing. By Lemmas 3.8 and 3.6 the event in (3.26) implies that there must be at least two resonant L_0 -box inside of $A_{L_1}(x_0)$ (here but only here we measure resonance with δ_0 and not $\delta_0/2$). Hence

$$\begin{aligned} & \mathbb{P}(\{A_{L_1}(x_0) \text{ is } E\text{-nres but } (\gamma_1, E)\text{-sing}\}) \\ & \leq \mathbb{P}(\{A_{L_1}(x_0) \text{ contains two disjoint } L_0\text{-boxes, both } E\text{-res}\}) \\ & \leq \sum_{A_{L_1}(x_0)} \sum_{A'_{L_0} \subset A_{L_1}(x_0)} \sum_{\tilde{A}'_{L_0} \subset A_{L_1}(x_0) \setminus A'_{L_0}} \mathbb{P}(\{\text{dist}(\text{spec}(H_{\tilde{A}'_{L_0}}), E_j) \leq 2\delta_0 \text{ for some } E_j \in \text{spec}(H_{A'_{L_0}})\}) \\ & \leq 8B(L_1)B(L_0)^2|Q_{L_1}|^2|Q_{L_0}|^2\delta_0 =: q_1 \end{aligned}$$

In the third line the energy is eliminated by δ_0 -closeness of E to some eigenvalue E_j of $H_{A'_{L_0}}$, and the fourth line is Wegner's estimate. The sum over $A_{L_1}(x_0)$ expresses the existence of some L_1 -box with the stated property. At scale L_k , $k \geq 2$, and suppressing x_0 for simplicity,

$$\begin{aligned} & \mathbb{P}(\{\text{for some } E \text{ a box } A_{L_k} \text{ is } E\text{-nres but } (\gamma_k, E)\text{-sing}\}) \\ & \leq \mathbb{P}(\{\text{for some } E \text{ a box } A_{L_k} \text{ is } E\text{-nres, contains two disjoint } L_{k-1}\text{-boxes, both } (\gamma_{k-1}, E)\text{-sing}\}) \\ & \leq \mathbb{P}(\{\text{for some } E \text{ a box } A_{L_k} \text{ contains two disjoint } L_{k-1}\text{-boxes, one } E\text{-res, the other } (\gamma_{k-1}, E)\text{-sing}\}) \\ & + \mathbb{P}(\{\text{for some } E \text{ a box } A_{L_k} \text{ contains two disjoint } L_{k-1}\text{-boxes, both } E\text{-nres, but } (\gamma_{k-1}, E)\text{-sing}\}) \end{aligned}$$

In analogy with $k = 1$ we bound the third line by

$$\begin{aligned} & \leq \sum_{A_{L_k}} \sum_{A'_{L_{k-1}} \subset A_{L_k}} \sum_{y_0 \in A_{L_k}} \mathbb{P}(\{\text{some box } \tilde{A}'_{L_{k-1}}(y_0) \subset A_{L_k} \setminus A'_{L_{k-1}} \text{ is } (\gamma_{k-1}, E)\text{-sing} \\ & \quad \text{with } |E - E_j| \leq \delta_{k-1}/2, E_j \in \text{spec}(H_{A'_{L_{k-1}}})\}) \\ & \leq B(L_k)B(L_{k-1})|Q_{L_k}|^2|Q_{L_{k-1}}|p_{k-1} \end{aligned}$$

where the final estimate is given by Corollary 3.12 with $E_* = E_j$. Note that while E_j is random, these variables are independent from $H_{\tilde{A}'_{L_{k-1}}(y_0)}$. Hence we may first condition on the random variables in $H_{A'_{L_{k-1}}}$. The fourth line above is bounded by the inductive assumption and independence, and so it is $\leq B(L_k)|Q_{L_k}|^2q_{k-1}^2$. In summary, by Proposition 3.10,

$$\begin{aligned} q_k & \leq B(L_k)B(L_{k-1})|Q_{L_k}|^2|Q_{L_{k-1}}|p_{k-1} + B(L_k)|Q_{L_k}|^2q_{k-1}^2 \\ & \leq (2(L_k + 1))^{5d}L_k^{-m} + (2(L_k + 1))^{3d}q_{k-1}^2 \end{aligned} \quad (3.27)$$

and we conclude as for (3.23) that $q_k \leq L_{k+1}^{-b}$ for all b provided L_0 is taken large enough depending on b . $\square$

Proof of Theorem 3.5. Let $\mathcal{B}_k(x_0)$ be the event in (3.26). We remove the 0-probability event

$$\mathcal{B} = \bigcup_{x_0 \in \mathbb{Z}^d} \limsup_{k \rightarrow \infty} \mathcal{B}_k(x_0).$$

Considering a realization of the random operator H off of this event, for spectrally almost every energy E relative to this operator we can find a nontrivial generalized eigenfunction $H\psi = E\psi$ which is at most polynomially growing, say $|\psi(n)| \leq C(\sigma, \psi)|n|^\sigma$ with $\sigma > \frac{d}{2}$ and all $n \in \mathbb{Z}^d$, $n \neq 0$. Let $\psi(x_0) \neq 0$. Suppose $\Lambda_{L_k}(x_0)$ is E -nonresonant for infinitely many k . Then by Lemma 3.6, for those k ,

$$\max_{(y', y) \in \partial \Lambda_{L_k}(x_0)} |G_{\Lambda_{L_k}(x_0)}(E)(x_0, y')| \leq \varepsilon^{\frac{1}{2}L_k}$$

Then $[(H_{\Lambda_{L_k}(x_0)} - E)\psi](y') = -\sum_{(y', y) \in \partial \Lambda_{L_k}(x_0)} \psi(y)$,

$$|\psi(x_0)| \leq \sum_{(y', y) \in \partial \Lambda_{L_k}(x_0)} |G_{\Lambda_{L_k}(x_0)}(E)(x_0, y')| |\psi(y)| \leq C(\sigma, \psi) \sum_{(y', y) \in \partial \Lambda_{L_k}(x_0)} \varepsilon^{\frac{1}{2}L_k} L_k^\sigma \quad (3.28)$$

which is impossible for infinitely many k . Hence for all $k \geq k_0(\psi)$, $\Lambda_{L_k}(x_0)$ is E -resonant. We now remove another 0-probability event, namely *double resonances* between disjoint boxes which are not too far from each other. To be specific, as above we conclude that, a.s. for every x_0 and all but finitely many k ,

$$\forall E \in \mathbb{R} \text{ if } \Lambda_{L_k}(x_0) \text{ is } E\text{-res then } \forall 2L_k < |y_0 - x_0| \leq 100L_{k+1}, \Lambda_{L_k}(y_0) \text{ is } (\gamma_k, E)\text{-reg} \quad (3.29)$$

Indeed, the resonance condition ensures that E is $\delta_k/2$ -close to one of the (random) eigenvalues of $H_{\Lambda_{L_k}(x_0)}$, and Corollary 3.12 bounds the probability that one of the boxes $\Lambda_{L_k}(y_0)$ is (γ_k, E) -regular by L_{k+1}^{-m} where $m > 2d$, say. Hence we can sum this up over all y_0 in a $100L_{k+1}$ -box and apply Borel–Cantelli as before. Consequently, *all boxes* $\Lambda_{L_k}(y_0)$ are regular as stated in (3.29). By a resolvent expansion as in the proof of Lemma 3.8, the reader will easily verify that all Green functions $G_{\Lambda_{L_k}(y_0)}(E)(x, y)$ have exponential decay if $\Lambda_{L_k}(y_0) \subset \Lambda_{100L_{k+1}}(x_0) \setminus \Lambda_{2L_k}(x_0)$ where we take $L_{k+1} \leq L \leq 50L_{k+1}$. By an estimate as in (3.28) one now concludes exponential decay of ψ . $\square$

4. The one-dimensional quasi-periodic model

4.1. The Fröhlich–Spencer–Wittwer theorem: even potentials

In this section we will provide a fairly complete proof sketch of the following result due to Fröhlich, Spencer, and Wittwer [23]. The dynamics (rotation) $T_\omega \theta = \theta + \omega \pmod{1}$ takes place on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and all “randomness” sits in a single parameter, namely $\theta \in \mathbb{T}$. The one-dimensional *random* model is treated by completely different techniques, starting from Fürstenberg’s classical theorem on positive Lyapunov exponents for random $\mathrm{SL}(2, \mathbb{R})$ cocycles, cf. [50] for a comprehensive exposition of this fundamental result as well as Lyapunov exponents in general. See the recent papers [12, 30], and [37] for streamlined elegant treatments of the 1-dimensional random Anderson model, including the Bernoulli case. For quasi-periodic (and other highly correlated) cocycles, Fürstenberg’s global theorem does not apply, and other techniques must be used. The proof of the following result will in fact be perturbative.

Theorem 4.1. *Let $v \in C^2(\mathbb{T})$ be even, with exactly two nondegenerate critical points. Define*

$$H_\varepsilon(\theta) = \varepsilon^2 \Delta_{\mathbb{Z}} + V_\theta, \quad V_\theta(n) = V(T_\omega^n \theta) \quad \forall n \in \mathbb{Z} \quad (4.1)$$

where $\omega \in \mathbb{T}$ is Diophantine, viz. $\|n\omega\| \geq b_0 n^{-2}$ for all $n \geq 1$ with some $b_0 > 0$. There exists $\varepsilon_0(b_0, V)$ such that for all $0 < \varepsilon \leq \varepsilon_0$ the operators $H_{\theta, \varepsilon}$ exhibit Anderson localization for a.e. $\theta \in \mathbb{T}$.

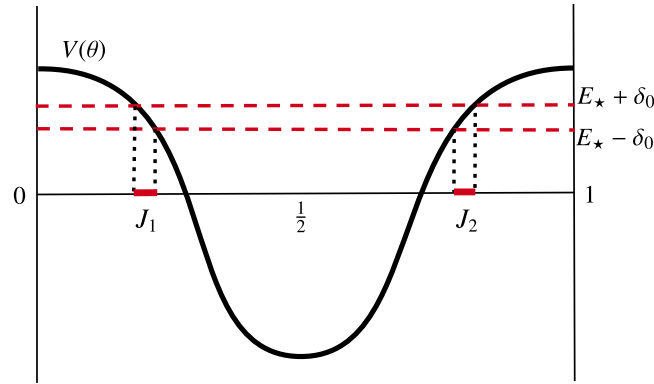


Fig. 3. The potential and energy strip at the initial step.

The evenness assumption allows for substantial simplifications as we shall see. Note that it entails that V is symmetric about $\frac{1}{2}$. Theorem 4.1 cannot hold for all θ , see [36]. As in the previous section, we shall drop the index ε and simply write $H(\theta)$ for (4.1), and $H_\Lambda(\theta)$ for its finite volume version. It is important to keep track of θ so we include it in the notation (while in the random case we could drop the ω , the variable in the probability space). Fix $\theta_* \in \mathbb{T}$ and $E_* \in \mathbb{R}$. The singular sites relative to θ_*, E_* are defined as

$$\begin{aligned} \mathcal{S}_0 &= \mathcal{S}_0(\theta_*, E_*) := \{n \in \mathbb{Z} \mid |V(\theta_* + n\omega) - E_*| \leq \delta_0\} \\ &= \{n \in \mathbb{Z} \mid T^n \theta_* \in V^{-1}([E_* - \delta_0, E_* + \delta_0])\} \end{aligned} \quad (4.2)$$

Fig. 3 shows one scenario for which $V^{-1}([E_* - \delta_0, E_* + \delta_0]) = J_1 \cup J_2$ with disjoint intervals. There might be just one interval or the set could be empty. By our assumption of V being a Morse function, $\max_{i=1,2} |J_i| \leq C_0(V) \delta_0^{\frac{1}{2}}$ for all cases. We choose the *resonance width* $\delta_0 = A_0 \varepsilon$ with a large constant A_0 . We investigate the structure of $\mathcal{S}_0$ by means of the example $V(\theta) = \cos(2\pi\theta)$. If $k, \ell \in \mathcal{S}_0$ are distinct, then

$$|\sin(\pi(k - \ell)\omega) \sin(\pi(2\theta_* + (k + \ell)\omega))| \leq \delta_0$$

which implies for small δ_0 that

$$m(k, \ell) := \min(\|(k - \ell)\omega\|, \|2\theta_* + (k + \ell)\omega\|) \leq 2\sqrt{\delta_0}$$

The first alternative here, viz. $\|(k - \ell)\omega\| \leq 2\sqrt{\delta_0}$ occurs precisely if both $T_\omega^k \theta_*$ and $T_\omega^\ell \theta_*$ fall into J_1 , or both fall into J_2 . The second one occurs if they fall into different intervals. The Diophantine assumption implies that

$$b_0 |k - \ell|^{-2} \leq \|(k - \ell)\omega\| \leq 2\sqrt{\delta_0}, \quad |k - \ell| \gtrsim \delta_0^{-\frac{1}{4}}$$

Henceforth $\gtrsim$ and $\lesssim$ will indicate multiplicative constants depending on b_0, V . On the other hand, $\|2\theta_* + (k + \ell)\omega\| \leq 2\sqrt{\delta_0}$ might occur for $\ell = k + 1$ which is the case if $T_\omega(J_1) \cap J_2 \neq \emptyset$. It is clear that the function m appears not just for cosine, but in fact for any V as in the theorem.

Lemma 4.2. *Any two distinct $k, \ell \in \mathcal{S}_0$ satisfy $m(k, \ell)^2 \lesssim \delta_0$, and any three distinct points $k, \ell, n \in \mathcal{S}_0$ satisfy*

$$\max(|k - \ell|, |\ell - n|) \gtrsim \delta_0^{-\frac{1}{4}}$$

Proof. The argument is essentially the same as for cosine, the trigonometric identities being replaced by the symmetry of V on $\mathbb{T}$ about $\frac{1}{2}$: if $\theta_* + k\omega \in J_1$ and $\theta_* + \ell\omega \in J_2 = -J_1 \pmod{1}$, then $2\theta_* + (k + \ell)\omega \in J_1 - J_1 \pmod{1}$ whence $\|2\theta_* + (k + \ell)\omega\|^2 \lesssim \delta_0$. $\square$

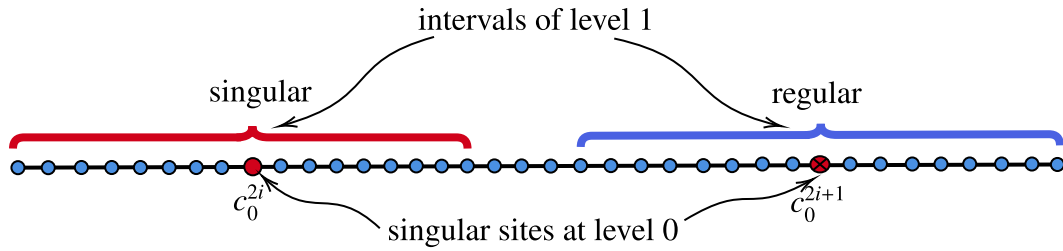


Fig. 4. Simple resonances at level 1.

Definition 4.3. We label $\mathcal{S}_0 = \{c_0^i\}_{i=-\infty}^\infty$ in increasing order (assuming $\mathcal{S}_0 \neq \emptyset$). These are the singular sites (or singular “intervals”) at level 0. Let $s_0 := \min\{c_0^i - c_0^j \mid i > j\}$. If $s_0 \geq 4|\log \varepsilon|^2$, then we speak of a *simple resonance*, otherwise of a *double resonance*, both at level 1. In the latter case, we replace $\mathcal{S}_0$ with $\tilde{\mathcal{S}}_0 = \mathcal{S}_0 \cup (\mathcal{S}_0 + s_0)$, which we again label as $\{c_0^i\}_{i=-\infty}^\infty$. In the simple resonant case, we let I_1^i be an interval of length $\ell_1 := \lceil \log(1/\varepsilon) \rceil^2$ centered at $c_1^i := c_0^i$, in the double resonant case I_1^i has length $\ell_1 := \lceil \log(1/\varepsilon) \rceil^4$, centered at $c_1^i := (c_0^{2i} + c_0^{2i+1})/2 \in \frac{1}{2}\mathbb{Z}$. By construction, all of the I_1^i are pairwise disjoint, and each c_0^i is contained in a unique interval at level 1. We classify those intervals I_1^i as *singular* provided

$$\text{dist}(\text{spec}(H_{I_1^i}(\theta_*)), E_*) \leq \delta_1 := \varepsilon \ell_1^{2/3} \quad (4.3)$$

and $\mathcal{S}_1 := \{c_1^i \mid I_1^i \text{ is singular}\}$. All other intervals I_1^i are called *regular*.

We shall see later, based on [Theorem 2.1](#), that for spectrally a.e. energy $E \in \text{spec}(H(\theta))$ the set of singular intervals, which are constructed iteratively at all levels (see below), is not empty. [Figs. 4](#), resp. [5](#) illustrate the two cases, with the blue dots being $\mathbb{Z} \setminus \mathcal{S}_0$. The terminology simple/double resonance is derived from the structure of the eigenfunctions at level 1 associated with the operators $H_{I_1^i}(\theta_*)$ and the unique (as we shall see) eigenvalue $E_1^i(\theta_*)$ satisfying (4.3). In the simple resonance case, the eigenfunction has most (say 99%) of its ℓ^2 mass at the center c_0^i , whereas in the double resonant case it may have significant mass at both sites c_0^{2i} and c_0^{2i+1} .

[Fig. 4](#) depicts only one of four possibilities for the intervals at level 1, they might both be singular, both regular, or the order could be reversed. The red dot with $\otimes$ is supposed to indicate a return of the trajectory $T_\omega^j \theta_*$ to J_2 , whereas the red dot on the left a return to J_1 , cf. [Fig. 3](#). While the distance between these two red dots is required to be at least $4(\log(1/\varepsilon))^2$, the Diophantine condition forces two red dots of the same kind (associated with J_1 , resp. J_2) to be separated by a distance exceeding $\delta_0^{-\frac{1}{4}}$. This is much larger than the length $\ell_1 = \lceil \log(1/\varepsilon) \rceil^2$ of the intervals I_1^i .

The reason for passing from $\mathcal{S}_0$ to $\tilde{\mathcal{S}}_0$ lies with the self-symmetry indicated in [Fig. 5](#) (i.e., $c_0^{2i+1} - c_0^{2i}$ does not depend on i). To see this, note that by definition of s_0 there exist $k_1, k_2 \in \mathcal{S}_0$ with $\theta_* + k_i \omega \in J_i$ with $i = 1, 2$ and $s_0 = |k_1 - k_2| \leq 4(\log(1/\varepsilon))^2$. We are again using the Diophantine condition here to ensure that we do not fall into the same interval (as a standing assumption ε needs to be small enough depending on v and c_0 so as to guarantee this). Next, take any $k \in \mathcal{S}_0$ with $\theta_* + k\omega \in J_1$ (everything modulo integers which will be henceforth understood tacitly). Then $\theta_* + (k + s_0)\omega \in \tilde{J}_2$, where $\tilde{J}_2$ has the same center as J_2 and twice the length. On the other hand, it might be that $\theta_* + (k + s_0)\omega \notin J_2$, but we must still include $k + s_0$ in $\mathcal{S}_0$ for the construction to work. In fact, [Lemma 4.2](#) remains valid for $\tilde{\mathcal{S}}_0$ and the defining inequality (4.2) is modified only slightly, viz. $|V(\theta_* + n\omega) - E_*| \lesssim \delta_0$ for all $n \in \tilde{\mathcal{S}}_0$.

We will establish the following analogue of [Lemma 4.2](#) at level 1. We emphasize again that this statement only exists for even V .

Lemma 4.4. For all $c_1^i, c_1^j \in \mathcal{S}_1$ one has $m(c_1^i, c_1^j)^2 \lesssim \delta_1$, with an absolute implied constant.

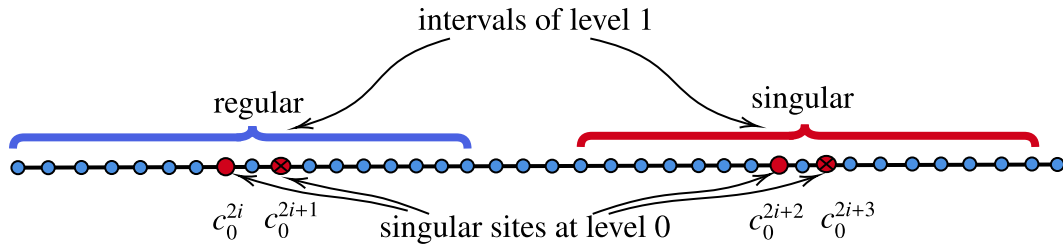


Fig. 5. Double resonances at level 1.

The idea is to carry out a similar argument as for Lemma 4.2, with the potential function V replaced by the parameterizations of the eigenvalues of the operator $H_{I_1}(\theta)$ localized both near θ_* and E_* . This stability hinges crucially on a *spectral gap* or on the *separation of the eigenvalues*. The latter can be seen as a quantitative version of the simplicity of the Dirichlet spectrum of Sturm–Liouville operators, such as $H_I(\theta)$. Before discussing the details of Lemma 4.4, we exhibit the entire strategy of the proof of Theorem 4.1.

- In analogy with Definition 4.3 define regular and singular intervals at level $n \geq 2$. More specifically, for $n \geq 1$ set

$$s_n := \min\{|c_n^i - c_n^j| \mid c_n^i, c_n^j \in \mathcal{S}_n, i \neq j\}$$

If $s_n > 4\ell_n^2$, then we call this a simple resonance and define $c_{n+1}^i = c_n^i$ for all i and $\ell_{n+1} = \ell_n^2$, otherwise for the double resonance we set $c_{n+1}^i = (c_n^{2i} + c_n^{2i+1})/2 \in \frac{1}{2}\mathbb{Z}$, $\ell_{n+1} = \ell_n^4$, and also augment $\mathcal{S}_n$ to $\bar{\mathcal{S}}_n$ by including the mirror image of each I_n^i if it was not already included in $\mathcal{S}_n$. By *mirror image* we mean an interval $\tilde{I}_n^i$ of the same length as I_n^i , with center $\tilde{c}_n^i \in \mathbb{Z}$ so that $|c_n^i - \tilde{c}_n^i| = s_n$ and $\|2\theta_* + (c_n^i + \tilde{c}_n^i)\omega\| \leq 12\sqrt{\delta_n}$. The meaning of mirror image is that

$$V(\theta_* + (-j + c_n^i)\omega) = V(\theta_* + (j + \tilde{c}_n^i)\omega) + O(\sqrt{\delta_n}) \quad \forall j \in \mathbb{Z}$$

The existence of such an interval is left to the reader, see also [23, Lemma 1.3]. By construction, the I_{n+1}^i are pairwise disjoint and each c_n^i is contained in a unique interval at level $n+1$. An interval I_{n+1}^i centered at c_{n+1}^i is called *singular* if

$$\text{dist}(\text{spec}(H_{I_{n+1}^i}(\theta_*)), E_*) \leq \delta_{n+1} = \varepsilon \ell_{n+1}^{2/3} \quad (4.4)$$

and regular otherwise. Define $\mathcal{S}_{n+1}$ to be the centers of the singular intervals. One can arrange for ∂I_{n+1}^i for all singular not to meet any singular interval of level m with $m \leq n$.

- An arbitrary interval $A \subset \mathbb{Z}$ is called n -regular provided every point in $A \cap \mathcal{S}_0$ is contained in a regular interval $I_m^j \subset A$ for some $m \leq n$, cf. Fig. 6. Note that every singular point at level 0 is either (i) contained in infinitely many singular intervals $I_n^{j_n}$ for each $n \geq 0$ or (ii) contained in a finite number of such intervals at successive levels followed by a regular one. By induction on scales one proves the following crucial decay and stability property of the Green function associated with n -regular intervals A : $|G_A(\theta, E)(x, y)| \leq \varepsilon^{\frac{1}{2}|x-y|}$ for all $x, y \in A$, $|x - y| \geq \ell_n^{5/6}$, $|E - E_*| \lesssim \delta_n$, and $|\theta - \theta_*| \lesssim \delta_n$.
- One has

$$m(c_n^i, c_n^j)^2 \lesssim |E_n^i(\theta_*) - E_n^j(\theta_*)| \lesssim \delta_n \quad \forall c_n^i, c_n^j \in \mathcal{S}_n \quad (4.5)$$

for all $n \geq 0$. Here $E_n^i(\theta_*)$, are the unique eigenvalues of $H_{I_n^i}(\theta_*)$ in the interval $[E_* - c\delta_n, E_* + c\delta_n]$ with c small. This hinges crucially on the separation property of the eigenvalues, see Lemma 4.6. For simple resonances, we will use mainly first order eigenvalue perturbation theory, and for double resonances, rely crucially on second order perturbation theory. However, some second order arguments are already needed for simple resonances.

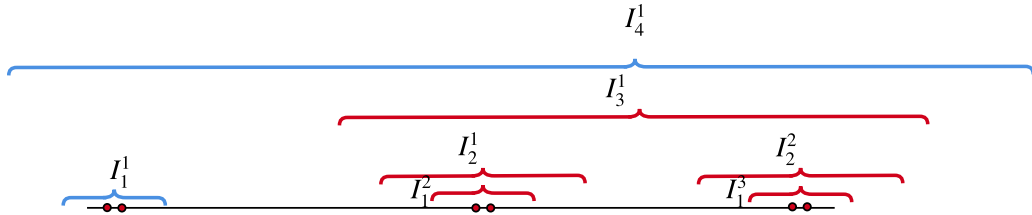


Fig. 6. Regular and singular intervals of 4 levels.

- Based on the estimate on m , we prove [Theorem 4.1](#) by double resonance elimination as in the previous section. In analogy with [Theorem 3.5](#) we start with the polynomially bounded Fourier basis provided by [Theorem 2.1](#), find an increasing nested family of resonant intervals which are resonant at the given energy, and thus due to the elimination of double resonances obtain exponential decay at all scale. The main departure from the proof of [Theorem 3.5](#) lies with the application of Borel–Cantelli to remove a zero measure set of bad $\theta \in \mathbb{T}$.

We begin with the Green function decay on regular intervals (this is the analogue of the regular Green function from [Definition 3.9](#)). We set $\ell_0 := \lceil \log(1/\varepsilon) \rceil$.

Lemma 4.5. *For all n -regular intervals Λ , $n \geq 0$, one has $|G_\Lambda(\theta, E)(x, y)| \leq \varepsilon^{\gamma_n |x-y|}$ for all $x, y \in \Lambda$, $|x - y| \geq \ell_n^{5/6}$, $|E - E_*| \lesssim \delta_n$, and $|\theta - \theta_*| \lesssim \delta_n$. The γ_n decrease, but $\gamma_n \geq \frac{1}{2}$ for all n .*

Proof. At $n = 0$ the interval Λ contains only regular lattice points, i.e., blue dots in the figures above. Then the Neumann series argument from [Lemma 3.6](#) implies that, for $\delta_0 = A_0 \varepsilon$ with A_0 large enough, and for all $|\theta - \theta_*| \ll \delta_0$, $|E - E_*| \ll \delta_0$ (meaning up to a small multiplicative constant),

$$|G_\Lambda(\theta, E)(x, y)| \leq \delta_0^{-1} (2\varepsilon^2 \delta_0^{-1})^{|x-y|} \leq \varepsilon^{|x-y|-1} \quad \forall x, y \in \Lambda$$

We are using that $|V(\theta_* + k\omega) - E_*| \gtrsim \delta_0$ implies that $|V(\theta + k\omega) - E| \gtrsim \delta_0$ in the specified range of parameters. If Λ is 1-regular, then let $\{I_1^i\}_{i=i_0}^{i_1}$ be a complete list of all level 1 intervals, in increasing order, which cover all points in $\Lambda \cap \mathcal{S}_0$. By construction, $I_1^i \subset \Lambda$ for all $i_0 \leq i \leq i_1$. The intervals I_1^i (which are all regular) do not really come off the axis in [Fig. 7](#), they are only depicted in this way to indicate that they are level 1 intervals. The line segment is supposed to depict Λ and it consists entirely of regular lattice points at level 0 apart from the red singular sites. For the double resonance case shown in [Fig. 7](#), one red pair is separated from another by $\gtrsim \delta_0^{-\frac{1}{4}} \simeq \varepsilon^{-\frac{1}{4}}$, which is much larger than the I_1^i which are of length $(\log \varepsilon)^4$. On the other hand, in the single resonant case recall that the I_1^i are of length $\lceil \log(1/\varepsilon) \rceil^2$, and the separation between the singular sites in $\mathcal{S}_0$ at least $4(\log \varepsilon)^2$ (but possibly much longer).

These long sections consisting entirely of regular lattice points between singular pairs in the double resonance case, resp. singular sites in the simple resonance case, allow us to iterate the resolvent identity similar to [Lemma 3.8](#). For general n , it is essential to use (4.5) up to level $n - 1$ in order to achieve this separation. See Appendix A in [23] for the details. $\square$

Proof of Theorem 4.1. For any $\theta \in \mathbb{T}$, by [Theorem 2.1](#) for spectrally a.e. $E \in \mathbb{R}$ there is a generalized eigenfunction $H(\theta)\psi = E\psi$ with at most linear growth. For any such E, ψ we claim that there exists $N = N(\theta, \psi) \geq 1$ so that all intervals $\Lambda_n = [-2\ell_n, 2\ell_n]$ are n -singular for (θ, E) provided $n \geq N$. If Λ_n is n -regular for infinitely many n , then by the Poisson formula (3.28) and by [Lemma 4.5](#) for any j and large n ,

$$|\psi(j)| \leq \sum_{(k', k) \in \partial \Lambda_n} |G_{\Lambda_n}(E)(j, k')| |\psi(k)| \leq C \varepsilon^{\frac{1}{2}(2\ell_n - |j|)} \ell_n$$

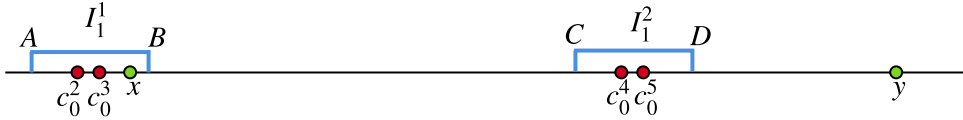


Fig. 7. A 1-regular interval A .

Taking the limit $n \rightarrow \infty$ yields $\psi \equiv 0$, whence our claim. Next, we claim that

$$A_n \cap I_n^i \neq \emptyset \text{ for some singular } I_n^i \quad (4.6)$$

for large n . Since A_n is not n -regular, we can pick some $c_0^j \in \mathcal{S}_0 \cap A_n$. By the recursive construction of singular intervals, see (4.4) there is the following dichotomy: either, there exists $1 \leq m \leq n$ with

$$c_0^j \in I_1^{j1} \subset \dots \subset I_m^{jm}$$

and I_m^{jm} is regular, or $m = n$ with I_n^{jn} singular. The first alternative cannot occur for every $c_0^j \in \mathcal{S}_0 \cap A_n$ by definition of regularity and so (4.6) holds. If $A'_n := [-\ell_{n+1}, \ell_{n+1}]$ contains another singular interval at level n , say I_n^j , then by (4.5) one has $m(c_n^i, c_n^j)^2 \lesssim \delta_n$. But $\|(c_n^i - c_n^j)\omega\| \lesssim \delta_n^{\frac{1}{2}}$ is impossible by the Diophantine condition whence

$$\|2\theta + (c_n^i + c_n^j)\omega\| \lesssim \delta_n^{\frac{1}{2}} \quad (4.7)$$

Given that there are at most $\lesssim \ell_{n+1}^2$ many choices of $c_n^i, c_n^j \in A'_n$, it follows that the measure of θ as in (4.7) is $\lesssim \ell_{n+1}^2 \delta_n^{\frac{1}{2}}$. This can be summed, whence by Borel–Cantelli there is a set $\mathcal{B}$ of measure 0 off of which for large n , A'_n contains a unique singular interval at level n . Furthermore, this singular interval has distance $< 3\ell_n$ from 0 and thus $[3\ell_n, \ell_{n+1}]$ and $[-\ell_{n+1}, -3\ell_n]$ are n -regular, for parameters (θ, E) with $\theta \in \mathbb{T} \setminus \mathcal{B}$. Lemma 4.5 and (3.28) conclude the proof. It is essential here that $\mathcal{B}$ does not depend on E , as evidenced by (4.7). $\square$

Proof. The remainder of this section is devoted to the proof of Lemma 4.4. We begin with the easier case of a simple resonance, i.e., $s_0 \geq 4(\log \varepsilon)^2$. Fix any $E_* \in \mathbb{R}$ with $[E_* - \delta_0, E_* + \delta_0] \cap [\min V, \max V] \neq \emptyset$, and some $\theta_* \in \mathbb{T}$. Then $\mathcal{S}_0 = \{c_0^i\}_{i=-\infty}^\infty$, and every c_0^i is contained in a unique level 1 interval I_1^i , with $|I_1^i| = \ell_1 = \lceil \log(1/\varepsilon) \rceil^2$. These are pairwise disjoint by construction, and they may be regular or singular. We discard the regular ones and only consider those $c_1^i = c_0^i$ for which I_1^i is singular. By the definitions,

$$\begin{aligned} |V(\theta_* + k\omega) - E_*| &> \delta_0 \quad \forall k \in I_1^i \setminus \{c_0^i\} \\ |V(\theta_* + c_0^i\omega) - E_*| &\leq \delta_0 \end{aligned} \quad (4.8)$$

Let $\{E_j^{i,1}(\theta)\}_{j=1}^{\ell_1}$ be the eigenvalue parameterizations (Rellich functions) of $H_{I_1^i}(\theta)$. By min–max, there exists $k(j, \theta) \in I_1^i$ so that

$$|V(\theta + k(j, \theta)\omega) - E_j^{i,1}(\theta)| \leq 2\varepsilon^2 \quad \forall 1 \leq j \leq \ell_1 \quad (4.9)$$

Fig. 8 exhibits² numerically computed Rellich functions for $\ell_1 = 7$ and the cosine potential. The graphs do not cross, but some of the gaps are too small to be visible. Subfigures (A) and (B) show how the gaps become wider with increasing ε . The figure demonstrates how we need to jump between different translates

² The graphs were produced by Yakir Forman at Yale.

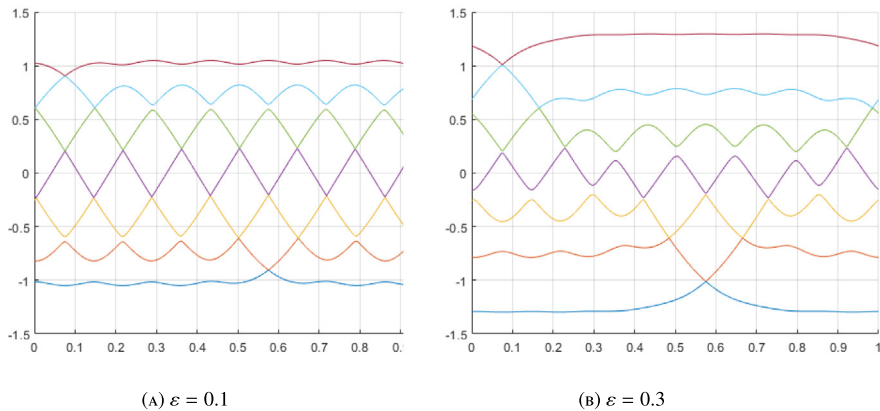


Fig. 8. Eigenvalue parameterizations for the cosine potential.

$V(\theta + k\omega)$ to approximate any given Rellich graph, hence $k(j, \theta)$, which are not unique near crossing points of V with its own translates by ω . Since $\delta_1 \ll \varepsilon^2 \ll \varepsilon$, (4.3) and (4.8), (4.9) imply that for all $|\theta - \theta_*| \ll \delta_0$

$$\begin{aligned} |V(\theta + c_0^i \omega) - E_*| &\ll \delta_0 \\ |V(\theta + k\omega) - E_*| &\gtrsim \delta_0 \quad \forall k \in I_1^i \setminus \{c_0^i\} \end{aligned} \quad (4.10)$$

with implied absolute constants (depending only on v, ω). We now claim that a normalized eigenfunction $\psi(\theta)$ associated with $H_{I_1^i}(\theta)\psi(\theta) = E_j^{i,1}(\theta)\psi(\theta)$ and $k(j, \theta) = c_1^i = c_0^i$ satisfies

$$\|P^\perp \psi(\theta)\|_{\ell^2(I_1^i)} \lesssim \varepsilon^2 \delta_0^{-1} \ll \varepsilon \quad (4.11)$$

where $P^\perp$ denotes the orthogonal projection onto all vectors perpendicular to $\delta_{c_0^i}$ in $\ell^2(I_1^i)$. Then

$$\begin{aligned} 0 &= \psi_{c_0^i}(\theta) P^\perp (H_{I_1^i}(\theta) - E_j^{i,1}(\theta)) \delta_{c_0^i} + P^\perp (H_{I_1^i}(\theta) - E_j^{i,1}(\theta)) P^\perp \psi(\theta) \\ &= \varepsilon^2 \psi_{c_0^i}(\theta) (\delta_{c_0^i-1} + \delta_{c_0^i+1}) + P^\perp (H_{I_1^i}(\theta) - E_j^{i,1}(\theta)) P^\perp \psi(\theta) \end{aligned}$$

Here and below we use δ both for the resonance width and in the Dirac sense, without any danger of confusion. By (4.9) and (4.10),

$$\|[P^\perp (H_{I_1^i}(\theta) - E_j^{i,1}(\theta)) P^\perp]^{-1}\|_{\ell^2(I_1^i \setminus \{c_0^i\})} \lesssim \delta_0^{-1} \quad (4.12)$$

which implies (4.11) and

$$|E_j^{i,1}(\theta) - V(\theta + c_0^i \omega)| = |\langle H_{I_1^i}(\theta) \psi(\theta), \psi(\theta) \rangle - V(\theta + c_0^i \omega)| \lesssim \varepsilon^2 \delta_0^{-1} \quad (4.13)$$

By first order eigenvalue perturbation (Feynman formula), writing $E = E_j^{i,1}$ and $V(\theta)$ for the multiplication operator by the potential,

$$|E'(\theta) - V'(\theta + c_0^i \omega)| = |\langle V'(\theta) \psi(\theta), \psi(\theta) \rangle - V'(\theta + c_0^i \omega)| \lesssim \varepsilon^2 \delta_0^{-1} \quad (4.14)$$

and by the second order perturbation formula, with $G^\perp$ being the resolvent on the left-hand side of (4.12),

$$\begin{aligned} |E''(\theta) - V''(\theta + c_0^i \omega)| &= |\langle V''(\theta) \psi(\theta), \psi(\theta) \rangle - V''(\theta + c_0^i \omega) - 2\langle \psi(\theta), V'(\theta) G^\perp (E(\theta)) V'(\theta) \psi(\theta) \rangle| \\ &\lesssim \varepsilon^2 \delta_0^{-1} + (\varepsilon^2 \delta_0)^2 \delta_0^{-1} \lesssim \varepsilon^2 \delta_0^{-1} \ll \varepsilon \end{aligned} \quad (4.15)$$

The estimates (4.13)–(4.15) hold for $|\theta - \theta_*| \ll \delta_0$. We conclude from (4.14), (4.15) that

$$\min_{|\theta - \theta_*| \ll \delta_0} (|E'(\theta)| + |E''(\theta)|) \gtrsim 1 \quad \forall \theta_* \in \mathbb{T}, \quad (4.16)$$

Recall that $E(\theta) = E_j^{i,1}(\theta)$ depends on θ_* and E_* , where the latter is chosen so that $\mathcal{S}_0 \neq \emptyset$. The constant in (4.16) is uniform in θ_*, E_* . The reader is invited to compare (4.16) to Fig. 8. Now suppose we have two distinct singular intervals I_1^i and I_1^j relative to (θ_*, E_*) . Then the previous analysis applies to both Rellich functions $E(\theta), \tilde{E}(\theta)$ defined for $|\theta - \theta_*| \ll \delta_0$ characterized by

$$\begin{aligned} \text{spec}(H_{I_1^i}(\theta)) \cap [E_* - \delta_0/2, E_* - \delta_0/2] &= \{E(\theta)\}, \\ \text{spec}(H_{I_1^j}(\theta)) \cap [E_* - \delta_0/2, E_* - \delta_0/2] &= \{\tilde{E}(\theta)\} \end{aligned} \quad (4.17)$$

By (4.3) we have $|E(\theta_*) - E_*| \leq \delta_1$, $|\tilde{E}(\theta_*) - E_*| \leq \delta_1$. By (4.9) we have

$$|V(\theta_* + c_0^i \omega) - E(\theta_*)| \leq 2\varepsilon^2, \quad |V(\theta_* + c_0^j \omega) - \tilde{E}(\theta_*)| \leq 2\varepsilon^2$$

whence

$$|V(\theta_* + c_0^i \omega) - V(\theta_* + c_0^j \omega)| \leq 4\varepsilon^2 + 2\delta_1 \leq 5\varepsilon^2$$

We showed in Lemma 4.2 that this implies

$$m(c_0^i, c_0^j) \lesssim \varepsilon \ll \delta_0. \quad (4.18)$$

Next, we improve this estimate to $m(c_0^i, c_0^j) \lesssim \delta_1$. Suppose (4.18) means $\|(c_0^i - c_0^j)\omega\| \ll \delta_0$. By (3.2) one has

$$\text{spec}(H_{I_1^i}(\theta)) = \text{spec}(H_{I_1^j}(\theta + (c_0^i - c_0^j)\omega)) \quad (4.19)$$

which, combined with (4.17) implies that $\tilde{E}(\theta) = E(\theta + (c_0^j - c_0^i)\omega)$ for all $|\theta - \theta_*| \ll \delta_0$. This finally implies that

$$|E(\theta_*) - E(\theta_* + (c_0^j - c_0^i)\omega)| \leq 2\delta_1$$

From (4.16) we obtain $\|(c_0^j - c_0^i)\omega\|^2 \lesssim \delta_1$. On the other hand, if (4.18) means $\|2\theta_* + (c_0^i + c_0^j)\omega\| \ll \delta_0$, then we have

$$\tilde{E}(\theta) = E(2\theta_* - \theta - \theta_{**}), \quad \theta_{**} := 2\theta_* + (c_0^i + c_0^j)\omega \quad (4.20)$$

for all $|\theta - \theta_*| \ll \delta_0$. In terms of Fig. 3 this corresponds to $E(\theta)$ being approximated by V over J_1 , whereas $\tilde{E}(\theta)$ is approximated by V over J_2 . Setting $\theta = \theta_*$ in (4.20), we find that

$$|E(\theta_*) - E(\theta_* - \theta_{**})| \leq 2\delta_1$$

which implies $\|\theta_{**}\|^2 \lesssim \delta_1$. We have thus proved Lemma 4.4 for simple resonances. Figs. 9 and 10 depict eigenfunctions on finite volume 200 computed with MATHEMATICA (the code is included in the appendix for the reader to experiment for themselves, for example by changing θ or trying rational ω). The choice of θ was made as a crossing point of V with one of its translates by a multiple of ω , since double resonances occur near those points. The first eigenfunction shown in the upper left of Figs. 9 shows the case of a simple resonance, whereas the second and third are more complicated — they exhibit a main peak with smaller ones due to resonances at later stages of the induction. On the other hand, Fig. 10 exhibits double resonances

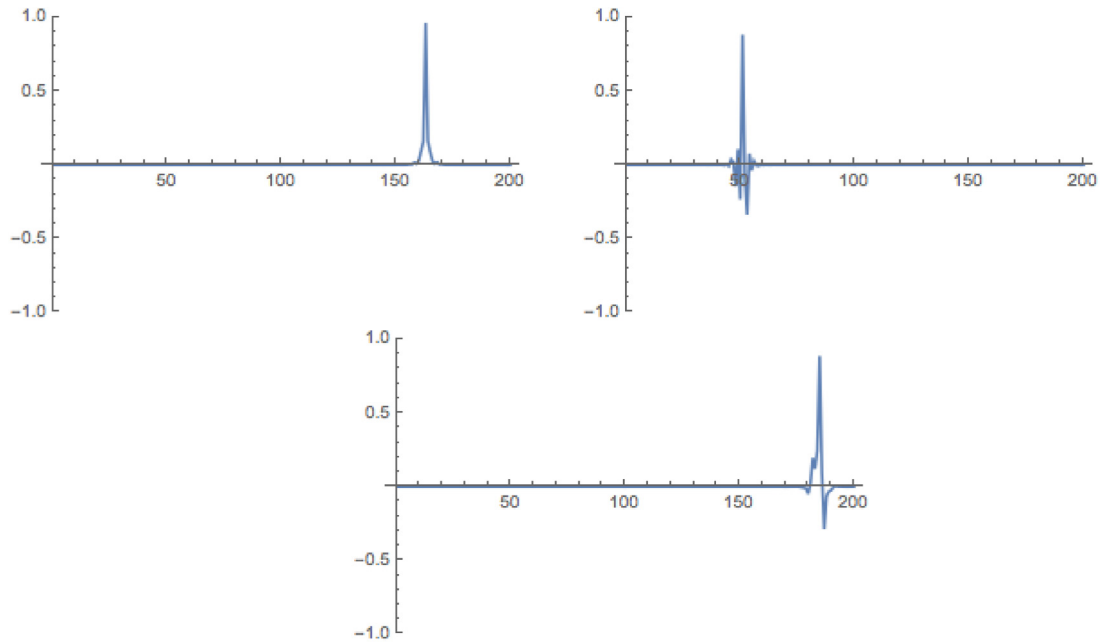


Fig. 9. Numerically computed eigenfunctions, $\varepsilon = 0.3$, $\omega = \sqrt{2}$, $\theta = -17\omega/2$.

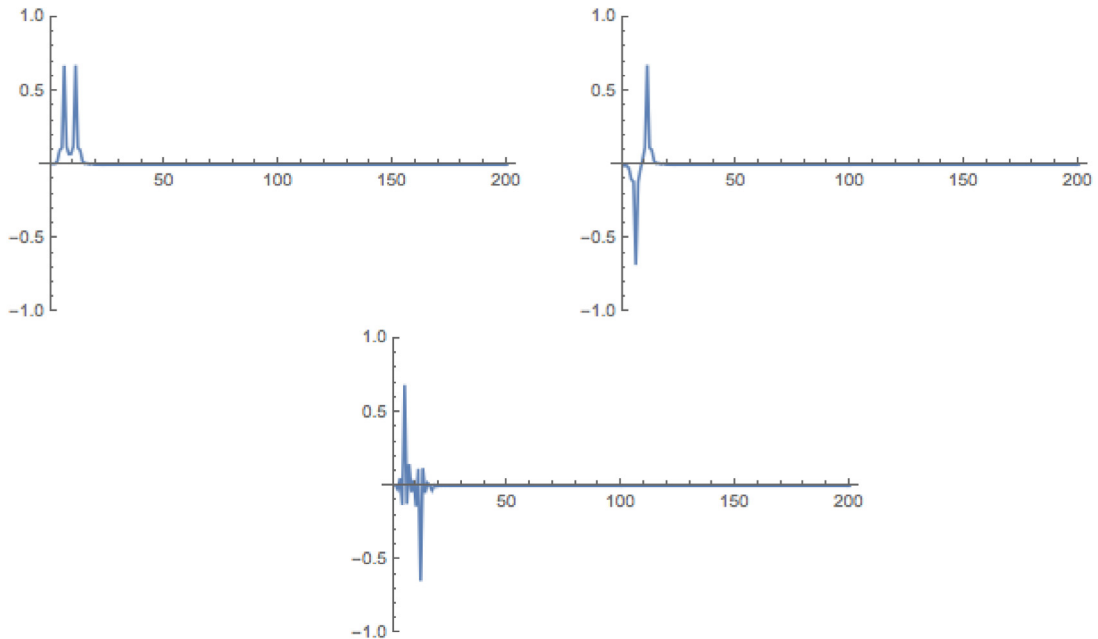


Fig. 10. Numerically computed eigenfunctions, $\varepsilon = 0.3$, $\omega = \sqrt{2}$, $\theta = -17\omega/2$.

quite clearly (with the bottom eigenfunction exhibiting a more complicated structure). The reader should note the distinct distribution of the ℓ^2 -mass which is quite apparent on the y -axes of these figures.

We now prove [Lemma 4.4](#) for double resonances. Let I_1^i be singular (as the red interval on the right-hand side of [Fig. 5](#)), centered at $c_1^i = \frac{1}{2}(c_0^{2i} + c_0^{2i+1}) \in \frac{1}{2}\mathbb{Z}$. As a side remark, suppose that $c_1^i \in \frac{1}{2} + \mathbb{Z}$. Then all $c_1^j \in \frac{1}{2} + \mathbb{Z}$ due to $c_0^{2j+1} - c_0^{2j} = \text{const.}$ for all j (since we passed to $\bar{S}_0$). At the next levels $n = 2, 3, \dots, N$ we will encounter only simple resonances, and so all $c_n^k \in \frac{1}{2} + \mathbb{Z}$ for all these n . If we then encounter a double

resonance at $N + 1$, it implies that $c_{N+1}^k \in \frac{1}{2}\mathbb{Z}$, and the pattern repeats itself. Continuing with the main argument, one then has

$$\begin{aligned} |V(\theta_* + c_0^{2i}\omega) - E_*| &\lesssim \delta_0, \quad |V(\theta_* + c_0^{2i+1}\omega) - E_*| \lesssim \delta_0 \\ |V(\theta_* + k\omega) - E_*| &\gg \delta_0^{\frac{1}{2}} \quad \forall k \in I_1^i \setminus \{c_0^{2i}, c_0^{2i+1}\} \end{aligned}$$

where the second line follows from the Diophantine condition since $I_1^i = \lceil \log(1/\varepsilon) \rceil^4$ (one can choose a larger lower bound such as δ_0^a for any fixed $0 < a \leq \frac{1}{2}$ at the expense of making ε smaller). By (4.9),

$$\text{spec}(H_{I_1^i}(\theta)) \cap [E_* - \delta_0^{\frac{1}{2}}, E_* + \delta_0^{\frac{1}{2}}] = \{E(\theta), \tilde{E}(\theta)\} \quad \forall |\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}} \quad (4.21)$$

with $E > \tilde{E}$. By the same type of argument as in the simple resonant case, cf. (4.11), (4.12), we see that the normalized eigenfunctions of $H_{I_1^i}(\theta)$ associated with E , resp. $\tilde{E}$, are

$$\begin{aligned} \psi &= A\delta_{c_0^{2i}} + B\delta_{c_0^{2i+1}} + O(\varepsilon^2\delta_0^{-\frac{1}{2}}) \\ \tilde{\psi} &= -B\delta_{c_0^{2i}} + A\delta_{c_0^{2i+1}} + O(\varepsilon^2\delta_0^{-\frac{1}{2}}) \end{aligned} \quad (4.22)$$

uniformly on $|\theta - \theta_*| \leq \delta_0^{\frac{1}{2}}$ with $A^2 + B^2 = 1$. In place of (4.12) we have

$$\| [P^\perp(H_{I_1^i}(\theta) - E_*)P^\perp]^{-1} \|_{\ell^2(I_1^i \setminus \{c_0^{2i}, c_0^{2i+1}\})} \lesssim \delta_0^{-\frac{1}{2}} \quad (4.23)$$

where $P^\perp$ is the orthogonal projection perpendicular to $\text{span}(\delta_{c_0^{2i}}, \delta_{c_0^{2i+1}})$ in $\ell^2(I_1^i)$. The eigenvalues at level 0 associated with the points c_0^{2i}, c_0^{2i+1} are, resp., $E_0^{2i}(\theta) := V(\theta + c_0^{2i}\omega)$, $E_0^{2i+1}(\theta) := V(\theta + c_0^{2i+1}\omega)$. This terminology is justified by the relations

$$\langle H(\theta)\delta_{c_0^{2i}}, \delta_{c_0^{2i}} \rangle = E_0^{2i}(\theta), \quad \langle H(\theta)\delta_{c_0^{2i+1}}, \delta_{c_0^{2i+1}} \rangle = E_0^{2i+1}(\theta)$$

By Lemma 4.2, $\|\theta_{**}\| \lesssim \delta_0^{\frac{1}{2}}$ with $\theta_{**} = 2\theta_* + (c_0^{2i} + c_0^{2i+1})\omega = 2(\theta_* + c_1^i\omega)$. It follows that either (a) $\|\theta_* + c_1^i\omega\| \lesssim \delta_0^{\frac{1}{2}}$ or (b) $\|\theta_* + \frac{1}{2} + c_1^i\omega\| \lesssim \delta_0^{\frac{1}{2}}$. These relations show that the unique solution of $E_0^{2i}(\theta) = E_0^{2i+1}(\theta)$ on $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$ is either (a) $\theta_s = -c_1^i\omega$ or (b) $\theta_s = \frac{1}{2} - c_1^i\omega$ (henceforth, θ_s will mean either of these whichever applies). These identities are a restatement of V being symmetric both (a) around 0 and (b) around $\frac{1}{2}$. Furthermore, one has

$$E_0^{2i}(\theta + \theta_s) = E_0^{2i+1}(-\theta + \theta_s)$$

whence $\partial_\theta E_0^{2i}(\theta_s) = -\partial_\theta E_0^{2i+1}(\theta_s)$. The configuration associated with a double resonance is shown in Fig. 11. Not only do the segments of the V -graphs (i.e., E_0^{2i} and E_0^{2i+1}) intersect at θ_s , but $E, \tilde{E}$ have their critical point at θ_s within the interval $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$. Indeed,

$$H_{I_1^i}(\theta + \theta_s) = UH_{I_1^i}(-\theta + \theta_s)U$$

where U is the reflection on $\mathbb{Z}$ about $c_1^i \in \frac{1}{2}\mathbb{Z}$. In particular, the eigenvalues are the same. In fact, using (4.21) one concludes that

$$E(\theta + \theta_s) = E(-\theta + \theta_s), \quad \tilde{E}(\theta + \theta_s) = \tilde{E}(-\theta + \theta_s) \quad \forall |\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}} \quad (4.24)$$

whence $E'(\theta_s) = \tilde{E}'(\theta_s) = 0$.

Next, we establish the lower bound

$$E(\theta) - \tilde{E}(\theta) > (c\varepsilon)^{5\ell_1^{1/2}} \gg \delta_1 = \varepsilon^{\ell_1^{2/3}} = \varepsilon^{\lceil \log \varepsilon \rceil^{8/3}} \quad (4.25)$$

which follows immediately from this separation lemma, see [23, Lemma 4.1]. This spectral gap is much larger than the resonance width δ_1 .

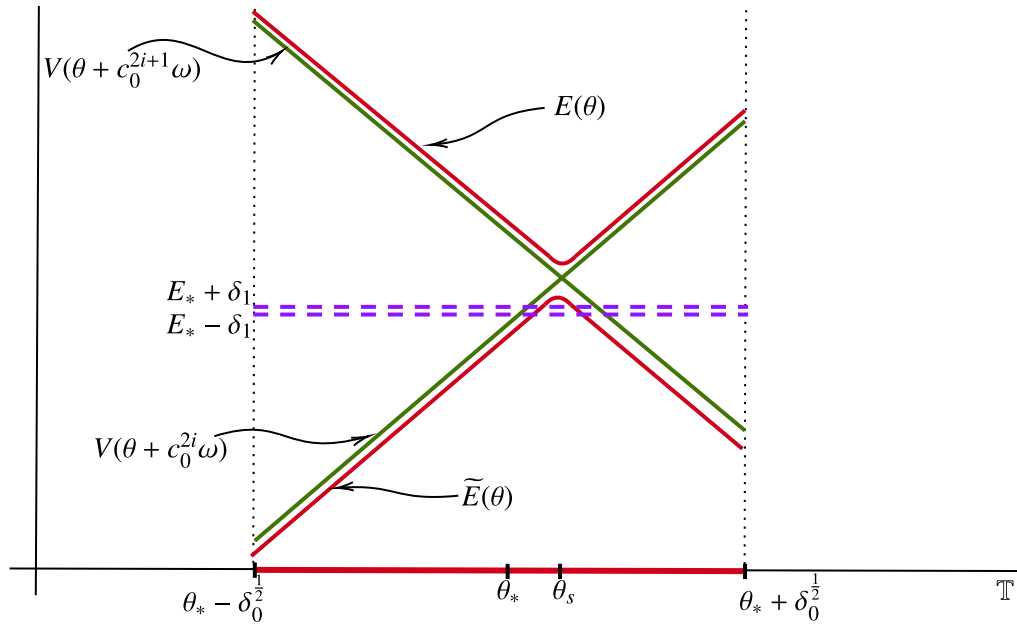


Fig. 11. Crossing graphs and double resonance.

Lemma 4.6. Let $H_\Lambda \psi_j = E_j \psi_j$, $j = 1, 2$ with nontrivial ψ_j and $E_1 \neq E_2$. If $\|\psi_j\|_{\ell^2(\Lambda_0)} \geq \frac{1}{2} \|\psi_j\|_{\ell^2(\Lambda)}$ for $j = 1, 2$ with $\Lambda_0 \subset \Lambda$ and $|\Lambda_0| \geq 2$, then

$$|E_1 - E_2| \geq (c_1 \varepsilon)^{|\Lambda_0|} (\varepsilon^{-2} + |\Lambda_0|)^{-1} \quad (4.26)$$

with a constant $c_1 = c_1(V) > 0$.

Proof. Let $\Lambda_0 = [n_0 - \ell_0, n_0 + \ell_0]$ or $\Lambda_0 = [n_0 - \ell_0, n_0 + \ell_0 - 1]$. By assumption, $\ell_0 \geq 1$. Normalize $\psi_j(n_0 - 1)^2 + \psi_j(n_0)^2 = 1$ for $j = 1, 2$. Setting $\tilde{\psi}_1(n) = \psi_1(n)$ if $n \in \Lambda$, $n \geq n_0$ and $\tilde{\psi}_1(n) = -\psi_1(n)$ if $n \in \Lambda$, $n < n_0$ one obtains from considering $\langle H_\Lambda \tilde{\psi}_1, \psi_2 \rangle = \langle \tilde{\psi}_1, H_\Lambda \psi_2 \rangle$ that

$$\begin{aligned} 2\varepsilon^2 |\vec{v}_1 \wedge \vec{v}_2| &= 2\varepsilon^2 |\psi_1(n_0)\psi_2(n_0 - 1) - \psi_2(n_0)\psi_1(n_0 - 1)| = \\ &\leq 2|E_1 - E_2| (\|\psi_1\|_2 \|\psi_2\|_2)^{\frac{1}{2}} (\|\psi_1\|_{\ell^2(\Lambda_0)} \|\psi_2\|_{\ell^2(\Lambda_0)})^{\frac{1}{2}} \\ &\leq 2(C\varepsilon^{-2})^{\ell_0} |E_1 - E_2| (\|\psi_1\|_2 \|\psi_2\|_2)^{\frac{1}{2}} \end{aligned} \quad (4.27)$$

where $\vec{v}_j = \begin{pmatrix} \psi_j(n_0) \\ \psi_j(n_0 - 1) \end{pmatrix}$ and $C = C(V)$. The final estimate is obtained from the transfer matrix representation of the eigenfunctions, viz. for $n \geq n_0 + 1$

$$\begin{pmatrix} \psi_j(n) \\ \psi_j(n - 1) \end{pmatrix} = \prod_{k=n_0}^{n-1} \begin{bmatrix} \varepsilon^{-2}(V_k - E_j) & -1 \\ 1 & 0 \end{bmatrix} \vec{v}_j$$

and for $n \leq n_0 - 1$

$$\begin{pmatrix} \psi_j(n - 1) \\ \psi_j(n) \end{pmatrix} = \prod_{k=n_0-1}^n \begin{bmatrix} \varepsilon^{-2}(V_k - E_j) & -1 \\ 1 & 0 \end{bmatrix}^{-1} \vec{v}_j$$

with $V_k = V(\theta + k\omega)$. On the one hand, with $B = C\varepsilon^{-2}$, and using that $\|\vec{v}_1 - \vec{v}_2\|_2 = 2|\vec{v}_1 \wedge \vec{v}_2|$,

$$\begin{aligned} \|\psi_1 - \psi_2\|_{\ell^2(\Lambda_0)} &\leq 2B^{\ell_0} |\vec{v}_1 \wedge \vec{v}_2| + \ell_0 B^{\ell_0} |E_1 - E_2| \\ &\leq 2\varepsilon^{-2} B^{\ell_0} |E_1 - E_2| (\|\psi_1\|_2 \|\psi_2\|_2)^{\frac{1}{2}} + \ell_0 B^{\ell_0} |E_1 - E_2| \end{aligned}$$

On the other hand, with $a_j := \|\psi_j\|_{\ell^2(\Lambda)}$,

$$\begin{aligned}\sqrt{a_1^2 + a_2^2} &\leq \frac{1}{2}(a_1 + a_2) + \|\psi_1 - \psi_2\|_{\ell^2(\Lambda_0)} \\ &\leq \frac{1}{2}(a_1 + a_2) + B^{\ell_0}(B\sqrt{a_1 a_2} + \ell_0)|E_1 - E_2| \\ &\leq \frac{1}{2}(a_1 + a_2) + B^{\ell_0}(B + \ell_0)|E_1 - E_2|\sqrt{a_1 a_2}\end{aligned}$$

If $|E_1 - E_2| < \frac{1}{4}B^{-\ell_0}(B + \ell_0)^{-1}$, then $\sqrt{a_1^2 + a_2^2} < \frac{5}{8}(a_1 + a_2)$ which is impossible. Adjusting the constants one obtains (4.26). $\square$

The critical points of V are $\theta = 0$ and $\theta = \frac{1}{2}$. We claim that $\min(\|\theta_*\|, \|\theta_* - \frac{1}{2}\|) \geq K\delta_0^{\frac{1}{2}}$ where K is any large constant, to be fixed below (as always, provided ε is small enough). This is immediate from the Diophantine condition due to $s_0 \leq 4(\log \varepsilon)^2$, cf. Fig. 3. In particular, $|V'(\theta)| \gg \delta_0^{\frac{1}{2}}$ on the interval $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$. By first order eigenvalue perturbation and (4.22), uniformly on this interval

$$\partial_\theta E(\theta) = \langle \psi(\theta), V'(\theta)\psi(\theta) \rangle = A^2(\theta)\partial_\theta E_0^{2i}(\theta) + B^2(\theta)\partial_\theta E_0^{2i+1}(\theta) + O(\varepsilon^2\delta_0^{-\frac{1}{2}})$$

where by the preceding $|\partial_\theta E_0^{2i}(\theta)| \gg \delta_0^{\frac{1}{2}}$ and $|\partial_\theta E_0^{2i+1}(\theta)| \gg \delta_0^{\frac{1}{2}}$. Setting $\theta = \theta_s$ it follows that $A^2(\theta_s) - B^2(\theta_s) = O(\varepsilon^2\delta_0^{-1}) = O(\varepsilon)$. Due to $A^2 + B^2 = 1$, $|A(\theta_s)|^2 = 1/2 - O(\varepsilon)$ and $|B(\theta_s)|^2 = 1/2 - O(\varepsilon)$. In fact, the same argument shows that $|A(\theta)| \simeq |B(\theta)| \simeq 1$ for all $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$ with $|\partial_\theta E(\theta)| \lesssim \delta_0^{\frac{1}{2}}$.

Using this property we can now establish closeness of all eigenvalues. In fact, $H_\Lambda(\theta)\psi(\theta) = E(\theta)\psi(\theta)$ and $H_\Lambda(\theta)\tilde{\psi}(\theta) = \tilde{E}(\theta)\tilde{\psi}(\theta)$ in combination with (4.22) imply that

$$E_0^{2i}(\theta) - E(\theta) = O(\varepsilon^2\delta_0^{-\frac{1}{2}}), \quad E_0^{2i+1}(\theta) - E(\theta) = O(\varepsilon^2\delta_0^{-\frac{1}{2}}) \quad (4.28)$$

and the same for $\tilde{E}$. In particular,

$$|E(\theta) - \tilde{E}(\theta)| \lesssim \varepsilon^2\delta_0^{-\frac{1}{2}} \quad (4.29)$$

for all $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$ with $|\partial_\theta E(\theta)| \lesssim \delta_0^{\frac{1}{2}}$. The final step in our analysis is to establish a lower bound on $|\partial_\theta^2 E(\theta)|$ and $|\partial_\theta^2 \tilde{E}(\theta)|$ for those θ . This hinges on the second order perturbation formulas (suppressing θ as argument)

$$\begin{aligned}\partial_\theta^2 E &= \langle \psi, V''\psi \rangle - 2\langle V'\psi, G(E)^\perp V'\psi \rangle \\ \partial_\theta^2 \tilde{E} &= \langle \tilde{\psi}, V''\tilde{\psi} \rangle - 2\langle V'\tilde{\psi}, G(\tilde{E})^\perp V'\tilde{\psi} \rangle\end{aligned}$$

on $\ell^2(I_1^i)$ with $G(E)^\perp = [P_\psi^\perp(H_{I_1^i} - E)P_\psi^\perp]^{-1}$ in $P_\psi^\perp\ell^2(I_1^i)$ and $P_\psi^\perp$ being the orthogonal projection onto the complement of ψ in $\ell^2(I_1^i)$. Analogous comments apply $G(\tilde{E})^\perp$ which is the resolvent orthogonal to $\tilde{\psi}$. We now write

$$\langle V'\psi, G(E)^\perp V'\psi \rangle = \frac{\langle \psi, V'\tilde{\psi} \rangle^2}{\tilde{E} - E} + \langle V'\psi, G(E)^{\perp\perp} V'\psi \rangle$$

where $G(E)^{\perp\perp} = P_\psi^\perp[P_\psi^\perp(H_{I_1^i} - E)P_\psi^\perp]^{-1}P_\psi^\perp$. By (4.21), $\|G(E)^{\perp\perp}\| \lesssim \delta_0^{-\frac{1}{2}}$. On the other hand, by (4.22)

$$\langle \psi, V'\tilde{\psi} \rangle = -2AB\partial_\theta E_0^{2i}(\theta_s) + O(\delta_0^{\frac{1}{2}})$$

whence from $|AB| \simeq 1$,

$$|\langle \psi, V'\tilde{\psi} \rangle| \gg \delta_0^{\frac{1}{2}}$$

Combining this with (4.29) we obtain

$$|\langle V'\psi, G(E)^\perp V'\psi \rangle| \gg \varepsilon^{-2} \delta_0^{\frac{3}{2}} - O(\delta_0^{-\frac{1}{2}}) \gg \varepsilon^{-2} \delta_0^{\frac{3}{2}} \gg 1$$

Since $|\langle \psi, V''\psi \rangle| \lesssim 1$, it follows that

$$|\partial_\theta^2 E(\theta)| \gg \varepsilon^{-2} \delta_0^{\frac{3}{2}}, \quad |\partial_\theta^2 \tilde{E}(\theta)| \gg \varepsilon^{-2} \delta_0^{\frac{3}{2}} \quad (4.30)$$

for all $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$ with $|\partial_\theta E(\theta)| \lesssim \delta_0^{\frac{1}{2}}$. The exact same analysis applies to $\tilde{E}$. To summarize, these are the main points concerning double resonances at level 1.

- $s_0 \leq (\log \varepsilon)^2$ and the level-0 singular sites are $\mathcal{S}_0 = \{c_0^j\}_{j=-\infty}^\infty$ with $c_0^{2i+1} - c_0^{2i} = s_0$ for all i . We have $|V(\theta + k\omega) - E_*| \lesssim \delta_0^{\frac{1}{2}}$ for $k \in \{c_0^{2i}, c_0^{2i+1}\}$, $|V(\theta + k\omega) - E_*| \gg \delta_0^{\frac{1}{2}}$ for all $k \in I_1^i \setminus \{c_0^{2i}, c_0^{2i+1}\}$, both for all $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$. Here $I_1^i \subset \mathbb{Z}$, $|I_1^i| \simeq (\log \varepsilon)^4$, centered at $c_1^i \in \frac{1}{2}\mathbb{Z}$ and $\text{dist}(\text{spec}(H_{I_1^i}(\theta_*)), E_*) \lesssim \delta_1 = \varepsilon^{\ell_1^2/3}$.
- $\text{spec}(H_{I_1^i}(\theta)) \cap [E_* - \delta_0^{\frac{1}{2}}, E_* + \delta_0^{\frac{1}{2}}] = \{E(\theta), \tilde{E}(\theta)\}$ (with $E > \tilde{E}$) for these θ , with all other eigenvalues being separated from E_* by $\gg \delta_0^{\frac{1}{2}}$. From the level-0 estimate $m(c_0^{2i}, c_0^{2i+1}) \lesssim \delta_0$, either $\theta_s = -c_1^i \omega$ or $\theta_s = \frac{1}{2} - c_1^i \omega$ satisfy $\|\theta_s - \theta_*\| \lesssim \delta_0^{\frac{1}{2}}$ and the unique critical points of $E, \tilde{E}$ in this interval are at θ_s . There is a spectral gap of size $E(\theta) - \tilde{E}(\theta) > \varepsilon^{5\ell_1^2}$.
- For every $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$ one has either both $|\partial_\theta E(\theta)| \gg \delta_0^{\frac{1}{2}}$ and $|\partial_\theta \tilde{E}(\theta)| \gg \delta_0^{\frac{1}{2}}$ (large slopes), or both $|\partial_\theta E(\theta)| \lesssim \delta_0^{\frac{1}{2}}$ and $|\partial_\theta \tilde{E}(\theta)| \lesssim \delta_0^{\frac{1}{2}}$ (small slopes). This follows from $|\partial_\theta E_0^{2i}(\theta_s)| \gg \delta_0^{\frac{1}{2}}$, and the first order eigenvalue perturbation formulas

$$\begin{aligned} \partial E(\theta) &= (A^2(\theta) - B^2(\theta))\partial_\theta E_0^{2i}(\theta_s) + O(\delta_0^{\frac{1}{2}}) \\ \partial \tilde{E}(\theta) &= (-A^2(\theta) + B^2(\theta))\partial_\theta E_0^{2i}(\theta_s) + O(\delta_0^{\frac{1}{2}}) \end{aligned}$$

for all $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$, cf. (4.22).

- If the small slope alternative occurs, then $|A(\theta)| \simeq |B(\theta)| \simeq 1$ and (4.28) holds for both E and $\tilde{E}$. In particular, the spectral gap is small as in (4.29), and the second derivatives are large and $\gg \delta_0^{-\frac{1}{2}}$, see (4.30). This means that the intervals of small slopes around the critical points at θ_s are of size $\ll \delta_0$.
- Fig. 11 depicts the situation for a double resonance: E reaches its minimum, resp. $\tilde{E}$ its maximum, at θ_s . The spectral gap is the smallest at this point and the quantitative estimates above hold. In particular, this gap is much larger than δ_1 , whence exactly one of E or $\tilde{E}$ achieve the resonance condition (4.3) at θ_* .

To conclude the proof of Lemma 4.4 we apply this description to two such level 1 intervals, say I_1^i and I_1^j . Because of the double resonance assumption, we have

$$\|1/2 + c_1^i \omega\| + \|1/2 + c_1^j \omega\| \lesssim \delta_0^{\frac{1}{2}}$$

which implies that $m(c_1^i, c_1^j) = \|(c_1^i - c_1^j)\omega\| \lesssim \delta_0^{\frac{1}{2}}$. As in the single resonance case, cf. (4.19), for all $|\theta - \theta_*| \lesssim \delta_0^{\frac{1}{2}}$

$$H_{I_1^i}(\theta) = H_{I_1^j}(\theta + (c_1^i - c_1^j)\omega) \text{ and } E_1^i(\theta) = E_1^j(\theta + (c_1^i - c_1^j)\omega), \quad \tilde{E}_1^i(\theta) = \tilde{E}_1^j(\theta + (c_1^i - c_1^j)\omega)$$

Finally, by (4.3), either

$$|E_1^j(\theta_*) - E_1^j(\theta_* + (c_1^i - c_1^j)\omega)| \lesssim \delta_1 \text{ or } |\tilde{E}_1^j(\theta_*) - \tilde{E}_1^j(\theta_* + (c_1^i - c_1^j)\omega)| \lesssim \delta_1$$

By the bounds derived above on the first and second derivatives on E_1^i etc. and elementary calculus, we finally conclude that $m(c_1^i, c_1^j) \lesssim \delta_1^{\frac{1}{2}}$. Indeed, in the large slopes case, $m(c_1^i, c_1^j) = \|(c_1^i - c_1^j)\omega\| \lesssim \delta_0^{-\frac{1}{2}} \delta_1$, whereas in the small slopes case, $m(c_1^i, c_1^j) \lesssim \delta_0^{\frac{1}{4}} \delta_1^{\frac{1}{2}}$. This is slightly better than what Lemma 4.4 claims, and we are done. $\square$

The full induction needed to establish (4.5) follows these exact same lines with no essentially new ideas needed. The reader can either convince themselves of this fact, or consult [23]. Note, however, that Lemma 5.2 in loc. cit. erroneously sets $\theta_s = -c_m^i \alpha$ forgetting the case (b) above in which $1/2$ has to be added. This is a systematic oversight in Section 5 in that paper which is rooted in a false identity at the conclusion of the proof of Lemma 5.3: $2\|\theta\| = \|2\theta\|$ for the metric on $\mathbb{T}$.

It seems very difficult to approach quasi-periodic localization in more general settings by relying on eigenvalue parameterization, as we did in this section.

4.2. The work of Forman and VandenBoom: dropping evenness of V

We will now discuss the highly challenging task of implementing some version of the Fröhlich–Spencer–Wittwer proof strategy without the symmetry assumption on the potential. This has recently been accomplished by Forman and VandenBoom, see [21]. For the relation of this section with Sinai’s classical work [46] we refer the reader to the introduction of [21]. We present a sketch³ of the proof of their result.

Theorem 4.7. *Let $V \in C^2(\mathbb{T})$ have exactly two nondegenerate critical points. Define*

$$H_\varepsilon(\theta) = \varepsilon^2 \Delta_{\mathbb{Z}} + V_\theta, \quad V_\theta(n) = V(T_\omega^n \theta) \quad \forall n \in \mathbb{Z} \quad (4.31)$$

where $\omega \in \mathbb{T}$ is Diophantine, viz. $\|n\omega\| \geq b_0 n^{-2}$ for all $n \geq 1$ with some $b_0 > 0$. There exists $\varepsilon_0(b_0, V)$ such that for all $0 < \varepsilon \leq \varepsilon_0$ the operators $H_{\theta, \varepsilon}$ exhibit Anderson localization for a.e. $\theta \in \mathbb{T}$.

This is precisely the result of Fröhlich, Spencer, and Wittwer without the evenness assumption, and we will make frequent references to the proof of that result. See the previous section.

As in the symmetric case, we can define singular sites S_0 relative to θ_*, E_* . However, the $m(k, \ell)$ function is no longer useful, as $\|2\theta_* + (k + \ell)\omega\|$ is no longer small if $T_\omega^k \theta_*$ and $T_\omega^\ell \theta_*$ fall into different connected components of $V^{-1}([E_* - \delta_0, E_* + \delta_0])$. Without symmetry, no such function m can be defined to be independent of E_* .

Instead, we divide the energy axis into several overlapping intervals, and we construct a collection $\mathcal{E}_1$ of well-separated Rellich functions of certain Dirichlet restrictions of H whose domains cover the circle $\mathbb{T}$ with the same structural properties as $\mathbf{E}_0$, cf. Fig. 12. We choose an initial interval length $\ell_1^{(1)}$ and consider energy regions of size $\mathcal{O}((\ell_1^{(1)})^{-16})$. Each energy region can be characterized as double-resonant, if it contains some E_n which satisfies $E_n = V(\theta_n) = V(\theta_n + n\omega)$ for some $\theta \in \mathbb{T}$ and $|n| \leq \ell_1^{(1)}$, or simple-resonant if it does not. Each function $\mathbf{E}_1 \in \mathcal{E}_1$ is a Rellich function of H_{A_1} , where $A_1 \subset \mathbb{Z}$ is an interval of length $\ell_1^{(1)}$ if the energy region is simple-resonant, or $\ell_1^{(2)} \approx (\ell_1^{(1)})^2$ if the energy region is double-resonant. The singular intervals are then characterized by

$$S_1 = \{A_1 + m \mid m \in \mathbb{Z}, |\mathbf{E}_1(\theta_* + m\omega) - E_*| < \delta_1\}$$

where $\mathbf{E}_1 \in \mathcal{E}_1$ is the Rellich function defined in the energy region containing E_* , and δ_1 is defined as Fröhlich, Spencer, and Wittwer define it.

³ The remainder of this subsection was written by Forman and VandenBoom.

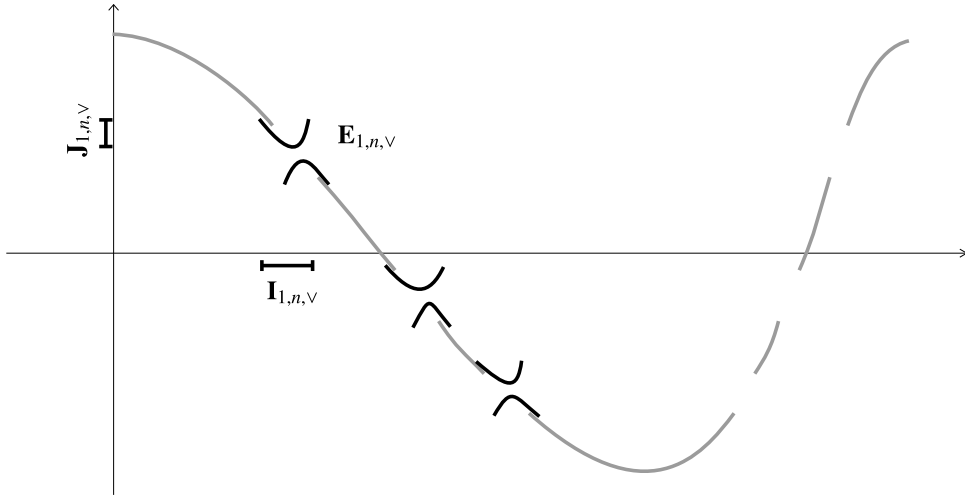


Fig. 12. A cartoon output of the first inductive step: a collection $\mathcal{E}_1$ of Rellich functions of various Dirichlet restrictions of H , whose domains (and their relevant translates) cover the circle $\mathbb{T}$. The curves in black come from double resonances, and the curves in gray are simple resonant.

Assuming the constructed Rellich functions satisfy a Morse condition, maintain two monotonicity intervals, and are well-separated from other Rellich functions on the same domain (i.e., we have an upper bound on $\|[P^\perp(H_{A_1} - \mathbf{E}_1)P^\perp]^{-1}\|$, as considered above), we can iterate this procedure inductively and conclude the proof as Fröhlich, Spencer, and Wittwer do. While we cannot control the bad set of $\theta \in \mathbb{T}$ by the m function as they do, we can bound it by controlling the number of Rellich functions we construct in $\mathcal{E}_n$ at each scale. Since the energy regions at scale s are of size at least $\mathcal{O}(\delta_{s-2}^3)$, each energy region at scale $s-1$ gives rise to at most $\mathcal{O}(\delta_{s-2}^{-3})$ Rellich functions at scale s ; thus, we inductively bound $|\mathcal{E}_n| \leq \mathcal{O}(\delta_{n-2}^{-4})$. The bad set of θ at scale n for a specific $\mathbf{E}_n \in \mathcal{E}_n$ is bounded in measure by $\ell_{n+1}^2 \delta_{n-1}^{1/4}$ by a calculus argument. Since $\delta_{n-2}^{-4} \ell_{n+1}^2 \delta_{n-1}^{1/4}$ is still summable, we can apply Borel–Cantelli.

It remains to show that the Rellich functions in $\mathcal{E}_{n+1}$ inherit the structural properties of those in $\mathcal{E}_n$; namely, a Morse condition and a uniform separation estimate. By construction, simple resonant Rellich functions are well-separated from others, so they satisfy $\|\mathbf{E}_{n+1} - \mathbf{E}_n\|_{C^2} \ll \delta_n$ by the same arguments used above. In the double-resonant case, the Morse lower bound on the second derivative follows by a slight modification of the above argument to allow for V 's asymmetry. A new argument is required to separate the pair of double-resonant Rellich functions uniformly by a stable, quantifiable gap. We thus show

Lemma 4.8. *In our setting, double resonances of a Rellich function $\mathbf{E}_n$ of H_{A_n} resolve as a pair of uniformly locally separated Morse Rellich functions $\mathbf{E}_{n+1,\vee} > \mathbf{E}_{n+1,\wedge}$ of $H_{A_{n+1}}$ with at most one critical point, cf. Fig. 13. The size of the gap is larger than the next resonance scale:*

$$\inf \mathbf{E}_{n+1,\vee} - \sup \mathbf{E}_{n+1,\wedge} \gg \delta_{n+1}$$

This gap ensures that any Rellich function $\mathbf{E}_n$ can resonate only with itself at future scales, which ultimately enables our induction.

To prove Lemma 4.8, we interlace two auxiliary curves between the double-resonant Rellich pair. Specifically, let $\mathbf{E}_n(\theta), \tilde{\mathbf{E}}_n(\theta)$ be the two resonant Rellich functions with corresponding eigenvectors $\psi(\theta), \tilde{\psi}(\theta)$. By the Min–Max Principle, there must be an eigenvalue $\tilde{\lambda}$ of $P_\psi^\perp H_{A_{n+1}} P_\psi^\perp$ satisfying

$$\mathbf{E}_{n+1,\wedge}(\theta) \leq \tilde{\lambda}(\theta) \leq \mathbf{E}_{n+1,\vee}(\theta)$$

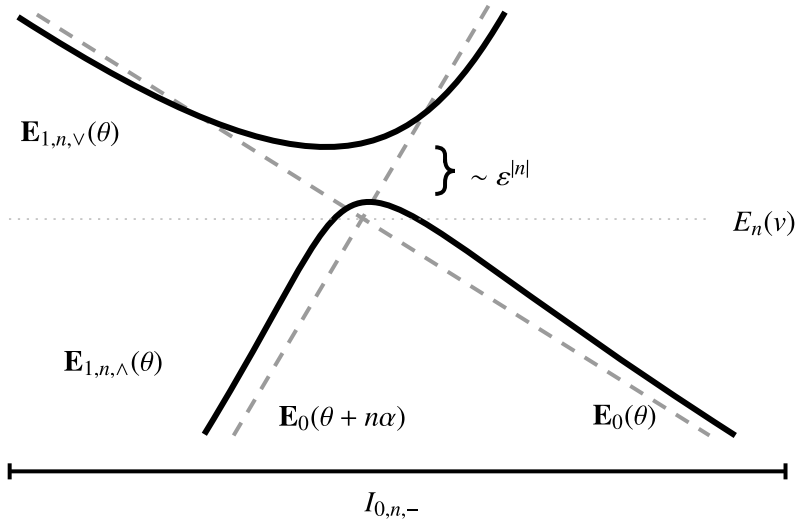


Fig. 13. The resolution of a double resonance of $\mathbf{E}_0 = V$ into a pair of uniformly locally well-separated Rellich curves of a Dirichlet restriction H_{Λ_1} of H . The curves $\mathbf{E}_0(\theta)$ and $\mathbf{E}_0(\theta + n\alpha)$ need not interlace the Rellich curves $\mathbf{E}_1$, but the auxiliary curves $\lambda, \tilde{\lambda}$ (not pictured) must.

Moreover, since we have projected away from one resonance, the arguments from the simple-resonance case can be used to show that $\|\tilde{\lambda} - \tilde{\mathbf{E}}_n\|_{C^1} \ll \delta_n$. As a consequence of the Morse condition, $|\partial_\theta \tilde{\mathbf{E}}_n| \gg \delta_n$, so $|\partial_\theta \tilde{\lambda}|$ is similarly bounded below. By repeating this process to construct an eigenvalue λ of $P_\psi^\perp H_{\Lambda_{n+1}} P_\psi^\perp$ with $\|\lambda - \mathbf{E}_n\|_{C^1} \ll \delta_n$, we construct two curves, with large opposite-signed first derivatives, which separate $\mathbf{E}_{n+1,\vee}$ and $\mathbf{E}_{n+1,\wedge}$. Combining this with the pointwise separation bound gives a uniform separation bound, proving Lemma 4.8 and allowing the inductive argument to proceed.

No version of this proof currently exists for more than two critical points. In higher dimensions, which can mean both a higher-dimensional lattice Laplacian, as well as potentials defined on $\mathbb{T}^d$ with $d \geq 2$, it is even more daunting to implement this perturbative proof strategy. This is why we will impose a much more rigid assumption on the potential function, namely analyticity, for the remainder of these lectures. Smooth potentials are a largely uncharted territory, especially in higher dimensions.

5. Subharmonic functions in the plane

This section⁴ establishes some standard facts about harmonic and subharmonic functions in the plane. In the subsequent development of the theory of quasi-periodic localization for analytic potentials, we will make heavy use of such results as Riesz' representation of subharmonic functions, and the Cartan estimate. A reader familiar with this material can move on to the following section.

5.1. Motivation and definition

Let $\Omega \subset \mathbb{C}$ be a domain (open and connected). Let $\mathcal{H}(\Omega)$ denote the holomorphic functions on Ω . What sort of function is $\log|f(z)|$ for $f \in \mathcal{H}(\Omega)$ with $f \not\equiv 0$? Recall that for $f \in \mathcal{H}(\Omega)$ if $f \neq 0$ in Ω simply connected then there exists $g \in \mathcal{H}(\Omega)$, unique up to an additive constant in $2\pi i\mathbb{Z}$, such that $f = e^g$. Indeed, if $f = e^g$ then $f' = g'e^g$ so that $g' = \frac{f'}{f}$. Then for any $z_0 \in \Omega$, set $g(z) = g(z_0) + \int_{z_0}^z \frac{f'(w)}{f(w)} dw$, where this integral is well-defined because the integrand is holomorphic and Ω is simply connected. The upshot of this

⁴ Based on notes written and typed by Adam Black during a graduate class by the author at Yale.

is that for non-vanishing f , $\log|f| = \log e^{\operatorname{Re} g} = \operatorname{Re} g$ so that $\log|f|$ is harmonic. Notice that this is still true if Ω is not simply connected because being harmonic is a local property and we can always find the existence of such a g in a disk around any point. Now, if $f(z_0) = 0$, then we may write $f(z) = (z - z_0)^n \tilde{f}(z)$ where $\tilde{f}(z)$ does not vanish in some neighborhood of z_0 . In this neighborhood, we have

$$\log|f(z)| = n \log|z - z_0| + \log|\tilde{f}(z)|$$

which we can make sense of in the entire neighborhood by declaring $\log|z - z_0| = -\infty$ at $z = z_0$. Indeed, this function is continuous as map into $\mathbb{R} \cup \{-\infty\}$ relative to the natural topology. More generally, if $K \subset\subset \Omega$ (that is, compactly contained) then we let $\{\zeta_j\}_{j=1}^N$ be the zeros of f in K counted with multiplicity so that $f(z) = \prod_{j=1}^N (z - \zeta_j) F(z)$ where F is holomorphic on some $\Omega' \supset K$ and $F \neq 0$ in Ω' . Then $\log|f(z)| = \sum_{j=1}^N \log|z - \zeta_j| + \log|F(z)|$. From this we infer what type of function $\log|f|$ is, namely it is harmonic away from the zeros of f , and $-\infty$ there, so the value of the function should be *lower* than its average on a small disk. This motivates the following definition, which applies to all dimensions. However, throughout we limit ourselves to the plane.

Definition 5.1. A function u is *subharmonic* on $\Omega \subset \mathbb{R}^2$, denoted $u \in \mathcal{SH}(\Omega)$, if

- $u : \Omega \rightarrow [-\infty, \infty)$ is upper semi-continuous (usc)
- u satisfies the subharmonic mean value property (smvp):

$$u(x_0) \leq \int_{\partial \mathbb{D}(x_0, r)} u(y) \sigma(dy)$$

for any disk $\mathbb{D}(z_0, r)$ such that $\overline{\mathbb{D}(z_0, r)} \subset \Omega$.

One should think of subharmonic functions as lying below harmonic ones, see [Corollary 5.9](#). Hence, in one dimension, subharmonic functions are convex as they lie below lines, which are the one-dimensional harmonic functions. The integral in the above definition is well defined (although it may be $-\infty$) because of the following lemma.

Lemma 5.2. Let $f : K \rightarrow [-\infty, \infty)$ be usc with K compact. Then f attains its maximum.

Proof. Let $M := \sup_{x \in K} f(x)$. Let $f(x_i) \rightarrow M$ as $i \rightarrow \infty$. By compactness, pass to a subsequence if necessary so that $x_i \rightarrow x$. Then $M = \limsup_{i \rightarrow \infty} f(x_i) \leq f(x) \leq M$. $\square$

5.2. Basic properties

In this section we prove some basic properties of subharmonic functions. Readers familiar with the properties of harmonic functions may find these proofs rather familiar.

Proposition 5.3. If $u \in \mathcal{SH}(\Omega)$ then $u(z_0) \leq \int_{\mathbb{D}(z_0, r)} u(z) m(dz)$ for all $\overline{\mathbb{D}(z_0, r)} \subset \Omega$ where m is Lebesgue measure in the plane.

Proof. For all $0 < s \leq r$ we have that

$$u(z_0) |\partial \mathbb{D}(z_0, s)| \leq \int_{\partial \mathbb{D}(z_0, s)} u(z) \sigma(dz)$$

so that the result follows immediately by integrating both sides from 0 to r with respect to s . $\square$

Corollary 5.4. *Let $u, v \in \mathcal{SH}(\Omega)$ such that $u(z) = v(z)$ for almost every z . Then $u \equiv v$.*

Proof. By the smvp and the fact that u and v are equal almost everywhere, we see that for every z_0 for any $r > 0$ such that $\mathbb{D}(z_0, r) \subset \Omega$

$$u(z_0) \leq \int_{\mathbb{D}(z_0, r)} v(z) m(dz)$$

Let $r_i \rightarrow 0$ and let $v(z)$ attain its maximum on $\mathbb{D}(z_0, r_i)$ at z_i so that $z_i \rightarrow z_0$. Thus for all i

$$u(z_0) \leq \int_{\mathbb{D}(z_0, r_i)} v(z) m(dz) \leq v(z_i)$$

so taking limsup we see that

$$u(z_0) \leq v(z_0)$$

by usc. By symmetry, we have also that $v(z_0) \leq u(z_0)$, so we are done. $\square$

Lemma 5.5. *Suppose $u \in C^2(\Omega)$. Then $u \in \mathcal{SH}$ iff $\Delta u(z) \geq 0$ for all $z \in \Omega$.*

Proof. Define

$$(Mu)_{x_0}(r) := \int_{\partial\mathbb{D}(x_0, r)} u(y) \sigma(dy) = \int_{|w|=1} u(x_0 + rw) \sigma(dw)$$

where σ is the surface measure on the circle. We compute

$$\partial_r (Mu)_{x_0}(r) = \int_{|w|=1} \nabla u(x_0 + rw) \cdot w \sigma(dw) = \frac{1}{|\partial\mathbb{D}(0, 1)|r} \int_{\partial\mathbb{D}(x_0, r)} \nabla u(y) \cdot \vec{n} \sigma(dw)$$

which by the divergence theorem is equal to

$$\frac{1}{|\partial\mathbb{D}(0, 1)|r} \int_{\mathbb{D}(x_0, r)} \Delta u(y) m(dy)$$

Thus, we see that if $\Delta u \geq 0$ then $(Mu)_{x_0}(r)$ is non-decreasing with r , and as its limit as $r \rightarrow 0$ is $u(x_0)$, one direction follows. For the other direction, note that if $\Delta u(x_0) < 0$ then there exists some disk $\mathbb{D}(x_0, r)$ on which $\Delta u(x) < 0$. The above computation then shows that $(Mu)_{x_0}$ is decreasing for small enough r , which contradicts the smvp. $\square$

Proposition 5.6. *The function $f(z) = \log|z|$ is subharmonic on $\mathbb{R}^2$.*

Proof. Let $f_n = \frac{1}{2} \log(|z|^2 + 1/n)$. Then it is easy to compute in polar coordinates that $\Delta f_n = \frac{2(1/n)}{(r^2 + 1/n)^2} \geq 0$ so that because f_n is $C^2(\mathbb{R}^2)$, it is subharmonic. On $\partial\mathbb{D}(z_0, r)$, the sequence $\{f_n\}$ is bounded above by some M so that $M - f_n$ is a positive monotone sequence of integrable functions. By applying the monotone convergence theorem to this sequence we see that

$$\lim_{n \rightarrow \infty} \int_{\partial\mathbb{D}(z_0, r)} f_n(z) \sigma(dz) = \int_{\partial\mathbb{D}(z_0, r)} f(z) \sigma(dz)$$

from which the result follows. $\square$

Lemma 5.7. *The maximum or sum of finitely many subharmonic functions is subharmonic.*

Proof. Follows directly from the definition. $\square$

Lemma 5.8 (*Maximum principle*). *Let $u \in \mathcal{SH}(\Omega)$ with Ω connected and suppose there exists $z_0 \in \Omega$ such that $u(z) \leq M := u(z_0)$ for all $z \in \Omega$. Then u is constant.*

Proof. Consider $S = \{z \in \Omega \mid u(z) = M\}$. This set is closed because u is usc. Furthermore it is open because if $f(z) = M$, then $f(z) \leq \int_{\partial \mathbb{D}(z,r)} f(w) \sigma(dw)$ implies that $f(w) = M$ for all $w \in \partial \mathbb{D}(z, r)$. $\square$

The following result explains the terminology *subharmonic*.

Corollary 5.9. *Let $u \in \mathcal{SH}(\Omega)$. If v is harmonic on $\overline{\Omega'} \subset \Omega$ for Ω' bounded and $v \geq u$ on $\partial \Omega'$ then $v \geq u$ in Ω' .*

Proof. The function $u - v$ is subharmonic so that if $u - v > 0$ in Ω' then it would have a maximum in this region, violating the above. $\square$

5.3. Review of harmonic functions

In the next section we will need some basic facts about harmonic functions, which we now briefly recall. They can be found in many places, such as [38]. For $\Omega \subset \mathbb{R}^2$ a bounded region with smooth boundary, say, we would like to solve the boundary value problem

$$\begin{cases} \Delta u(z) = f & z \in \Omega \\ u(z) = g & z \in \partial \Omega \end{cases}$$

Recall Green's identity for $u, v \in C^2(\Omega)$:

$$\int_{\Omega} u(\zeta) \Delta v(\zeta) m(d\zeta) = \int_{\Omega} \Delta u(\zeta) v(\zeta) m(d\zeta) + \int_{\partial \Omega} \left(u \frac{\partial v}{\partial n} - \frac{\partial u}{\partial n} v \right) d\sigma \quad (5.1)$$

If $v = G(z, \zeta)$ is such that (in the sense of distributions) $\Delta_z G(z, \zeta) = \delta_{\zeta}(z)$ and $G(z, \zeta) = 0$ for $z \in \partial \Omega$ then

$$u(z) = \int_{\Omega} G(z, \zeta) f(\zeta) m(d\zeta) + \int_{\partial \Omega} \frac{\partial G}{\partial n}(z, \zeta) g(\zeta) \sigma(d\zeta) \quad (5.2)$$

with m Lebesgue measure in the plane and σ surface measure on the boundary. Such a *Green function* $G(z, \zeta)$ exists for any bounded domain Ω for which $\partial \Omega$ satisfies an exterior cone condition. This is a standard application of Perron's method, see [38] (this method applies to any dimension). For the case of a disk $\mathbb{D}(0, R) \subset \mathbb{C}$, there is the explicit formula given by the logarithm of the absolute value of the conformal automorphism of the disk:

$$G(z, \zeta) = \frac{1}{2\pi} \log |z - \zeta| + \frac{1}{2\pi} \log \left| \frac{R}{R^2 - z\bar{\zeta}} \right| \quad (5.3)$$

In particular, by (5.1) a harmonic function on Ω which is $C^2(\bar{\Omega})$ with boundary values g is given by

$$u(z) = \int_{\partial \Omega} \frac{\partial G}{\partial n}(z, \zeta) g(\zeta) \sigma(d\zeta) \quad (5.4)$$

This is *Poisson's formula* and $P_{\zeta}(z) = \frac{\partial G}{\partial n}(z, \zeta)$ is the Poisson kernel of Ω . If $g \in C(\partial \Omega)$, then (5.4) defines a harmonic function in Ω which is the unique solution of the boundary value problem (uniqueness by the maximum principle). For the disk of radius r in the plane we have

$$P_{\zeta}(z) = \frac{1}{2\pi} \frac{r^2 - |z|^2}{r|z - \zeta|^2}$$

and there is an analogous expression in higher dimensions. This implies Harnack's inequality, which controls the value of a positive harmonic function on a disk by its value at the center.

Proposition 5.10. *Let u be positive harmonic function on the disk $\overline{\mathbb{D}(z_0, R)} \subset \mathbb{C}$. Then for $|z - z_0| < r$ one has*

$$\frac{R-r}{R+r}u(z_0) \leq u(z) \leq \frac{R+r}{R-r}u(z_0).$$

Proof. Simply bound the Poisson kernel and then apply the mean value property. $\square$

Finally, we recall the following compactness property of families of harmonic functions (the analogue of normal families in complex analysis). It is valid in all dimensions but we state it only in the plane.

Theorem 5.11. *A sequence of harmonic functions on $\Omega \subset \mathbb{C}$ that is uniformly bounded on each compact subset of Ω has a subsequence which converges to some harmonic u uniformly on each compact subset.*

Proof. If $u(z)$ is harmonic on $\mathbb{D}(a, r)$ then taking derivatives of (5.4) shows that $|D^\alpha u(a)| \leq \frac{C_\alpha \|u\|_{L^\infty}}{r^\alpha}$ for some universal constant C_α . Thus, any uniformly bounded sequence of harmonic functions is in fact equicontinuous. We can then take a convergent subsequence on any compact subset by Arzela–Ascoli at which point a diagonal argument with increasing compact sets finds the desired u . By the mean value property u is harmonic. $\square$

5.4. Riesz representation of subharmonic functions in $\mathbb{C}$

As noted earlier, any subharmonic function of the form $\log|f|$ for $f \in \mathcal{H}(\Omega)$ admits the representation for any $\Omega' \Subset \Omega$ (compact containment):

$$\log|f| = \sum_{j=1}^N \log|z - \zeta_j| + h(z)$$

with h harmonic in Ω' and $\zeta_j \in \Omega'$. We can think of this expression as $\int_\Omega \log|z - \zeta| \mu(d\zeta) + h(z)$ where $\mu = \sum_{j=1}^N \delta_{\zeta_j}$. Note that h is bounded on any $\Omega'' \Subset \Omega'$ but not necessarily on Ω' . This section develops an analogous representation for all subharmonic functions, known as Riesz representation. The difference is that we can allow any positive finite measure μ . We begin with some basic properties of logarithmic potentials of such measures.

Proposition 5.12. *Let Ω be a bounded domain and $\mu \in \mathcal{M}^+(\Omega)$, that is, a positive finite Borel measure on Ω . Then with $u(z) := \int_\Omega \log|z - \zeta| \mu(d\zeta)$*

- $u \in \mathcal{SH}(\Omega)$
- $u > -\infty$ (Lebesgue) almost everywhere
- u is bounded above.

Proof. Note that for $z, \zeta \in \Omega$, $\log|z - \zeta| \leq \log(\text{diam}\Omega)$ so that $u(z) \leq \log(\text{diam}\Omega)\mu(\Omega)$, which shows that u is bounded above.

Consider $\mathbb{D} \Subset \Omega$ a disk of radius R . Then with m the Lebesgue measure in $\mathbb{R}^2$,

$$\int_{\mathbb{D}} u(z) m(dz) = \int_\Omega \int_{\mathbb{D}} \log|z - \zeta| m(dz) \mu(d\zeta)$$

by Fubini–Tonelli for positive measurable functions because the integrands are bounded from above. In fact, $\log|z - \zeta|$ is Lebesgue integrable on $\mathbb{D}$:

$$\int_{\mathbb{D}} \log|z - \zeta| m(dz) \geq \int_{\mathbb{D}(0,1)} \log|z| m(dz) = 2\pi \int_0^1 r \log r \cdot dr > -\infty$$

which also shows that the total integral is $> -\infty$. Since this holds for any disk, we have shown that $u > -\infty$ a.e. in Ω . To see that u is usc, observe that if $z_j \rightarrow z$ then by (the reverse) Fatou's lemma

$$\limsup_{j \rightarrow \infty} u(z_j) = \limsup_{j \rightarrow \infty} \int_{\Omega} \log|z_j - \zeta| \mu(d\zeta) \leq \int_{\Omega} \limsup_{j \rightarrow \infty} \log|z_j - \zeta| \mu(d\zeta) = u(z)$$

where the use of Fatou's lemma is justified due to the uniform upper bound on $\log|z - \zeta|$. Finally, note that for $\mathbb{D}$ a disk centered at z_0

$$\int_{\partial \mathbb{D}} u(z) dz = \int_{\Omega} \int_{\partial \mathbb{D}} \log|z - \zeta| dz \mu(d\zeta) \geq \int_{\Omega} \log|z_0 - \zeta| d\mu(\zeta) = u(z_0)$$

which shows that $u(z)$ satisfies the smvp because $\log|z|$ does. $\square$

Remark 5.13. We cannot hope for any better than usc from this construction. For instance, consider $\mu = \sum_{n=1}^{\infty} 2^{-n} \delta_{2^{-n}}$ so that $u(z) = \sum_{n=1}^{\infty} 2^{-n} \log|z - 2^{-n}|$. Then $u(0) = -2\log 2$ but $u(2^{-n}) = -\infty$ for all n .

We will also require the following smooth approximation result.

Lemma 5.14. *Let $u \in \mathcal{SH}(\Omega)$ where Ω is a bounded domain. Then there exists a sequence $u_n \in \mathcal{SH}(\Omega_{1/n}) \cap C^{\infty}(\Omega_{1/n})$ where $\Omega_{1/n} := \{z \in \Omega \mid \text{dist}(z, \partial\Omega) > 1/n\}$ such that $u_n \rightarrow u$ pointwise and monotone decreasing (in Ω_{1/n_0} for $n > n_0$).*

Proof. We accomplish this via mollification, so let $\varphi \in C^{\infty}(\mathbb{R}^2)$ be a radial function satisfying $\varphi(x) \geq 0$, $\varphi(x) = 0$ for $|x| \geq 1$ and $\int_{\mathbb{R}^2} \varphi(x) m(dx) = 1$. Define also $\varphi_n = n^2 \varphi(nx)$. We claim that $u_n(z) = (u * \varphi_n)(z)$ satisfies the desired properties. It is clearly smooth and well-defined on $\Omega_{1/n}$. The smvp for u_n follows from Fubini's theorem and $\varphi_n \geq 0$. To see that u_n is decreasing, write

$$u_n(z) = n^2 \int_{\mathbb{R}^2} u(z - w) \varphi(nw) dw = 2\pi \int_0^{\infty} \int_0^1 u(z - \frac{r}{n} e(\theta)) d\theta r \varphi(r) dr \geq u(z)$$

with $e(\theta) = e^{2\pi i \theta}$, the final inequality implied by the smvp. First, $v(\zeta) := \int_0^1 u(z - \zeta e(\theta)) d\theta$ is subharmonic since it is easily seen to be usc, and the smvp follows by Fubini (note that u remains subharmonic after a rotation and translation). Second, it is radial and thus an increasing (but not necessarily in the strict sense) function of $|\zeta|$ by the maximum principle. Finally, $u_n(z) \leq \max_{|z-w| \leq 1/n} u(w)$ for $z \in \Omega_{1/n}$ so that by usc $u_n(z) \rightarrow u(z)$ as $n \rightarrow \infty$. $\square$

We are now ready to prove Riesz's representation theorem for subharmonic functions.

Theorem 5.15. *Let $u \in \mathcal{SH}(\Omega)$ where Ω is some neighborhood of $\overline{\mathbb{D}(0, 4)}$. Suppose that $u \leq M$ on $\overline{\mathbb{D}(0, 4)}$ and $u(0) \geq m > -\infty$. Then there exists $\mu \in \mathcal{M}^+(\mathbb{D}(0, 3))$ and h harmonic in $\mathbb{D}(0, 3)$ such that for all $z \in \mathbb{D}(0, 3)$*

$$u(z) = \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + h(z)$$

Furthermore, there exists $C_0 > 0$ universal such that $\|h - M\|_{L^{\infty}(\mathbb{D}(0,2))} \leq C_0(M - m)$ and $\mu(\mathbb{D}(0, 3)) \leq C_0(M - m)$. In fact, for any $\delta \in (0, 1)$ there exists $C_0(\delta)$ so that $\|h - M\|_{L^{\infty}(\mathbb{D}(0,3-\delta))} \leq C_0(\delta)(M - m)$.

Proof. We first reduce to the smooth case. To this end, suppose that the claim holds for all $v \in \mathcal{SH}(\overline{\mathbb{D}(0,4)}) \cap C^\infty(\overline{\mathbb{D}(0,4)})$. Choose any $u \in \mathcal{SH}(\overline{\mathbb{D}(0,4)})$ and let $u_n \rightarrow u$ in $\mathbb{D}(0,4)$ be as in Lemma 5.14. We then have with some decreasing $M_n \rightarrow M$

$$u_n \leq M_n \text{ on } \overline{\mathbb{D}(0,4)}, \quad u_n(0) \geq m$$

By validity of the theorem in the smooth case we may write

$$u_n(z) = \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu_n(d\zeta) + h_n(z) \quad (5.5)$$

and because u_n is monotone decreasing and uniformly bounded above on any compact set, for any $\varphi \in C(\overline{\mathbb{D}(0,3)})$ we have that

$$\langle u_n, \varphi \rangle \rightarrow \langle u, \varphi \rangle = \int_{\mathbb{D}(0,3)} u(x + iy) \varphi(x + iy) dx dy$$

by the monotone convergence theorem. By assumption, the above measures are uniformly bounded, so by Banach–Alaoglu we may take a weak-* limit in $C(\overline{\mathbb{D}(0,3)})^*$, thus $\mu_n \rightarrow \mu$ in the weak-* sense where μ is a finite Borel measure on $\overline{\mathbb{D}(0,3)}$ which satisfies

$$\mu(\overline{\mathbb{D}(0,3)}) \leq C_0(M - m)$$

Since, with $m(dz)$ being Lebesgue measure in the plane,

$$\psi(\zeta) := \int_{\mathbb{D}(0,3)} \log|z - \zeta| \varphi(z) m(dz)$$

is a continuous function of $\zeta \in \mathbb{R}^2$, we conclude that

$$\lim_{n \rightarrow \infty} \int_{\mathbb{D}(0,3)} \psi(\zeta) \mu_n(d\zeta) = \int_{\mathbb{D}(0,3)} \psi(\zeta) \mu(d\zeta)$$

which implies that

$$\left\langle \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu_n(d\zeta), \varphi \right\rangle \rightarrow \left\langle \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta), \varphi \right\rangle. \quad (5.6)$$

By the theorem in the smooth case,

$$\limsup_{n \rightarrow \infty} \|h_n - M\|_{L^\infty(\mathbb{D}(0,3-\delta))} \leq C_0(\delta)(M - m)$$

so by Theorem 5.11 there exists some h harmonic in $\mathbb{D}(0,3)$ such that a subsequence of $\{h_n\}$ converges to h uniformly on all compact subsets of $\mathbb{D}(0,3)$. Thus, for any $\varphi \in C(\overline{\mathbb{D}(0,3)})$ of compact support, $\langle h_n, \varphi \rangle \rightarrow \langle h, \varphi \rangle$ along this sequence. In combination with (5.5), (5.6) we conclude that

$$\langle u, \varphi \rangle = \left\langle \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + h(z), \varphi \right\rangle$$

Thus

$$u(z) = \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + h(z) \text{ almost everywhere in } \mathbb{D}(0,3),$$

which in turn implies equality everywhere by Corollary 5.4. Finally, to obtain the desired form we write

$$\int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) = \int_{\partial\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta)$$

and notice that $\int_{\partial\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta)$ is harmonic in $\mathbb{D}(0, 3)$. Thus,

$$u(z) = \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + h_0(z)$$

where $h_0(z) = \int_{\partial\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + h(z)$ is harmonic in $\mathbb{D}(0, 3)$. This harmonic function h_0 satisfies similar L^∞ bounds as before, albeit with different constants.

It remains to prove the theorem for smooth subharmonic functions on $\overline{\mathbb{D}(0, 4)}$. In view of (5.2)

$$u(z) = \int_{\mathbb{D}(0,4)} G(z, \zeta) \Delta u(\zeta) m(d\zeta) + \int_{\partial\mathbb{D}(0,4)} \frac{\partial G}{\partial n}(z, \zeta) u(\zeta) \sigma(d\zeta) \quad (5.7)$$

so that by using the particular form $G(z, \zeta)$ in (5.3), and defining $\mu(dz) := \frac{1}{2\pi} \Delta u(z) dz$ we rewrite the above as

$$u(z) = \int_{\mathbb{D}(0,3)} \log|z - \zeta| \mu(d\zeta) + \int_{\mathbb{D}(0,3)} \log \frac{4}{|16 - z\bar{\zeta}|} \mu(d\zeta) + \int_{\mathbb{D}(0,4) \setminus \mathbb{D}(0,3)} G(z, \zeta) \Delta u(\zeta) d\zeta + h_0(z)$$

where $h_0(z) := \int_{\partial\mathbb{D}(0,4)} \frac{\partial G}{\partial n}(z, \zeta) u(\zeta) \sigma(d\zeta)$ is the harmonic extension of u to $\mathbb{D}(0, 4)$, see (5.4). The second term is harmonic for $z \in \mathbb{D}(0, 3)$ because $16 - z\bar{\zeta} \neq 0$ and the third term because $\zeta \in \mathbb{D}(0, 4) \setminus \mathbb{D}(0, 3)$ and thus

$$h(z) = \int_{\mathbb{D}(0,3)} \log \frac{4}{|16 - z\bar{\zeta}|} \mu(d\zeta) + \int_{\mathbb{D}(0,4) \setminus \mathbb{D}(0,3)} G(z, \zeta) \Delta u(\zeta) d\zeta + h_0(z) \quad (5.8)$$

is harmonic in $\mathbb{D}(0, 3)$. We have therefore obtained the desired form for u , we only have left to show the stated bounds. To bound $\mu(\mathbb{D}(0, 3))$, use (5.7) to see that

$$\begin{aligned} u(0) &= \int_{\mathbb{D}(0,4)} G(0, \zeta) \mu(d\zeta) + h_0(0) = \int_{\mathbb{D}(0,4)} \log \frac{|\zeta|}{4} \mu(d\zeta) + h_0(0) \\ \log \frac{4}{r} \mu(\mathbb{D}(0, r)) &\leq \int_{\mathbb{D}(0,r)} \log \frac{4}{|\zeta|} \mu(d\zeta) = h_0(0) - u(0) \leq M - m \end{aligned}$$

where we have used that $u(0) = m$ and the fact that $u \leq M$ on $\partial\mathbb{D}(0, 4)$ implies that $h_0 \leq M$. Setting $r = 3$, we see that $\mu(\mathbb{D}(0, 3)) \leq C(M - m)$ as desired. For $z \in \mathbb{D}(0, 3)$, the first term in (5.8) is negative by inspection, the second negative since $G < 0$, and the third is bounded above by M as before. Therefore, $h(z) \leq M$. For the reverse bound, Harnack's inequality on $|z| \leq 3 - \delta$ yields

$$M - h(z) \leq \frac{3+r}{3-r} (M - h(0)) \leq \frac{6-\delta}{\delta} (M - h(0))$$

and

$$h(0) = u(0) - \int_{\mathbb{D}(0,3)} \log|\zeta| \mu(d\zeta) \geq m - \int_{\mathbb{D}(0,3) \setminus \mathbb{D}(0,1)} \log|\zeta| \mu(d\zeta) \geq m - C(M - m)$$

so putting these together implies that

$$M - C_\delta(M - m) \leq h(z)$$

for all $|z| \leq 3 - \delta$. $\square$

In fact, by essentially the same proof one can obtain the following more general Riesz representation. Note that one can move the point z_0 to 0 by an automorphism of the disk, which retains the property of being subharmonic.

Theorem 5.16. *Let $u \in \mathcal{SH}(\overline{\mathbb{D}(0, R_1)})$ and suppose that $u \leq M$ on $\overline{\mathbb{D}(0, R_1)}$ and $u(z_0) \geq m > -\infty$ where $|z_0| < R_1$. Let $R_1 > R_2 > R_3 > 0$. There exists $\mu \in \mathcal{M}^+(\mathbb{D}(0, R_2))$ and h harmonic in $\mathbb{D}(0, R_2)$ such that for all $z \in \mathbb{D}(0, R_2)$*

$$u(z) = \int_{\mathbb{D}(0, R_2)} \log|z - \zeta| \mu(d\zeta) + h(z)$$

Furthermore, there exist $C_0 = C_0(z_0, R_1, R_2) > 0$ and $C_1 = C_1(z_0, R_1, R_2, R_3) > 0$ universal such that $\mu(\mathbb{D}(0, R_2)) \leq C_0(M - m)$ and $\|h - M\|_{L^\infty(\mathbb{D}(0, R_3))} \leq C_1(M - m)$.

See Theorem 2.2 in [31] for explicit constants.

5.5. Cartan's lower bound

Next, we prove Cartan's theorem which controls large negative values of logarithmic potentials. Levin's book [41] has much more on this topic, see page 76.

Theorem 5.17. *Let μ be a finite positive measure in $\mathbb{C}$ and consider the logarithmic potential*

$$u(z) = \int_{\mathbb{R}^2} \log|z - \zeta| \mu(d\zeta)$$

For any $H \in (0, 1)$ there exist disks $\{\mathbb{D}(z_j, r_j)\}_{j=1}^J$, for $1 \leq J \leq \infty$ with $\sum_{j=1}^J r_j \leq 5H$ and

$$u(z) \geq -\|\mu\| \log(e/H) \quad \forall z \in \mathbb{C} \setminus \bigcup_{j=1}^J \mathbb{D}(z_j, r_j) \quad (5.9)$$

Proof. Let $z \in \mathbb{C}$ be a good point if $n(z, r) := \mu(\mathbb{D}(z, r)) \leq pr$ for all $r > 0$. Here p depends on H and will be determined. For every bad z there exists $r(z) > 0$ with $n(z, r(z)) > r(z)p$. Note that $r(z) \leq \|\mu\|/p$. By Vitali's covering lemma there exist bad points z_j so that $\{\mathbb{D}(z_j, r(z_j))\}_j$ are pairwise disjoint and

$$\mathcal{B} := \{z \in \mathbb{C} \mid z \text{ is a bad point}\} \subset \bigcup \mathbb{D}(z_j, r_j) \text{ with } r_j := 5r(z_j).$$

In particular, $\sum_j r_j \leq 5\|\mu\|/p$ whence we need to set $p = \|\mu\|/H$. If $z \in \mathbb{C} \setminus \bigcup \mathbb{D}(z_j, r_j)$, then z is good and we obtain by integrating by parts

$$\begin{aligned} u(z) &\geq \int_0^1 \log r \, d(n(z, r)) = - \int_0^1 \frac{n(z, r)}{r} \, dr \geq - \int_0^H p \, dr + \|\mu\| \log H \\ &= -pH + \|\mu\| \log H = \|\mu\| \log(H/e) \end{aligned}$$

as claimed. $\square$

We call $\|\mu\|$ the Riesz mass of u . We leave it to the reader to check that Theorem 5.17 with the same proof generalizes as follows.

Theorem 5.18. *Under the same assumptions as in the previous theorem, suppose $0 < \delta \leq 1$. Then for any $H \in (0, 1)$ there exist disks $\{\mathbb{D}(z_j, r_j)\}_{j=1}^J$, for $1 \leq J \leq \infty$ with $\sum_{j=1}^J r_j^\delta \leq 5^\delta H$ and*

$$u(z) \geq -\frac{1}{\delta} \|\mu\| \log(e/H) \quad \forall z \in \mathbb{C} \setminus \bigcup_{j=1}^J \mathbb{D}(z_j, r_j) \quad (5.10)$$

We chose $0 < \delta \leq 1$ here instead of $0 < \delta \leq 2$ since the range $1 < \delta \leq 2$ is weaker than [Theorem 5.17](#). As an immediate corollary we conclude that $\dim(\{z \in \mathbb{C} | u(z) = -\infty\}) = 0$ in the sense of Hausdorff dimension, for any logarithmic potential of a finite positive measure. By [Theorem 5.15](#), this same property therefore holds locally on Ω for any subharmonic function on Ω which is not constant $-\infty$. For our applications, Cartan's theorem, i.e., [Theorem 5.17](#), will suffice. The following serves to illustrate this result.

- Consider the logarithm of a polynomial of degree N with roots $\zeta_j \in \mathbb{C}$. Thus, $P(z) = \prod_{j=1}^N (z - \zeta_j)$ and

$$u(z) = \log \left| \prod_{j=1}^N (z - \zeta_j) \right| = \int \log |z - \zeta| \mu(d\zeta), \quad \mu = \sum_{j=1}^N \delta_{\zeta_j}$$

Given $0 < H < 1$, there exist disks $\mathbb{D}(z_j, r_j)$, $1 \leq j \leq J$, with $\sum_j r_j \leq 5H$ and $|P(z)| \geq (H/e)^N$ for all $z \in \mathbb{C} \setminus \bigcup \mathbb{D}(z_j, r_j)$. By the maximum principle, each disk contains a zero of P (otherwise it can be removed without changing the conclusion). Thus, $J \leq N$. The bound on the Riesz mass in [Theorem 5.15](#) is nothing other than Jensen's formula counting the roots of analytic functions, see [41, page 10].

- If $\zeta_j = 0$ for all j , then $|P(z)| = |z|^N \geq H^N$ if $|z| \geq H$. This shows that Cartan's theorem is optimal up to multiplicative constants on H .
- On the other hand, suppose $\zeta_j = e(j/N)$ for $1 \leq j \leq N$ where $e(\theta) = e^{2\pi i \theta}$. Then $P(z) = z^N - 1$ and we can take the Cartan disks centered at ζ_j of radius $\rho = 1/N$. Then for any z with $z = \zeta_j + \rho e(\theta)$ we have

$$|P(z)| = |z^N - 1| = |(\zeta_j + \rho e(\theta))^N - 1| \geq \rho N - \sum_{\ell=2}^N \binom{N}{\ell} \rho^\ell \geq 1 - \sum_{\ell=2}^N \frac{(N\rho)^\ell}{\ell!} = 3 - e \quad (5.11)$$

It follows from the maximum (minimum) principle for analytic functions that $|P(z)| \geq 3 - e$ for all $z \in \mathbb{C} \setminus \bigcup_{j=1}^N \mathbb{D}(\zeta_j, 1/N)$. Therefore Cartan's estimate is woefully imprecise in this example. Indeed, for the polynomial P with roots at the N th roots of unity, $u(z) = \log |P(z)|$ behaves in [Theorem 5.17](#) like a subharmonic function with Riesz mass 1, at least for $H = 1/N$.

In applications of Cartan's theorem to quasi-periodic localization, the distribution of the zeros plays a decisive role and it is therefore essential to improve on the Cartan bound. In other words, we are in a situation much closer to the roots-of-unity example where Cartan falls far short from the true estimate. Nevertheless, combining Cartan's bound with the dynamics, one can still obtain a nontrivial statement as we shall see in the following section.

To conclude this section, we prove Riesz's representation theorem on general domains from the one for disks which we proved above. We will do this by connection points by chains of disks, which uses Cartan.

Corollary 5.19. *Let $\Omega \subset \mathbb{C}$ be a bounded domain, u subharmonic on Ω with $\sup_{\Omega} u \leq M$. Suppose $u(z_0) \geq m > -\infty$ for some $z_0 \in \Omega$. For any $\Omega_2 \Subset \Omega_1 \Subset \Omega$, there exist a positive measure μ on Ω_1 and a harmonic function h on Ω_1 such that*

$$\begin{aligned} u(z) &= \int_{\Omega_1} \log |z - \zeta| \mu(d\zeta) + h(z) \quad \forall z \in \Omega_1 \\ \mu(\Omega_1) &\leq C_1(\Omega, z_0, \Omega_1)(M - m) \\ \|h - M\|_{L^\infty(\Omega_2)} &\leq C_2(\Omega, z_0, \Omega_1, \Omega_2)(M - m) \end{aligned} \quad (5.12)$$

Proof. By [Lemma 5.14](#) we can assume that u is smooth, although this is strictly speaking not necessary. The measure $\mu(dz) = \frac{1}{2\pi} \Delta u \, dx dy$ is unique and therefore h harmonic on Ω_1 if it satisfies (5.12). Let $\sup_K u = u(z_0)$, $z_0 \in K$. By compactness, there exists $\delta > 0$ and N finite so that for any $z \in \Omega_1$ we can find disks $\mathbb{D}(z_j, \delta) \subset \Omega$, $0 \leq j \leq N$, with $z_N = z$, and $z_j \in \mathbb{D}(z_{j-1}, \delta/2)$ for all $j \geq 1$. Moreover, we

may assume that $\Omega_2 \subset \bigcup_{z \in \Omega_1} \mathbb{D}(z, \delta/2)$ and by compactness this can be chosen as a finite union. By Riesz's representation as in [Theorem 5.16](#) we have

$$\begin{aligned} u(z) &= \int_{\mathbb{D}(z_0, \delta/2)} \log|z - \zeta| \mu(d\zeta) + h_0(z) \quad \forall z \in \mathbb{D}(z_0, \delta/2) \\ \mu(\mathbb{D}(z_0, \delta/2)) &\leq C_0(\delta)(M - m), \quad \|h_0 - M\|_{L^\infty(\mathbb{D}(z_0, \delta/4))} \leq C_0(\delta)(M - m) \end{aligned} \quad (5.13)$$

Next, apply [Theorem 5.17](#) to the logarithmic potential in (5.13) with $H = \delta/100$. Hence, there exists $w_1 \in \mathbb{D}(z_0, \delta/4) \subset \mathbb{D}(z_1, 3\delta/4)$ with

$$u(w_1) \geq m - C_1(\delta)(M - m) \quad (5.14)$$

while $u \leq M$ on $\mathbb{D}(z_1, \delta)$. We now apply Riesz's representation as in [Theorem 5.16](#) on this disk, followed by Cartan to find a good point $w_2 \in \mathbb{D}(z_2, 3\delta/4)$ for which and analogue of (5.14) holds. We may repeat this procedure to finitely many times to cover all of Ω_1 by such disks leading to the stated upper bound on the measure $\mu(\Omega_1)$. For the estimate on the harmonic function h defined by (5.12), pick any $z_* \in \Omega_2$. Then with $\varepsilon_0 := \text{dist}(\partial\Omega_1, \Omega_2)$ we have $\mathbb{D}(z_*, \varepsilon_0) \subset \Omega_1$. On the one hand, for all $z \in \Omega_1$,

$$h(z) \geq u(z) - \log(\text{diam}(\Omega_1))\mu(\Omega_1) \geq u(z) - C(M - m)$$

with the same type of constant as before. By the previous Cartan estimate and chaining argument, we can find $z_{**} \in \mathbb{D}(z_*, \varepsilon_0/4)$ which satisfies a bound (5.14) with a purely geometric constant. Hence

$$h(z_{**}) \geq m - C(M - m) \quad (5.15)$$

On the other hand, again by [Theorem 5.17](#) we may find $\varepsilon_1 \in (3\varepsilon_0/4, \varepsilon_0)$ so that for all $|z - z_*| = \varepsilon_1$ one has

$$\int_{\Omega_1} \log|z - \zeta| \mu(d\zeta) \geq -C(M - m)$$

whence

$$h(z) \leq M + C(M - m) \quad \forall |z - z_*| = \varepsilon_1 \quad (5.16)$$

By Harnack's inequality, (5.15) and (5.16) imply that h satisfies the desired bound on $\mathbb{D}(z_*, \varepsilon_0/2)$ and hence everywhere on Ω_2 . $\square$

Alternatively, one can rely on the proof strategy of [Theorem 5.15](#), and use the Green function on general subdomains of Ω with sufficiently regular boundary. But this seems technically more involved, at least to the author.

6. The Bourgain–Goldstein theorem

In this section we will sketch a proof of the main theorem in [10]. Similar to [Theorem 4.1](#) it addresses Anderson localization for the operators

$$(H_{x,\omega}\psi)_n = \psi_{n-1} + \psi_{n+1} + V(T_\omega^n x)\psi_n \quad (6.1)$$

on $\ell^2(\mathbb{Z})$, where $T_\omega : \mathbb{T} \rightarrow \mathbb{T}$ is the rotation $x \mapsto x + \omega \pmod{1}$ and $V : \mathbb{T} \rightarrow \mathbb{R}$ is analytic.

Theorem 6.1. *Suppose the Lyapunov exponents $L(E, \omega)$ associated with (6.1) satisfy*

$$\inf_{E, \omega} L(E, \omega) > 0.$$

Then for almost every $\omega \in \mathbb{T}$, the operator $H_{0,\omega}$ exhibits pure point spectrum with exponentially decaying eigenfunctions. Moreover, for almost every $\omega \in \mathbb{T}$, the operator $H_{x,\omega}$ exhibits Anderson localization for almost every $x \in \mathbb{T}$.

The final statement of the theorem follows simply by Fubini and the fact that one may replace 0 in $H_{0,\omega}$ with any other $x \in \mathbb{T}$. See [8,10] for versions of this theorem with V analytic on higher-dimensional tori. This section is only meant to serve as a motivation for higher-dimensional techniques involving $\Delta_{\mathbb{Z}^d}$ with $d \geq 2$, and less as a review of [10] itself. We will often drop ω from the notation and write H_x or $H(x)$.

No explicit Diophantine condition arises here in contrast to Theorem 4.1. In fact, it is not known if Theorem 6.1 holds for all Diophantine ω . For $V(x) = \cos(2\pi x)$, Jitomirskaya proved [34] that this is indeed the case. Although Diophantine conditions play a decisive role in the proof of Theorem 6.1, one does remove a measure 0 set of “bad” ω in addition to a measure 0 set of non-Diophantine ω . The smallness condition on ε in Section 4 is replaced by positive Lyapunov exponents, a non-perturbative condition. No assumption on the number of monotonicity intervals of V is made, nor do we impose an explicit nondegeneracy condition. Note, however, that the most degenerate case $V = \text{const}$ cannot arise by positive Lyapunov exponents. By analyticity, V therefore cannot be infinitely degenerate anywhere. No analogue of Theorem 6.1 is known if V is merely smooth, nor is it clear what the results might be for smooth V .

We quickly review some elementary background on Lyapunov exponents. Consider (6.1) with the rotation T_ω replaced by an ergodic transformation $T : X \rightarrow X$ on a probability space (X, ν) , and V is a real-valued measurable function. Define

$$L(E) = \lim_{n \rightarrow \infty} \frac{1}{n} \int_X \log \|M_n(x, E)\| \nu(dx) = \inf_{n \geq 1} \frac{1}{n} \int_X \log \|M_n(x, E)\| \nu(dx) \quad (6.2)$$

where M_n are the transfer matrices

$$M_n(x, E) = \prod_{k=n}^1 \begin{bmatrix} E - V(T^k x) & -1 \\ 1 & 0 \end{bmatrix} \quad (6.3)$$

of (6.1), i.e., the column vectors of M_n are a fundamental system of the equation $H_x \psi = E\psi$. The limit in (6.2) exists as stated due to fact that $a_n := \int_X \log \|M_n(x, E)\| \nu(dx)$ is a subadditive sequence, and it is known that $\lim_{n \rightarrow \infty} \frac{1}{n} a_n = \inf_{n \geq 1} \frac{1}{n} a_n$ exists for such sequences. Since $M_n \in SL(2, \mathbb{R})$ we have $\|M_n\| \geq 1$ and thus $L(E) \geq 0$. It is an important and often difficult question to decide whether $L(E) > 0$ for (6.1), see [31,32] for an example of this. But this circle of problems will not concern us here. It was shown by Fürstenberg and Kesten [24], later generalized in Kingman’s subadditive ergodic theorem, that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|M_n(x, E)\| = L(E) \quad (6.4)$$

for a.e. $x \in X$. This does use ergodicity of T , whereas (6.2) does not. See Viana’s book [50] for all this.

The Thouless formula, see [15],

$$L(E) = \int_{\mathbb{R}} \log |E - E'| N(dE') \quad \forall E \in \mathbb{C} \quad (6.5)$$

relates the Lyapunov exponent to the density of states. Here N is the integrated density of states (IDS), i.e., the limiting distribution of the eigenvalues of (6.1) restricted to intervals $\Lambda = [-N, N]$ in the limit $N \rightarrow \infty$. In other words, there exists a deterministic nondecreasing function N so that for a.e. $x \in X$ one has

$$|\Lambda|^{-1} |\{j \in [1, |\Lambda|] \mid E_j^{(\Lambda)}(x) < t\}| \rightarrow N(t),$$

where $E_j^{(\Lambda)}(x)$ are the eigenvalues of H_x^Λ , the restriction of (6.1) to Λ with Dirichlet boundary conditions. The existence of this limit holds in great generality, see [20]. The Lyapunov exponent is a subharmonic function on $\mathbb{C}$, and harmonic on $\mathbb{C} \setminus \mathbb{R}$. The Thouless formula identifies the IDS N as the Riesz measure of $L(E)$, and also shows that L and $\frac{dN}{dE}$ are related to each other by the Hilbert transform. For far-reaching considerations involving these concepts see for example Avila’s global work on phase transitions [3].

6.1. Large deviation theorems

We now present a key ingredient in the proof of [Theorem 6.1](#), namely the large deviation estimates (LDTs), see also [\[26\]](#) where they are essential in the study of the regularity of the IDS. For the operators [\(6.1\)](#) defined in terms of rotations of $\mathbb{T}$, define

$$L_n(E) = \frac{1}{n} \int_{\mathbb{T}} \log \|M_n(x, E)\| dx.$$

The following LDT can be viewed as a quantitative form of [\(6.4\)](#).

Definition 6.2. By Diophantine, we will now mean any irrational ω so that $\|n\omega\| \geq bn^{-a}$ for all $n \geq 1$.

It is easy to see that for every $a > 1$ a.e. ω satisfies such a condition for some $b = b(\omega)$.

Proposition 6.3. For Diophantine ω there exist $0 < \sigma, \tau < 1$ depending on V, a so that for all $E \in [-E_0, E_0]$,

$$|\{x \in \mathbb{T} : |\log \|M_n(x, E)\| - nL_n(E)| > n^{1-\sigma}\}| \leq \exp(-n^\tau). \quad (6.6)$$

for all sufficiently large $n \geq n_0(V, a, b, E_0)$.

To motivate [\(6.6\)](#), consider the following scalar, or commutative, model:

$$u(x) = \sum_{k=1}^q \log |e(x) - e(k\omega)| \quad (6.7)$$

where $\omega = \frac{p}{q}$ and $e(x) = e^{2\pi i x}$. Then $u(x) = \log |e(xq) - 1|$ and $\int_{\mathbb{T}} u(x) dx = 0$ so that for $\lambda < 0$

$$|\{x \in \mathbb{T} : u(e(x)) < \lambda\}| = |\{x \in \mathbb{T} : |e(x) - 1| < e^\lambda\}| \quad (6.8)$$

which is of size e^λ . In this model case, $u(x + 1/q) = u(x)$. Returning to $u(x) = \log \|M_n(x, E)\|$, this exact invariance needs to be replaced by the almost invariance

$$\sup_{x \in \mathbb{T}} |u(x) - u(x + k\omega)| \leq Ck \text{ for any } k \geq 1. \quad (6.9)$$

The logarithm in our model case [\(6.7\)](#) is a reasonable choice because of Riesz's representation theorem for subharmonic functions applied to the function $u(z) = \log \|M_n(z, E)\|$ which is subharmonic on a neighborhood of $[0, 1]$ in $\mathbb{C}$ by analyticity of V . The subharmonicity can be seen by writing

$$u(z) = \sup_{\|\vec{v}\|=\|\vec{w}\|=1} \log |\langle M_n(z, E)\vec{v}, \vec{w} \rangle|$$

First, $\log |\langle M_n(z, E)\vec{v}, \vec{w} \rangle|$ is subharmonic by analyticity of $\langle M_n(z, E)\vec{v}, \vec{w} \rangle$. Second, the sub-mean value property (smvp) survives under suprema, and so u satisfies the smvp. Finally, the function $u(z)$ is clearly continuous.

Proof (Proof of [Proposition 6.3](#) by Riesz and Cartan). Fix a rectangle R which compactly contains $[0, 1]$. By Riesz representation as stated in [Theorem 5.15](#), there exists a positive measure μ on R and a harmonic function on R such that

$$u(z) = \log \|M_n(z, E)\| = \int_R \log |z - \zeta| \mu(d\zeta) + h(z) \quad (6.10)$$

Since $\|M_n(z)\| \leq e^{Cn}$, $0 \leq u(z) \lesssim n$ on R (with a constant that depends on V , R and E_0) and thus $\|\mu\| \lesssim n$ as well as $\|h\|_{L^\infty(R')} \lesssim n$, where $[0, 1] \Subset R' \Subset R$ is a slightly smaller rectangle. Fix a small $\delta > 0$ and take n large. Then there is a disk $\mathbb{D}_0 = \mathbb{D}(x_0, n^{-2\delta})$, $x_0 \in [0, 1]$ with the property that $\mu(\mathbb{D}_0) \lesssim n^{1-2\delta}$. Write

$$\int_R \log|z - \zeta| \mu(d\zeta) = u_1(z) + u_2(z) = \int_{\mathbb{D}_0} \log|z - \zeta| \mu(d\zeta) + \int_{\mathbb{C} \setminus \mathbb{D}_0} \log|z - \zeta| \mu(d\zeta)$$

Set $\mathbb{D}_1 = \mathbb{D}(x_0, n^{-3\delta})$. Then

$$|u_2(z) - u_2(z')| \lesssim n^{1-\delta} \quad \forall z, z' \in \mathbb{D}_1$$

since

$$|u_2(z) - u_2(z')| = \left| \int_{\mathbb{C} \setminus \mathbb{D}_0} \log \left| 1 + \frac{z' - z}{z - \zeta} \right| \mu(d\zeta) \right| \lesssim \frac{n^{-3\delta}}{n^{-2\delta}} \mu(\mathbb{C}) \lesssim n^{1-\delta}$$

Cartan's theorem applied to $u_1(z)$ yields disks $\{\mathbb{D}(z_j, r_j)\}_j$ so that $\sum_j r_j \lesssim \exp(-2n^\delta)$ and with the property that

$$u_1(z) \gtrsim -n^{1-\delta} \quad \forall z \in \mathbb{C} \setminus \bigcup_j \mathbb{D}(z_j, r_j)$$

From $u_1 \leq 0$ on $\mathbb{D}_1$ and $|h(z) - h(z')| \lesssim n|z - z'|$ on R' , it follows that

$$|u(z) - u(z')| \lesssim n^{1-\delta} \quad \forall z, z' \in \mathbb{D}_1 \setminus \bigcup_j \mathbb{D}(z_j, r_j) \quad (6.11)$$

From the Diophantine property with $1 < a < \frac{4}{3}$, say, for any $x, x' \in \mathbb{T}$ there are positive integers $k, k' \lesssim n^{4\delta}$ such that

$$x + k\omega, x' + k'\omega \in \mathbb{D}_1 \quad \text{mod } \mathbb{Z}$$

An elementary way of seeing this is to use Dirichlet's approximation principle, viz. for any $Q > 1$ there exists a reduced fraction $\frac{p}{q}$ so that $|\omega - p/q| \leq (qQ)^{-1}$ and $1 \leq q < Q$. Then use the Diophantine property to bound q from below in terms of Q . In order to avoid the Cartan disks $\bigcup_j \mathbb{D}(z_j, r_j)$ we need to remove a set $\mathcal{B} \subset \mathbb{T}$ of measure $\lesssim \exp(-n^\delta)$. For this step is important that Cartan controls the sum of the radii, i.e., $\sum_j r_j \lesssim \exp(-2n^\delta)$ since then the disks remove at most measure $\lesssim \exp(-2n^\delta)$ from the real line. Then from the almost invariance (6.9), for any $x, x' \in \mathbb{T} \setminus \mathcal{B}$,

$$|u(x) - u(x')| \lesssim n^{4\delta} + n^{1-\delta} \lesssim n^{1-\delta}$$

This implies (6.6) with $\sigma = \tau = \delta$. $\square$

This proof generalizes to other types of dynamics such as higher-dimensional shifts $Tx = x + \omega \pmod{\mathbb{Z}^d}$, on $\mathbb{T}^d$ with $d \geq 2$.

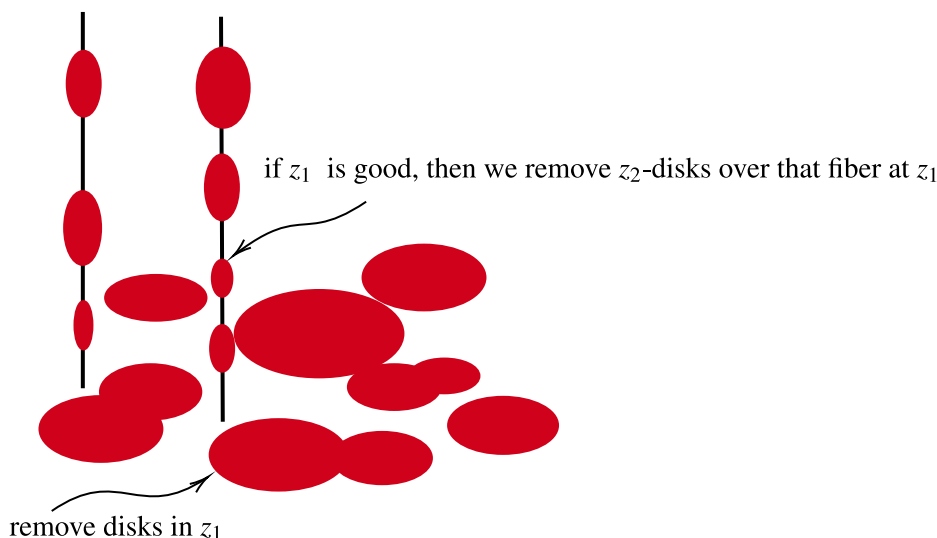
Definition 6.4. Let $0 < H < 1$. For any subset $\mathcal{B} \subset \mathbb{C}$ we define $\mathcal{B} \in \text{Car}_1(H)$ if $\mathcal{B} \subset \bigcup_j \mathbb{D}(z_j, r_j)$ with

$$\sum_j r_j \leq C_0 H. \quad (6.12)$$

If d is a positive integer greater than one and $\mathcal{B} \subset \mathbb{C}^d$, then we define recursively $\mathcal{B} \in \text{Car}_d(H)$ if there exists $\mathcal{B}_0 \in \text{Car}_{d-1}(H)$ so that

$$\mathcal{B} = \{(z_1, z_2, \dots, z_d) \mid (z_2, \dots, z_d) \in \mathcal{B}_0 \text{ or } z_1 \in \mathcal{B}(z_2, \dots, z_d) \text{ with } \mathcal{B}(z_2, \dots, z_d) \in \text{Car}_1(H)\}.$$

We refer to the sets in $\text{Car}_d(H)$ for any d and H summarily as Cartan sets (see Fig. 14).

Fig. 14. Cartan-2 sets in $\mathbb{C}^2$.

The following theorem from [26] furnishes the key property allowing one to extend the previous proof of (6.6) to higher-dimensional shifts. We state the case $d = 2$, with $d > 2$ being similar (see also [43]).

Theorem 6.5. Suppose u is continuous on $\mathbb{D}(0, 2) \times \mathbb{D}(0, 2) \subset \mathbb{C}^2$ with $|u| \leq 1$. Assume further that

$$\begin{cases} z_1 \mapsto u(z_1, z_2) & \text{is subharmonic for each } z_2 \in \mathbb{D}(0, 2) \\ z_2 \mapsto u(z_1, z_2) & \text{is subharmonic for each } z_1 \in \mathbb{D}(0, 2). \end{cases}$$

Fix some $\gamma \in (0, 1/2)$. Given $r \in (0, 1)$ there exists a polydisk $\Pi = \mathbb{D}(x_1, r^{1-\gamma}) \times \mathbb{D}(x_2, r) \subset \mathbb{D}(0, 1) \times \mathbb{D}(0, 1)$ with $x_1, x_2 \in [-1, 1]$ and a set $\mathcal{B} \in \text{Car}_2(H)$ so that

$$|u(z_1, z_2) - u(z'_1, z'_2)| < C_\gamma r^{1-2\gamma} \log \frac{1}{r} \quad \text{for all } (z_1, z_2), (z'_1, z'_2) \in \Pi \setminus \mathcal{B}. \quad (6.13)$$

$$H = \exp(-r^{-\gamma}).$$

This theorem replaces (6.11) in the previous proof. For the sake of completeness, we now also sketch a proof by Fourier series as in [8,10].

Proof (Proof of Proposition 6.3 by Fourier series). For this technique, it is more convenient to view $u(z)$ as a subharmonic function on an annulus around $|z| = 1$. This is based on viewing the periodic analytic potential $V(x)$ as an analytic function of $z = e(x) = e^{2\pi i x}$ instead and then extending analytically to the annulus $\mathcal{A} := \{z \in \mathbb{C} \mid 1 - \delta < |z| < 1 + \delta\}$ for some $0 < \delta < 1$. Thus, write $V(x) = W(e(x))$ with W analytic on that annulus. Accordingly, $u(x) = w(e(x))$, and the Riesz representation takes the form

$$w(z) = \int_{\mathcal{A}} \log|z - \zeta| \mu(d\zeta) + h(z) \quad \forall z \in \mathcal{A} \quad (6.14)$$

with μ a positive measure on $\mathcal{A}$ with $\mu(\mathcal{A}) \lesssim n$ and $\|h\|_{L^\infty(\mathcal{A}')} \lesssim n$ for a slightly thinner annulus $\mathcal{A}'$. Note that $u \geq 0$ on $|z| = 1$. In particular,

$$u(x) = \int_{\mathcal{A}} \log|e(x) - \zeta| \mu(d\zeta) + h(e(x)) \quad \forall x \in \mathbb{T}$$

Next, we claim that

$$f_\zeta(x) := \log|e(x) - \zeta| \text{ satisfies } \sup_{\zeta \in \mathbb{C}} |\widehat{f_\zeta}(k)| \leq C|k|^{-1} \quad \forall k \neq 0 \quad (6.15)$$

with an absolute constant. First, it suffices to prove this for $|\zeta| \leq 1$ by pulling out $\log|\zeta|$ otherwise. By translation in x we may further assume that $0 \leq \zeta \leq 1$. One checks that

$$\begin{aligned} \partial_x f_1(x) &= \partial_x \log|e(x) - 1| = \pi \cot(\pi x), \\ \partial_x f_r(x) &= \partial_x \log|e(x) - r| = \pi \frac{2r \sin(2\pi x)}{1 + r^2 - 2r \cos(2\pi x)} \end{aligned}$$

the latter for $0 \leq r < 1$. These are, respectively, the kernel of the Hilbert transform on $\mathbb{T}$ and the conjugate Poisson kernel. Both have uniformly bounded Fourier coefficients, uniformly in $0 \leq r \leq 1$, whence our claim (6.15). Indeed,

$$\widehat{f_r}(k) = (2\pi i k)^{-1} \widehat{\partial_x f_r}(k) = O(k^{-1})$$

We conclude that $|\hat{u}(k)| \leq Cn|k|^{-1}$ for all $k \neq 0$ by integrating over the Riesz mass. For the harmonic function we simply use that $|\partial_x h(e(x))| \lesssim n$ and the decay of the Fourier coefficients follows. By the almost invariance property (6.9),

$$u(x) - \langle u \rangle = \frac{1}{k} \sum_{j=1}^k u(x + j\omega) - \langle u \rangle + O(k) = \sum_{\nu \neq 0} \hat{u}(\nu) e(x\nu) \frac{1}{k} \sum_{j=1}^k e(j\nu\omega) + O(k)$$

Then one has that

$$\left| \frac{1}{k} \sum_{j=1}^k e(j\nu\omega) \right| \lesssim \min(1, k^{-1} \|\nu\omega\|^{-1})$$

for all $\nu \geq 1$. Also, it follows from (6.10) that $|\hat{u}(\nu)| \lesssim n|\nu|^{-1}$ which in turn implies that

$$|u(x) - \langle u \rangle| \lesssim \frac{1}{k} \sum_{j=1}^k \left| \sum_{|\nu| > K} \hat{u}(\nu) e(\nu(x + k\omega)) \right| + \sum_{0 < |\nu| \leq K} n|\nu|^{-1} \min(1, k^{-1} \|\nu\omega\|^{-1})$$

On the one hand, by Plancherel and the decay of the Fourier coefficients,

$$\left\| \frac{1}{k} \sum_{j=1}^k \sum_{|\nu| > K} \hat{u}(\nu) e(\nu(x + j\omega)) \right\|_{L_x^2} \leq \left\| \sum_{|\nu| > K} \hat{u}(\nu) e(\nu x) \right\|_2 \lesssim n K^{-1/2}$$

On the other hand, setting $K = e^{n^\tau}$ it follows from the Diophantine condition (with $a = 2$ for simplicity) that

$$\sum_{0 < |\nu| \leq K} n|\nu|^{-1} \min(1, k^{-1} \|\nu\omega\|^{-1}) \lesssim n k^{-\frac{1}{2}} \log K \lesssim n^{1+\tau} k^{-\frac{1}{2}} \quad (6.16)$$

Choosing $\tau > 0$ small and $k = n^{\frac{1}{2}}$, say, yields (6.6). To prove (6.16), partition $0 < |\nu| \leq K$ into sets corresponding to the size of $\|\nu\omega\|$. First $\|\nu\omega\| \leq k^{-1}$ and then $\|\nu\omega\| \in k^{-1}(2^{j-1}, 2^j]$ for $j \geq 1$ and $k^{-1}2^j < 1$. The Diophantine condition implies that the recurrences into these sets cannot be more frequent than specific arithmetic conditions, which the reader can easily check. The $\log K$ term results from summing the harmonic series over a finite arithmetic progression. $\square$

Remark 6.6. Write the Diophantine condition in the form $\|k\omega\| \geq h(k)$ for all $k \geq 1$. Later we will need to exploit the fact that the previous proofs require this condition only in the range $1 \leq k \leq n$.

For applications related to the study of fine properties of the IDS it turns out to be important to obtain sharp versions of (6.6). The commutative model example suggests that the optimal relation is $0 \leq 1 - \sigma = \tau \leq 1$. Here $\sigma = 1 - \tau = 0$ corresponds to the largest possible deviations and smallest measures. The previous two proofs do not easily yield such a statement, but it was proved in [26] by a more involved argument. The book [8] contains an elegant Fourier series proof, see Theorem 5.1 on page 25. Both these references require stronger Diophantine conditions.

There is a close connection between the Wegner estimate in Section 3 and the LDT from above. We refer to reader to [27, Lemma 5.5] for the precise formulation of a Wegner estimate derived via LDT for the quasi-periodic model (6.1).

6.2. LDT and regular Green functions

As in Sections 3 and 4 the key to proving localization in Theorem 6.1 is to exclude arbitrarily long chains of resonances (absence of infinite tunneling). In fact, one shows that one cannot have double resonances on sufficiently long scales, in exact analogy with the localization results we proved above. The LDT theorems from above enter into this analysis through the Green function associated with (6.1) on finite intervals. In fact, from Cramer's rule for any $\Lambda = [a, b] \in \mathbb{Z}$, and $a \leq j \leq k \leq b$,

$$(H_\Lambda(x) - E)^{-1}(j, k) = \frac{\det(H_{[a, j-1]}(x) - E) \det(H_{[k, b]}(x) - E)}{\det(H_\Lambda(x) - E)} \quad (6.17)$$

for fixed x, ω , the latter Diophantine. We denote $f_n(x, E) = \det(H_{[1, n]}(x) - E)$. In explicit form, the matrix is

$$H_{[1, n]}(x) - E = \begin{bmatrix} v_1(x) - E & 1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 1 & v_2(x) - E & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & v_3(x) - E & 1 & 0 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & v_n(x) - E \end{bmatrix} \quad (6.18)$$

with $v_j(x) = V(T^j x)$ and $T = T_\omega$. Thus, (6.17) implies that

$$G_{[1, n]}(x, E)(j, k) = (H_{[1, n]}(x) - E)^{-1}(j, k) = \frac{f_{j-1}(x, E) f_{n-k}(T^k x, E)}{f_n(x, E)}, \quad 1 \leq j \leq k \leq n \quad (6.19)$$

with the convention $f_0 = 1$. The transfer matrices defined in (6.3) satisfy for all $n \geq 1$

$$M_n(x, E) = \begin{bmatrix} (-1)^n f_n(x, E) & (-1)^n f_{n-1}(Tx, E) \\ (-1)^{n-1} f_{n-1}(x, E) & (-1)^{n-1} f_{n-2}(Tx, E) \end{bmatrix} \quad (6.20)$$

where we set $f_{-1} = 0$. The following uniform upper bound from [28, Proposition 4.3] improves on the LDT. As expected, as a subharmonic function $\log \|M_n(x, E)\|$ can only have large deviations towards values which are much smaller than $nL_n(E)$ but cannot exhibit deviations in the opposite direction. The following inequality requires positive Lyapunov exponents and relies on some machinery which we have not discussed here, such as the avalanche principle from [26]. Moreover, [28] imposes a Diophantine condition of the form

$$\|n\omega\| \geq \frac{b}{n(\log n)^2}, \quad n \geq 2$$

which holds for some $b > 0$ for a.e. ω . Of course one needs V analytic since the following lemma heavily relies on subharmonic functions and the LDT from above.

Lemma 6.7. Assume $L(E) \geq \gamma > 0$ for all $E \in I$, some interval. For all $n \geq 1$ one has

$$\sup_{x \in \mathbb{T}} \log \|M_n(x, E)\| \leq nL_n(E) + C(\log n)^B,$$

for some absolute constant B and $C = C(V, \gamma, b, I)$.

In view of (6.20) and the Thouless formula (6.5) it is natural to ask if each entry of M_n , i.e., the determinants f_n satisfy an LDT individually. This was proven to hold in [28, Section2].

Proposition 6.8. There exists $\sigma > 0$ so that for large n

$$|\{x \in \mathbb{T} \mid \log |f_n(x, E)| < nL_n(E) - n^\sigma\}| \leq e^{-n^\sigma}$$

A stronger statement is possible if we assume positive Lyapunov exponents. See [27, Lemma 5.1]. Below BMO refers to *bounded mean oscillation*, see e.g. [48].

Proposition 6.9. Assume $L(E) \geq \gamma > 0$ for all $E \in I$. For some constants A and C depending on ω , V , and γ , every $n \geq 1$ satisfies

$$\begin{aligned} \left| \int_0^1 \log |\det(H_{[1,n]}(x) - E)| dx - nL_n(E) \right| &\leq C \\ \|\log |\det(H_{[1,n]}(x) - E)|\|_{\text{BMO}} &\leq C(\log n)^A. \end{aligned} \quad (6.21)$$

Thus,

$$|\{x \in \mathbb{T} \mid |\log |\det(H_{[1,n]}(x) - E)| - nL_n(E)| > H\}| \leq C \exp\left(-\frac{H}{(\log n)^A}\right) \quad (6.22)$$

for any $H > (\log n)^A$. If V is a trigonometric polynomial, then the set on the left-hand side is covered by $2 \deg(V)n$ many intervals each not exceeding in length the measure bound of (6.22).

The final statement follows from the fact that $z^{dn} \det(H_{[1,n]}(x) - E)$ with $z = e(x)$, is a polynomial of degree $2dn$. The estimate (6.22) follows from the BMO bound (6.21) by means of the classical John–Nirenberg inequality. The large deviation estimate for the determinants $f_n(x, E)$ do not appear in the original proof of Theorem 6.1, and they were established later in [28]. However, they help to streamline some of the technical aspects of [10]. For example, in view of (6.19), (6.20) and Lemma 6.7, the Green function satisfies for large n (and of course for positive Lyapunov exponents)

$$\begin{aligned} |G_{[1,n]}(x, E)(j, k)| &\leq \exp((j-1)L_{j-1}(E) + (n-k)L_{n-k}(E) - nL_n(E) + (\log n)^{2A}) \quad \text{provided} \\ |f_n(x, E)| &\geq nL_n(E) - (\log n)^{2A} \end{aligned} \quad (6.23)$$

which therefore holds up to a set of measure $\lesssim \exp(-(\log n)^A)$ (assuming $A \geq B$). It was proved in [26] that $L_n(E) - L(E) \leq Cn^{-1}$ whence it follows from (6.23) that

$$|G_{[1,n]}(x, E)(j, k)| \leq \exp(-|j-k|L(E) + (\log n)^{2A}) \quad (6.24)$$

up to a set of measure $\lesssim \exp(-(\log n)^A)$. By the preceding this set can be made $\lesssim \exp(-n^\sigma)$ with $0 < \sigma < 1$ if we settle for the weaker Green function bound

$$|G_{[1,n]}(x, E)(j, k)| \leq \exp(-|j-k|L(E) + n^\sigma(\log n)^A) \quad \forall j, k \in [1, n] \quad (6.25)$$

and large n . This is precisely the notion of regular Green functions from Section 3. Since

$$\text{dist}(\text{spec}(H_{[1,n]}(x)), E) = \|G_{[1,n]}(x, E)\|^{-1}$$

the connection with a Wegner-type estimate is also immediately apparent. For example, from (6.25) one concludes the following statement. We assume throughout that

$$|E| \leq 2 + \|V\|_\infty \quad (6.26)$$

since this range contains $\text{spec}(H_x)$.

Corollary 6.10. *Under the same assumptions as Proposition 6.9 one has*

$$|\{x \in \mathbb{T} \mid \text{dist}(\text{spec}(H_{[1,n]}(x)), E) < \exp(-n^{1/3})\}| \leq e^{-n^{\frac{1}{4}}} \quad (6.27)$$

for large n . In addition, the set on the left-hand side is contained in $O(n)$ many intervals assuming V is a trigonometric polynomial.

6.3. Eliminating double resonances

We will assume for convenience that V is a trigonometric polynomial. As in the proof of localization in Sections 3 and 4 we begin from a generalized (nonzero) eigenfunction $H(x) = E\psi(x)$ which by Theorem 2.1 grows at most linearly: $|\psi(n)| \leq C(1 + |n|)$. We claim that for any n sufficiently large there exists a window $\Lambda_0 = [-m, m]$ with $n \leq m \leq n^3$ such that $H_{\Lambda_0}(0)$ is resonant with E . Quantitatively, we claim

$$\text{dist}(\text{spec}(H_{[-m,m]}(0)), E) \leq e^{-m^{1/4}} \quad (6.28)$$

Indeed, denote the set in Proposition 6.8 by $\mathcal{B}_n$. It consists of $O(n)$ intervals of length e^{-n^σ} . Therefore, by the Diophantine condition the set

$$\{n \leq m \leq n^2 \mid m\omega \in \mathcal{B}_n(E) \cup (-\mathcal{B}_n(E))\}$$

has cardinality $O(n)$. Pick an $m \in [n, n^2]$ which is not in this set. Then by (6.25)

$$|G_{[m-n, m+n]}(0, E)(m+1, m \pm n)| + |G_{[-m-n, -m+n]}(0, E)(-m-1, -m \pm n)| \leq \exp(-\gamma n/2),$$

whence by (3.28) for large n

$$\sqrt{\psi(m+1)^2 + \psi(-m-1)^2} \leq 2C \exp(-\gamma n/2)(1 + n^2) \leq \exp(-\gamma n/3)$$

Combined with $((H_{[-m,m]}(0)) - E)\psi = \psi(m+1)\delta_{m+1} + \psi(-m-1)\delta_{-(m+1)}$ this estimate implies that

$$\|(H_{[-m,m]}(0) - E)\psi\| \leq \exp(-\gamma n/3) \leq e^{-m^{1/4}}$$

which is what we claimed in (6.28). Let us denote by $\text{Dioph}_n(b)$ the Diophantine condition

$$\|k\omega\| \geq \frac{b}{k(1 + \log k)^2} \quad \forall 1 \leq k \leq n \quad (6.29)$$

and $\text{Dioph}(b) = \bigcap_{n=1}^\infty \text{Dioph}_n(b)$. Then under this condition we have the following stronger LDT for the determinants, see [29, Corollary 2.15]:

Lemma 6.11. *Assume $\omega \in \text{Dioph}_n(b)$ and positive Lyapunov exponents as above. For any E_0 in the range (6.26),*

$$|\{x \in \mathbb{T} \mid \log|f_n(x, \omega, E)| < nL_n(E, \omega) - n^{\frac{1}{2}} \text{ for some } |E - E_0| \leq e^{-n}\}| \leq e^{-n^{\frac{1}{3}}} \quad (6.30)$$

if $n \geq n_0(V, b, \gamma)$ is large.

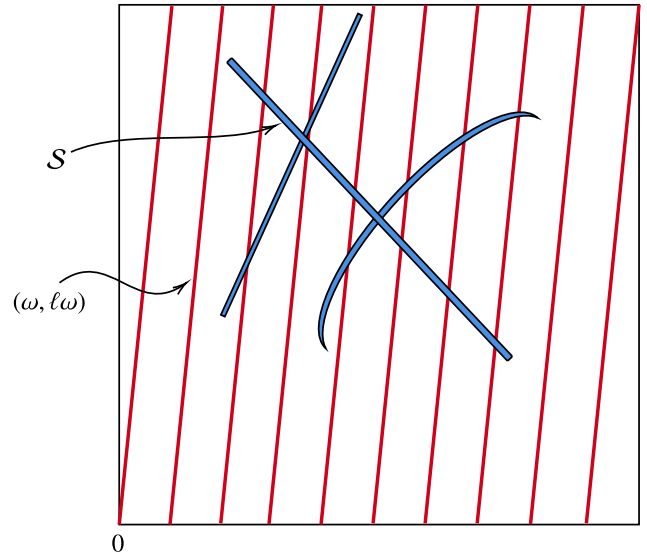


Fig. 15. The lemma on steep lines.

Proof. For fixed E_0 we already stated this LDT for the determinant in Proposition 6.9. The stability in E over the exponentially small interval $[E_0 - e^{-n}, E_0 + e^{-n}]$ is precisely what [29, Corollary 2.15] provides. The statement in loc. cit. is slightly weaker, but replacing the upper bound of [29, Corollary 2.14] with the stronger one of Lemma 6.7 implies (6.30). $\square$

In view of this lemma, and (6.28) we now introduce the following set which will allow us to eliminate double resonances: for any $b > 0$

$$\mathcal{S}_n(b) := \left\{ (\omega, x) \in \text{Dioph}_n(b) \times \mathbb{T} \mid \exists E \in \mathbb{R} \text{ with } \text{dist}(\text{spec}(H_{[-n,n]}(0, \omega)), E) \leq e^{-n^{1/4}}, \text{ and } \log |f_m(x, \omega, E)| \leq mL_m(E, \omega) - m^{1/2} \text{ for some } m \in [n^{1/4}/2, n^{1/4}] \right\} \quad (6.31)$$

If $\text{dist}(\text{spec}(H_{[-n,n]}(0, \omega)), E) \leq e^{-n^{1/4}}$, then $|E - E_{j,n}(\omega)| \leq e^{-n^{1/4}}$ for some eigenvalue $E_{j,n}(\omega)$ of $H_{[-n,n]}(0, \omega)$. Applying Lemma 6.11 with $E_0 = E_{j,n}(\omega)$ and summing over $1 \leq j \leq 2n+1$ one concludes by Fubini that

$$|\mathcal{S}_n(b)| \leq 3n^{\frac{5}{4}} e^{-n^{1/12}}. \quad (6.32)$$

The set of bad ω , which we will need to exclude in order to prevent double resonances, is

$$\mathcal{B}_n(b) := \{ \omega \in \mathbb{T} \mid (\omega, \ell\omega) \in \mathcal{S}_n(b) \text{ for some } \pm \ell \in [n^{2s}, 2n^{2s}] \} \quad (6.33)$$

Here $s \geq 2$ is an absolute constant, which we will specify later. The following *lemma on steep lines* from [10] guarantees that $\mathcal{B}_n(b)$ has very small measure. This hinges not only on the small measure estimate of (6.32), which by itself is insufficient, but also on the structure of the set $\mathcal{S}_n(b)$. Specifically, the fact that the horizontal slices

$$(\mathcal{S}_n(b))_x := \{ \omega \in \mathbb{T} \mid (\omega, x) \in \mathcal{S}_n(b) \} \quad (6.34)$$

are contained in no more than $O(n^s)$ many intervals of very small measure (see Fig. 15).

Lemma 6.12. *Suppose the Borel set $\mathcal{S} \subset \mathbb{T}^2$ has the property that for every $x \in \mathbb{T}$ the horizontal slice $\mathcal{S}_x$, viewed as a subset of $[0, 1]$, consists of no more than M intervals. Then*

$$|\{\omega \in \mathbb{T} \mid (\omega, \ell\omega) \in \mathcal{S} \pmod{\mathbb{Z}^2} \text{ for some } \ell \in [N, 2N]\}| \leq \frac{M}{N} + 8N^{\frac{5}{2}} |\mathcal{S}|^{\frac{1}{2}} \quad (6.35)$$

Proof. By Fubini, for each $\gamma > 0$,

$$|\{x \in \mathbb{T} \mid |\mathcal{S}_x| > \gamma\}| \leq |\mathcal{S}| \gamma^{-1}$$

We define the set of good $x \in \mathbb{T}$ as

$$\mathcal{G} := \{x \in \mathbb{T} \mid |\mathcal{S}_x| \leq \gamma \text{ and for all } j \in [1, N] \text{ one has } \|xj\| > 4N^2\gamma\} \quad (6.36)$$

with $\|\cdot\|$ the norm of $\mathbb{T}$. Then

$$|\mathbb{T} \setminus \mathcal{G}| \leq |\mathcal{S}| \gamma^{-1} + 4N^3\gamma$$

We optimize here by setting $\gamma = (4N^3)^{-\frac{1}{2}} |\mathcal{S}|^{\frac{1}{2}}$ whence

$$|\mathbb{T} \setminus \mathcal{G}| \leq 4N^{\frac{3}{2}} |\mathcal{S}|^{\frac{1}{2}} \quad (6.37)$$

Correspondingly,

$$\mathcal{S} = \mathcal{S}_* \cup \mathcal{S}_{**} := (\mathcal{S} \cap \mathcal{G}) \cup (\mathcal{S} \cap (\mathbb{T} \setminus \mathcal{G}))$$

We eliminate $\mathcal{S}_{**}$ as follows:

$$\begin{aligned} & |\{\omega \in \mathbb{T} \mid (\omega, \ell\omega) \in \mathcal{S}_{**} \pmod{\mathbb{Z}^2} \text{ for some } \ell \in [N, 2N]\}| \\ & \leq \sum_{\ell=N}^{2N} |\{\omega \in \mathbb{T} \mid \ell\omega \in \mathbb{T} \setminus \mathcal{G}\}| \leq (N+1) |\mathbb{T} \setminus \mathcal{G}| \leq 8N^{\frac{5}{2}} |\mathcal{S}|^{\frac{1}{2}} \end{aligned} \quad (6.38)$$

On the other hand, where $\{x\} = x - \lfloor x \rfloor$ for $x > 0$ denotes the fractional part,

$$\begin{aligned} & |\{\omega \in \mathbb{T} \mid (\omega, \ell\omega) \in \mathcal{S}_* \pmod{\mathbb{Z}^2} \text{ for some } \ell \in [N, 2N]\}| \\ & \leq \sum_{\ell=N}^{2N} \int_0^1 \mathbb{1}_{\mathcal{S}_*}(\omega, \{\ell\omega\}) d\omega = \sum_{\ell=N}^{2N} \frac{1}{\ell} \sum_{k=0}^{\ell-1} \int_0^1 \mathbb{1}_{\mathcal{S}_*}((x+k)/\ell, x) dx \\ & \leq \sum_{\alpha=1}^M \frac{1}{N} \int_{\mathcal{G}} \sum_{\ell=N}^{2N} \sum_{k=0}^{\ell-1} \mathbb{1}_{I_{\alpha}(x)}((x+k)/\ell) dx \end{aligned} \quad (6.39)$$

Here, for $x \in \mathcal{G}$, $(\mathcal{S}_*)_x = \mathcal{S}_x = \bigcup_{\alpha=1}^M I_{\alpha}(x)$ with $I_{\alpha}(x)$ intervals of length $|I_{\alpha}(x)| \leq \gamma$, possibly empty. We claim that for all $x \in \mathcal{G}$ one has

$$\sum_{\ell=N}^{2N} \sum_{k=0}^{\ell-1} \mathbb{1}_{I_{\alpha}(x)}((x+k)/\ell) \leq 1 \quad (6.40)$$

Indeed, suppose $\ell \neq \ell'$ both in $[N, 2N]$ and $\frac{x+k}{\ell}, \frac{x+k'}{\ell'} \in I_{\alpha}(x)$. Then

$$\left| \frac{x+k}{\ell} - \frac{x+k'}{\ell'} \right| \leq |I_{\alpha}(x)|$$

whence $|x(\ell - \ell') + k\ell' - k'\ell| \leq \ell\ell'|I_{\alpha}(x)|$ and thus $\|jx\| \leq 4N^2\gamma$ for some $1 \leq j \leq N$. But this is excluded by x being in the good set. So it follows that $\ell = \ell'$, which implies that for $k \neq k'$

$$\gamma \geq |I_{\alpha}(x)| \geq \left| \frac{x+k}{\ell} - \frac{x+k'}{\ell} \right| \geq \frac{1}{\ell} \geq \frac{1}{N}$$

which contradicts that $N^2\gamma^2 = (4N)^{-1} |\mathcal{S}| < 1$. So Claim (6.40) is correct, and the entire contribution to (6.39) is at most M/N . $\square$

To obtain the *complexity bound* of $O(n^s)$ on the set (6.34), we use semi-algebraic methods. A closed set $\mathcal{S} \subset \mathbb{R}^N$ is called *semi-algebraic* if there are polynomials $P_j \in \mathbb{R}[X_1, \dots, X_N]$, $1 \leq j \leq s$ of degrees bounded by d so that

$$\mathcal{S} = \bigcup_k \bigcap_{j \in \mathcal{F}_k} \{P_j \sigma_{kj} 0\}$$

with $\sigma_{kj} \in \{\leq, \geq, =\}$ and $\mathcal{F}_k \subset \{1, 2, \dots, s\}$. The degree of $\mathcal{S}$ is bounded by sd and is in fact the infimum of sd over all such representations.

One might expect to get away with more elementary arguments based on zero counts alone. Note, however, that E is projected out of in the set $\mathcal{S}_n(b)$ which makes it necessary to perform quantifier elimination. In fact, we will need to use a quantitative Seidenberg–Tarski theorem to control the complexity parameter M in Lemma 6.12. This fundamental result states that any projection of $\mathcal{S}$ onto a subspace of $\mathbb{R}^N$ is again semi-algebraic and the degree can only grow at a power rate (depending on N). See [4,5].

These semi-algebraic techniques are available here since V is a trigonometric polynomial although by approximation and truncation, V analytic can also be handled in [10]. Heuristically speaking, the semi-algebraic quantitative complexity bounds replace the explicitly imposed complexity in Theorem 4.1 where exactly two monotonicity intervals of V are assumed.

We claim that $\mathcal{S}_n(b)$ is contained in

$$\begin{aligned} \tilde{\mathcal{S}}_n(b) := \Pi_{\mathbb{R}^2} \{ (\omega, x, E) \in \text{Dioph}_n(b) \times \mathbb{T} \times \mathbb{R} \mid & \log|f_{2n+1}(n\omega, \omega, E)| \leq (2n+1)L_{2n+1}(E, \omega) - n^{1/4}/2, \\ & \text{and } \log|f_m(x, \omega, E)| \leq mL_m(E, \omega) - m^{1/2} \text{ for some } m \in [n^{1/4}/2, n^{1/4}] \} \end{aligned} \quad (6.41)$$

where $\Pi_{\mathbb{R}^2}$ projects on to (ω, x) and moreover, that $\tilde{\mathcal{S}}_n(b)$ has essentially the same measure bound as $\mathcal{S}_n(b)$. And conversely,

$$\begin{aligned} \tilde{\mathcal{S}}_n(b) \subset \{ (\omega, x) \in \text{Dioph}_n(b) \times \mathbb{T} \mid \exists E \in \mathbb{R} \text{ with } \text{dist}(\text{spec}(H_{[-n,n]}(0, \omega)), E) \leq e^{-n^{1/4}/4}, \text{ and} \\ \log|f_m(x, \omega, E)| \leq mL_m(E, \omega) - m^{1/2} \text{ for some } m \in [n^{1/4}/2, n^{1/4}] \} \end{aligned}$$

These relations follow from noting that

$$\text{dist}(\text{spec}(H_{[-n,n]}(0, \omega)), E) = \|((H_{[-n,n]}(0, \omega)) - E)^{-1}\| \quad \text{and} \quad \|A\| \leq \|A\|_{HS} \leq \sqrt{d} \|A\|$$

for any $d \times d$ matrix A , and using the relation (6.25). In particular, we obtain essentially the same estimates on their two-dimensional measure. The sets $\tilde{\mathcal{S}}_n(b) := \Pi_{\mathbb{R}^2} \mathcal{D}_n(b)$ with

$$\begin{aligned} \mathcal{D}_n(b) = \{ (\omega, x, E) \in \text{Dioph}_n(b) \times \mathbb{T} \times \mathbb{R} \mid & |f_{2n+1}(n\omega, \omega, E)| \leq \exp((2n+1)L_{2n+1}(E, \omega) - n^{1/4}/2) \} \\ \cap \bigcup_{m \in [n^{1/4}/2, n^{1/4}]} \{ (\omega, x, E) \in \text{Dioph}_n(b) \times \mathbb{T} \times \mathbb{R} \mid & |f_m(x, \omega, E)| \leq \exp(mL_m(E, \omega) - m^{1/2}) \} \end{aligned} \quad (6.42)$$

are already quite close to our sought after polynomial description. However, a polynomial expression for the Lyapunov exponents in finite volume needs to be found. Note that while we may pass to their infinite volume versions due to the [26] rate of convergence estimate $L_m(E, \omega) - L(E, \omega) \leq Cm^{-1}$, it would be counter productive to do so at this point. Rather, we will use that uniformly in x ,

$$nL_n(E, \omega) = \frac{1}{n^2} \sum_{j=1}^{n^2} \log \|M_n(x + j\omega, \omega, E)\| + O((\log n)^A)$$

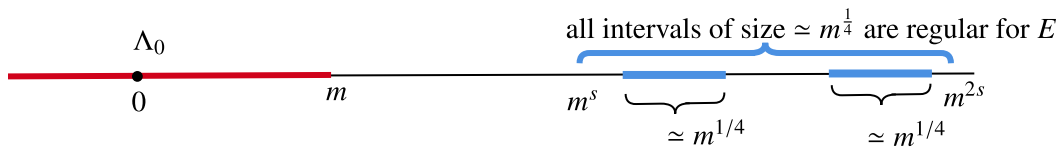


Fig. 16. Absence of double resonances.

This follows from $\|M_n\| \geq 1$, (6.20), and the same arguments which we used in the proof of (6.28). Therefore, we can replace

$$|f_m(x, \omega, E)| \leq \exp(mL_m(E, \omega) - m^{1/2})$$

with

$$|f_m(x, \omega, E)|^{2m^2} \leq e^{-m^{\frac{5}{2}}} \prod_{j=1}^{m^2} \|M_m(j\omega, \omega, E)\|_{HS}^2$$

This is a polynomial inequality in all variables of degree $O(m^4) = O(n)$. The set on the first line of $\mathcal{D}_n$ is described by a polynomial inequality of degree $O(n^4)$. Since there are $\lesssim n^{\frac{1}{4}}$ polynomials involved in the description of the semi-algebraic set $\mathcal{D}_n(b)$ above, it is of degree $\lesssim n^5$. Projecting out E , we conclude that $\tilde{\mathcal{S}}_n(b)$ has degree $O(n^s)$ for some finite s as claimed. Finally, each horizontal slice consists of at most $O(n^s)$ many connected components, i.e., intervals.

Proof of Theorem 6.1. The set of admissible ω for the theorem is

$$\Omega := \text{Dioph} \setminus \bigcup_{j=1}^{\infty} \limsup_{n \rightarrow \infty} \mathcal{B}_n(1/j), \quad \text{Dioph} := \bigcup_{j=1}^{\infty} \text{Dioph}(1/j)$$

where $\mathcal{B}_n(b)$ is defined in (6.33). By Lemma 6.12,

$$|\mathcal{B}_n(1/j)| \leq C(j)(n^{-s} + n^C e^{-n^{1/12}}), \quad \sum_{n=1}^{\infty} |\mathcal{B}_n(1/j)| < \infty$$

whence by Borel–Cantelli $|\limsup_{n \rightarrow \infty} \mathcal{B}_n(1/j)| = 0$. Since Dioph has full measure in $\mathbb{T}$, so does Ω . Now freeze some $\omega \in \Omega$. Note in particular that $\omega \in \text{Dioph}(b)$ for some $b > 0$ whence the LDT results all hold. Given a generalized eigenfunction $H(x)\psi = E\psi$ by Theorem 2.1, we showed that for all sufficiently large n , (6.28) holds for some $n \leq m \leq n^3$. By definition of $\mathcal{B}_n(b)$, see Fig. 16, we conclude that all Green functions $G_A(0, \omega, E)$ with $A \subset [m^s, 2m^s]$ and $|A| \simeq m^{\frac{1}{4}}$ satisfy

$$\|G_A(0, \omega, E)\| \leq e^{|A|^{\frac{1}{2}}}, \quad |G_A(0, \omega, E)(x, y)| \leq e^{-\gamma|x-y|+|A|^{\frac{1}{2}}} \quad \forall x, y \in A$$

Using the resolvent identity iteratively as in Lemma 3.8, albeit with all subintervals being regular for E , we conclude that the Green function on the large window is also regular for E :

$$\|G_{\pm[m^s, 2m^s]}(0, \omega, E)\| \leq e^{m^{\frac{1}{8}}}, \quad |G_{\pm[m^s, 2m^s]}(0, \omega, E)(x, y)| \leq e^{-\gamma|x-y|+m^{\frac{1}{8}}} \quad \forall x, y \in \pm[m^s, 2m^s]$$

from which the exponential decay of ψ immediately follows. $\square$

```

In[404]:= n = 200; epsilon = 0.3; omega = Sqrt[2]; theta = -17 * omega / 2;

In[405]:= Laplace = epsilon * Total[
    {DiagonalMatrix[Array[-1 &, n - 1], 1], DiagonalMatrix[Array[-1 &, n - 1], -1]}};

In[406]:= f = Cos[2 * Pi * (theta + ## omega)] &;

In[407]:= V = DiagonalMatrix[Array[f, n]];

In[408]:= H = Laplace + V;

In[409]:= With[{eigs = Eigenvectors[H]},
    Manipulate[ListLinePlot[eigs[[k]], ImageSize -> 4 * 72, PlotRange -> {-1, 1}],
    {{k, 1}, 1, Length[eigs], 1}]]

```

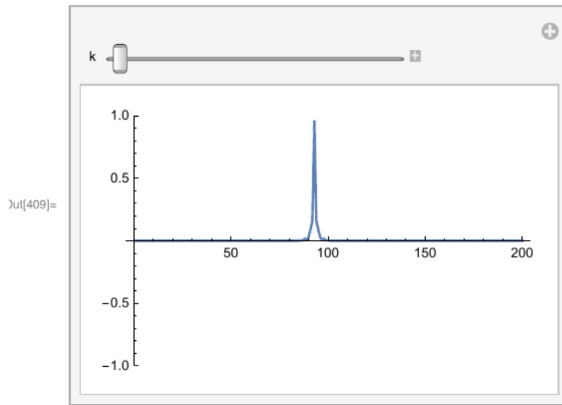


Fig. 17. MATHEMATICA code for Figs. 9 and 10.

References

- [1] M. Aizenman, S. Molchanov, Localization at large disorder and at extreme energies: an elementary derivation, *Commun. Math. Phys.* 157 (2) (1993) 245–278.
- [2] M. Aizenman, S. Warzel, *Random Operators. Disorder Effects on Quantum Spectra and Dynamics*, in: Graduate Studies in Mathematics, vol. 168, American Mathematical Society, Providence, RI, 2015.
- [3] A. Avila, Global theory of one-frequency Schrödinger operators, *Acta Math.* 215 (1) (2015) 1–54.
- [4] S. Basu, R. Pollack, M. Roy, On the combinatorial and algebraic complexity of quantifier elimination, *J. ACM* 43 (6) (1996) 1002–1045.
- [5] S. Basu, R. Pollack, M. Roy, *Algorithms in Real Algebraic Geometry*, Springer-Verlag, Berlin, 2003.
- [6] J. Berezanskii, *Expansions in eigenfunctions of selfadjoint operators*, in: Translations of Mathematical Monographs, Vol. 17, Amer. Math. Soc. Providence, R.I., 1968.
- [7] P. Bougerol, J. Lacroix, Products of random matrices with applications to Schrödinger operators, in: *Progress in Probability and Statistics*, Vol. 8, Birkhäuser Boston, Inc., Boston, MA, 1985.
- [8] J. Bourgain, Green's Function Estimates for Lattice Schrödinger Operators and Applications, in: *Annals of Mathematics Studies*, vol. 158, Princeton University Press, Princeton, NJ, 2005.
- [9] J. Bourgain, Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, d arbitrary, *Geom. Funct. Anal.* 17 (3) (2007) 682–706.
- [10] J. Bourgain, M. Goldstein, On nonperturbative localization with quasi-periodic potential, *Ann. Math. (2)* 152 (3) (2000) 835–879.
- [11] J. Bourgain, M. Goldstein, W. Schlag, Anderson localization for Schrödinger operators on $\mathbb{Z}^2$ with quasi-periodic potential, *Acta Math.* 188 (1) (2002) 41–86.
- [12] V. Bucaj, D. Damanik, J. Fillman, V. Gerbuz, T. VandenBoom, F. Wang, Z. Zhang, Localization for the one-dimensional Anderson model via positivity and large deviations for the Lyapunov exponent, *Trans. Amer. Math. Soc.* 372 (5) (2019) 3619–3667.
- [13] R. Carmona, A. Klein, F. Martinelli, Anderson localization for Bernoulli and other singular potentials, *Commun. Math. Phys.* 108 (1) (1987) 41–66.
- [14] R. Carmona, J. Lacroix, *Spectral Theory of Random Schrödinger Operators*, in: *Probability and Its Applications*, Birkhäuser Boston, Inc., Boston, MA, 1990.
- [15] W. Craig, B. Simon, Subharmonicity of the lyapunov index, *Duke Math. J.* 50 (2) (1983) 551–560.

- [16] D. Damanik, J. Fillman, One-Dimensional Ergodic Schrödinger Operators I. General Theory, in: Graduate Studies in Mathematics, AMS, in press.
- [17] D. Damanik, J. Fillman, One-dimensional ergodic Schrödinger operators II. Special cases, in: Graduate Studies in Mathematics, AMS, in press.
- [18] J. Ding, C. Smart, Localization near the edge for the Anderson Bernoulli model on the two dimensional lattice, *Invent. Math.* 219 (2) (2020) 467–506.
- [19] W. Faris, Perturbations and non-normalizable Eigenvectors, *Helv. Phys. Acta* 44 (1971) 930–936.
- [20] A. Figotin, L. Pastur, Spectra of Random and Almost-Periodic Operators, Springer, 1992.
- [21] Y. Forman, T. VandenBoom, Localization and cantor spectrum for quasiperiodic discrete Schrödinger operators with asymmetric, smooth, cosine-like sampling functions, 2021, preprint.
- [22] J. Fröhlich, T. Spencer, Absence of diffusion in the Anderson tight binding model for large disorder or low energy, *Commun. Math. Phys.* 88 (2) (1983) 151–184.
- [23] J. Fröhlich, T. Spencer, P. Wittwer, Localization for a class of one-dimensional quasi-periodic Schrödinger operators, *Commun. Math. Phys.* 132 (1) (1990) 5–25.
- [24] H. Fürstenberg, H. Kesten, Products of random matrices, *Ann. Math. Stat.* 31 (1960) 457–469.
- [25] F. Germinet, A. Klein, Bootstrap multiscale analysis and localization in random media, *Commun. Math. Phys.* 222 (2) (2001) 415–448.
- [26] M. Goldstein, W. Schlag, Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions, *Ann. Math.* (2) 154 (1) (2001) 155–203.
- [27] M. Goldstein, W. Schlag, On Schrödinger operators with dynamically defined potentials, *Mosc. Math. J.* 5 (3) (2005) 577–612.
- [28] M. Goldstein, W. Schlag, Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues, *Geom. Funct. Anal.* 18 (3) (2008) 755–869.
- [29] M. Goldstein, W. Schlag, M. Voda, On localization and the spectrum of multi-frequency quasi-periodic operators, <http://arxiv.org/abs/1610.00380>.
- [30] A. Gorodetski, V. Kleptsyn, Parametric Fürstenberg theorem on random products of $SL(2, \mathbb{R})$ matrices, *Adv. Math.* 378 (2021) paper no. 107522.
- [31] R. Han, M. Lemm, W. Schlag, Effective multi-scale approach to the Schrödinger cocycle over a skew shift base, *Ergodic Theory Dynam. Syst.* 40 (10) (2020) 2788–2853.
- [32] M. Herman, Une méthode pour minorer les exposants de lyapounov et quelques exemples montrant le caractère local d'un theoreme d'arnold et de moser sur le tore de dimension 2, *Comment. Math. Helv.* 58 (3) (1983) 453–502.
- [33] D. Hundertmark, A short introduction to Anderson localization, in: *Analysis and Stochastics of Growth Processes and Interface Models*, Oxford Univ. Press, Oxford, 2008, pp. 194–218.
- [34] S. Ya Jitomirskaya, Metal–insulator transition for the almost mathieu operator, *Ann. Math.* (2) 150 (3) (1999) 1159–1175.
- [35] S. Jitomirskaya, C.A. Marx, Dynamics and spectral theory of quasi-periodic Schrödinger-type operators, *Ergodic Theory Dyn. Syst.* 37 (8) (2017) 2353–2393.
- [36] S. Jitomirskaya, B. Simon, Operators with singular continuous spectrum. III. Almost periodic schrödinger operators, *Commun. Math. Phys.* 165 (1) (1994) 201–205.
- [37] S. Jitomirskaya, X. Zhu, Large deviations of the Lyapunov exponent and localization for the 1D Anderson model, *Commun. Math. Phys.* 370 (1) (2019) 311–324.
- [38] F. John, Partial differential equations, in: *Applied Mathematical Sciences*, Vol. 1, Reprint of the fourth edition, Springer-Verlag, New York, 1991.
- [39] W. Kirsch, Random Schrödinger operators a course, in: H. Holden, A. Jensen (Eds.), *Schrödinger Operators*, in: *Lecture Notes in Physics*, vol. 345, Springer, Berlin, Heidelberg, 1989.
- [40] P. Kuchment, An overview of periodic elliptic operators, *Bull. Am. Math. Soc.* 53 (3) (2016) 343–414.
- [41] B. Ya Levin, Lectures on entire functions, in: *Translations of Mathematical Monographs*, vol. 150, American Mathematical Society, Providence, RI, 1996.
- [42] W. Magnus, S. Winkler, Hill's equation, Corrected reprint of the 1966 edition, Dover Publications, Inc., New York, 1979.
- [43] W. Schlag, On the integrated density of states for Schrödinger operators on $\mathbb{Z}^2$ with quasi periodic potential, *Comm. Math. Phys.* 223 (2001) 47–65.
- [44] B. Simon, Schrödinger semigroups, *Bull. AMS* 7 (3) (1982).
- [45] B. Simon, T. Wolff, Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians, *Commun. Pure Appl. Math.* 39 (1) (1986) 75–90.
- [46] Ya G. Sinai, Anderson localization for one-dimensional difference Schrödinger operator with quasiperiodic potential, *J. Stat. Phys.* 46 (5–6) (1987) 861–909.
- [47] È.È. Šnol', On the behavior of the eigenfunctions of Schrödinger's equation, *Mat. Sb. (N.S.)* 42 (84) (1957) 273–286, erratum 46 (88) 259.
- [48] Elias M. Stein, Monographs in harmonic analysis, III, in: *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, in: *Princeton Mathematical Series*, vol. 43, Princeton University Press, Princeton, NJ, 1993.
- [49] P. Stollmann, Wegner estimates and localization for continuum Anderson models with some singular distributions, *Arch. Der Math.* 75 (2000) 307–311.
- [50] M. Viana, Lectures on Lyapunov exponents, *Cambridge Studies in Advanced Mathematics*, 145, Cambridge University Press, Cambridge, 2014.
- [51] H. von Dreifus, A. Klein, A new proof of localization in the Anderson tight binding model, *Commun. Math. Phys.* 124 (2) (1989) 285–299.
- [52] F. Wegner, The density of states for disordered systems, *Z. Phys. B* 44 (1981) 9–15.