AN INTRODUCTION TO MULTISCALE TECHNIQUES IN THE THEORY OF
ANDERSON LOCALIZATION. PART I.

WILHELM SCHLAG

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 er-
godic transformation (or of several commuting such transformations on higher-dimensional lattices).
Classical localization theorems by Fröhlich–Spencer for large disorders are presented, both for ran-
dom potentials in all dimensions, as well as even quasi-periodic 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.

Contents
1. Introduction 2
2. Polynomially bounded Fourier basis 2
3. The Anderson model and localization 5
4. The one-dimensional quasi-periodic model 18
  4.1. The Fröhlich-Spencer-Wittwer theorem: even potentials 18
  4.2. The work of Forman and VandenBoom: dropping evenness of \( V \) 33
5. Subharmonic functions in the plane 36
  5.1. Motivation and definition 36
  5.2. Basic properties 37
  5.3. Review of harmonic functions 38
  5.4. Riesz representation of subharmonic functions in \( \mathbb{C} \) 40
  5.5. Cartan’s lower bound 44
6. The Bourgain-Goldstein theorem 46
  6.1. Large deviation theorems 48
  6.2. LDT and regular Green functions 52
  6.3. Eliminating double resonances 55
References 60

The author was partially supported by NSF 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.
1. Introduction

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

\[ H = \Delta_{Z^d} + \lambda V \]

where \( \Delta_{Z^d} \) is the discrete Laplacian on the \( d \)-dimensional lattice and \( V : Z^d \to \mathbb{R} \) a random field with i.i.d. components, and a real parameter \( \lambda \). His pioneering work suggested by physical arguments that for large \( \lambda \), 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 Bloch-Floquet waves, see [Kuc, MagWin]. 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 \( \lambda \). 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 [AizWARi], as well as the more classical treaties [BouLac, FigPas, CarLac] and the forthcoming texts [DamFil1, DamFil2]. 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 \( \omega \) 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 [FitMar].

Bourgain’s book [Bou1] 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 [Bou1] 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 [BouGolSch2], and the higher-dimensional theory as in [Bou2], with applications. This will then hopefully serve to make Chapters 14 through 19 of [Bou1] 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 [Ber] and the survey [Sim]. Results of this type go by the name of Shnol theorem. We follow the argument in [Far]. Throughout, the discrete Laplacian on \( \mathbb{Z}^d \) is defined as the sum over nearest
neighbors, i.e.,

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

(2.1)

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

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

(2.2)

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

**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 measure \([\Sigma]\) of \(H\) there exists \(\psi : \mathbb{Z}^d \to \mathbb{R}\) not identically vanishing with \(H\psi = E\psi\) and \(|\psi(n)| \leq C(d, \sigma, E)(n)^{\sigma}\) for all \(n \in \mathbb{Z}^d\).

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

\[|\langle (H - z)^{-1}(x, y) \rangle| \leq C \exp(-\beta|x - y|) \quad \forall \ x, y \in \mathbb{Z}^d\]  

(2.3)

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 \(\|S_{\beta,j}\|_{2 \to 2} \leq C|\beta|\) 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,

\[\|\delta_{\beta}(H - z)^{-1} \delta_{\beta}\| = \langle (H - z)^{-1} \rangle \leq C e^{-\beta|x - y|}\]

Since the sign of \(\beta\) and the choice of \(j\) are arbitrary, (2.3) follows.

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

\[(H - z)^{-1} : \ell^2_{\sigma}(\mathbb{Z}^d) \to \ell^2(\mathbb{Z}^d)\]

\(^1\)I.e., up to a set of measure 0 relative to the spectral measure of \(H\).
is a Hilbert-Schmidt operator. Here \( \ell^2_\sigma(\mathbb{Z}^d) := w^{-\sigma}_\sigma \ell^2(\mathbb{Z}^d) \). By the spectral theorem there exists a unitary \( U : \ell^2(\mathbb{Z}^d) \to L^2(\mathbb{X}, \mu) \) where \( \mu \) is a \( \sigma \)-finite measure, and \( \phi \in L^\infty(\mathbb{X}) \) real-valued, with 

\[
UHf = \phi Uf \quad \text{for all } f \in \ell^2(\mathbb{Z}^d). 
\]

The \( \mu \)-essential range of \( \phi \) equals \( \Sigma \). The composition 

\[
T := U(H - z)^{-1} = (\phi - z)^{-1}U : \ell^2_\sigma(\mathbb{Z}^d) \to L^2(\mathbb{X}, \mu)
\]

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

\[
\int \sum_{n \in \mathbb{Z}^d} |\kappa(x, n)|^2 w^2_\sigma(n) \mu(dx) < \infty
\]

and 

\[
Tf(x) = \sum_{n \in \mathbb{Z}^d} \kappa(x, n)f(n) \quad \forall f \in \ell^2_\sigma(\mathbb{Z}^d)
\]

The series converges in \( L^2(\mathbb{X}, \mu) \). Define \( \psi_x(\cdot) := (\phi(\cdot) - z_0)\kappa(\cdot, n) \). Then for all \( f \in \ell^2_\sigma(\mathbb{Z}^d) \), and \( \mu \)-a.e. \( x \), 

\[
(Uf)(x) = (\phi(x) - z_0)(Tf)(x) = \sum_{n \in \mathbb{Z}^d} (\phi(x) - z_0)\kappa(x, n)f(n) = \langle \psi_x, f \rangle_{\ell^2_\sigma(\mathbb{Z}^d)}
\]

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

\[
H \psi_x = \phi(x) \psi_x. 
\]

as well as \( \psi_x \not= 0 \). 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 \( \mu \)-a.e. \( x \). It follows that \( H \psi_x = \phi(x) \psi_x \) whence (2.5). Now suppose \( \psi_x \equiv 0 \) for all \( x \in \mathbb{S} \subset \mathbb{X}, \mu(\mathbb{S}) > 0 \). Then for all \( f \in \ell^2(\mathbb{Z}^d) \), this implies that \( Uf = 0 \) \( \mu \)-a.e. on \( \mathbb{S} \), and 

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

But this means that \( U^{*}\chi_\mathbb{S} = 0 \) which contradicts \( \|U^{*}\chi_\mathbb{S}\|^2_{\ell^2} = \mu(\mathbb{S}) > 0 \). To summarize, there is \( \mathbb{N} \subset \mathbb{X} \) with \( \mu(\mathbb{N}) = 0 \) so that for all \( E \in \mathcal{G} := \phi(\mathbb{X} \setminus \mathbb{N}) \) the equation \( H \psi = E \psi \) has a nonzero solution \( \psi \in \ell^2_{-\sigma}(\mathbb{Z}^d) \). 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 \( H = \int_{\lambda \in \mathcal{E}} \mathcal{E}(d\lambda) \). But for any Borel set \( B \subset \mathbb{R} \), \( UE(B) U^{*}g = \chi_{\phi^{-1}(B) \cap \mathbb{N}}g \) on all \( g \in L^2(\mathbb{X}, \mu) \). Since \( \phi^{-1}(\mathbb{R} \setminus \mathcal{G}) \subset \mathbb{N} \) is a \( \mu \)-nullset, we conclude that \( UE(\mathbb{R} \setminus \mathcal{G}) U^{*} = 0 \) whence \( \mathcal{E}(\mathbb{R} \setminus \mathcal{G}) = 0 \). So \( H \psi = E \psi \) has nonzero solution \( \psi \in \ell^2_{-\sigma}(\mathbb{Z}^d) \) spectrally a.e.

□

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

\[ H_\omega = \Delta + V_\omega \]  

(3.1)

on \( \ell^2(\mathbb{Z}^d) \) where \( V_\omega \) 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 \( H_\omega \) 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 \( T : X \to X \) be measure preserving, invertible, and ergodic on \( X \) relative to the probability measure \( \nu \). Then set \( V_x(n) := f(T^n x) \) where \( f : X \to \mathbb{R} \) is measurable and \( \nu \)-essentially bounded, and define \( H_x := \Delta + V_x \) in \( \ell^2(\mathbb{Z}) \). For this 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 \( \nu \)-almost every \( x \in X \).

**Proof.** One has the conjugation

\[ H_{Tx} = U^{-1} \circ H_x \circ U \]  

(3.2)

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

\[ E(B) := \{ x \in X \mid \text{rank}(E_x(B)) > 0 \} \]

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

\[ \Sigma = \mathbb{R} \setminus \bigcup_{a \prec b} [(a, b) \mid \nu(E((a, b))) = 0] \]

where the union is over rational \( a, b \). If any interval \( I \subset \mathbb{R} \) intersects \( \Sigma \), then \( \text{rank}(E_x(I)) > 0 \) for \( \nu \)-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 \( \Sigma \) 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 \( \Sigma_{ac}, \Sigma_{sc}, \text{ and } \Sigma_{pp} \text{ of } \Sigma \text{ such that } \Sigma = \Sigma_{ac} \cup \Sigma_{sc} \cup \Sigma_{pp} \) (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 \( \mu(A) := \langle E_x(A \cap B) f, f \rangle \) 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,

\[ \Sigma_{ac} = \mathbb{R} \setminus \bigcup_{a \prec b} [(a, b) \mid \forall f \in \ell^2(\mathbb{Z}), A \mapsto \langle E_x((a, b) \cap A) f, f \rangle] \]

\( \nu - a.s. \) 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 \( H_{T_x} \), respectively, by the shift, the set
\[
Y(a, b) := \{x \in X | \exists f \in l^2(\mathbb{Z}), A \mapsto \langle E_n((a, b) \cap A)f, f \rangle
\]
has an absolutely continuous component
is \( T \)-invariant and thus \( \nu(Y(a, b)) = 0 \) or \( \nu(Y(a, b)) = 1 \). Hence
\[
\Sigma_{ac} = \mathbb{R} \setminus \bigcup_{a < b}[(a, b) | a, b \in \mathbb{Q}, \nu(Y(a, b)) = 0]
\]
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 \( \nu \)-a.s., \( A \mapsto \langle E_n((a, b) \cap A)f, f \rangle \) 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. \( \Box \)

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 \( \{V_n(n) = f(T^n_0x) | n \in \mathbb{Z} \} \) and
the closure of this set is deterministic and equals \( \Sigma_{pp} \). Moreover, \( \Sigma_{ac} = \Sigma_{sc} = \emptyset \).

Propositions 3.1 and 3.2 apply as stated to the random model \( H_{\omega} \) from above, as the reader is
invited to explore. In fact, on \( l^2(\mathbb{Z}^d) \) we may consider \( d \) measure preserving, invertible, commuting
transformations \( T_j : X \to X \) 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
\( H_\lambda := \Delta_{2^d} + V_x \) with \( V_x(n) = f(T_1^{\alpha_1} \circ T_2^{\alpha_2} \circ \cdots \circ T_d^{\alpha_d})x \) for any \( \bar{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d \) with essentially
the same proofs. See [Kir, FigPas] 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 \( \Sigma \) in Proposition 3.1. Recall that we are assuming bounded support of the
single site distribution.

**Proposition 3.3.** For \( H_\omega \) 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, \( K = \mathbb{R} \setminus \bigcup \{|I | \mu(I) = 0 \} \) 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 n} for all \( n \in \mathbb{Z}^d \). The following holds almost surely: given \( L \geq 1, \epsilon > 0 \), and \( \lambda_1 \in K \),
there exists a cube \( \Lambda \subset \mathbb{Z}^d \) of side length \( L \) such that \( \|V - \lambda_1\|_{\ell^\infty(\Lambda)} \leq \epsilon \). Then with \( \lambda = \lambda_0 + \lambda_1 \),
\[
(H - \lambda)\chi_\Lambda e_\alpha = (V - \lambda_1)\chi_\Lambda e_\alpha + g
\]
with \( \|g\|_2^2 \leq |\partial \Lambda| \leq L^{d-1} \). Here \( \partial \Lambda \) is defined as those \( x \in \Lambda \) which have a nearest neighbor in \( \mathbb{Z}^d \setminus \Lambda \),
and \( |\cdot| \) is the cardinality (or volume). Hence, with the normalized function \( \varphi = \chi_\Lambda e_\alpha |\Lambda|^{-\frac{1}{2}} \)
\[
\|(H - \lambda)\varphi\|_2 \leq \epsilon + L^{-\frac{1}{2}}
\]
which implies that almost surely,
\[
\sup_{\epsilon > 0} \| (H - \lambda - i\epsilon)^{-1} \| = \infty
\]
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 \( l^2(\mathbb{Z}^d) \).

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

It is natural to ask about the probability that any given number \( E \in \mathbb{R} \) comes close to the spectrum of \( H_\Lambda \). In other words, what is
\[
\mathbb{P}( \{ \text{dist}(E, \text{spec}(H_\Lambda)) < \epsilon \} ) = \mathbb{P}( \{ \| (H_\Lambda - E)^{-1} \| > \epsilon^{-1} \} ) ?
\]
The diagonal operator given by the random potential \( V \) alone satisfies
\[
\mathbb{P}( \{ \text{dist}(E, \text{spec}(P_\Lambda VP_\Lambda)) < \epsilon \} ) \leq \mathbb{E} \# \{ n \in \Lambda | V(n) \in (E - \epsilon, E + \epsilon) \}
\]
\[
\leq |\Lambda| \mu( (E - \epsilon, E + \epsilon) ) \leq 2\epsilon |\Lambda| \left( \frac{d\mu}{dx} \right)_\infty \tag{3.4}
\]
where \( \mu \) 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 [Weg].

**Proposition 3.4.** Assume the single site distribution of the random operator (3.1) satisfies \( \mu' = \frac{d\mu}{dx} \in L^\infty(\mathbb{R}) \). Then for all \( E \in \mathbb{R} \),
\[
\mathbb{P}( \{ \text{dist}(E, \text{spec}(H_\Lambda)) < \epsilon \} ) \leq 4\epsilon \left\| \mu' \right\|_\infty |\Lambda| \tag{3.5}
\]
for all cubes \( \Lambda \subset \mathbb{Z}^d \) and \( \epsilon > 0 \).

**Proof.** We will present two proofs. For the first we follow Wegner’s original argument [Weg].

Denote by \( N_\Lambda(x) \) the integrated density of states for the random operator \( H_\Lambda \). To wit, if \( E_1^\Lambda \leq E_2^\Lambda \leq \ldots \leq E_m^\Lambda, m = |\Lambda| \), denote the eigenvalues of \( H_\Lambda \) with multiplicity, then
\[
N_\Lambda(x) = \# \{ 1 \leq j \leq m | E_j^\Lambda \leq x \}
\]
Let \( \varphi \geq 0 \) be a smooth bump function on \( \mathbb{R} \) supported in \([-1, 1]\), and set \( \varphi_\epsilon(x) = \epsilon^{-1} \varphi(x/\epsilon) \).
Normalize so that \( \int_\mathbb{R} \varphi(x) \, dx = 1 \). Then with \( F_{\Lambda,\epsilon} = N_\Lambda * \varphi_\epsilon \) one has
\[
N_\Lambda(E + \epsilon) - N_\Lambda(E - \epsilon) \leq F_{\Lambda,\epsilon}(E + 2\epsilon) - F_{\Lambda,\epsilon}(E - 2\epsilon) = \int_{E - 2\epsilon}^{E + 2\epsilon} F'_{\Lambda,\epsilon}(x) \, dx
\]
interlace. This in turn guarantees that and thus \((3.7)\). Combining \((3.6)\) with \((3.7)\) implies \((3.5)\).

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

Thus, with \([-L, L]\) containing the support of \(\mu\),

\[
\mathbb{P}(\text{dist}(E, \text{spec}(H_\Lambda)) < \varepsilon) \leq - \int_{E-2\varepsilon}^{E+2\varepsilon} \sum_{j \in A} \mathbb{E} \frac{\partial F_{\Lambda,e}}{\partial v_j}(x) \, dx \\
\leq \int_{E-2\varepsilon}^{E+2\varepsilon} \sum_{j \in A} \mathbb{E}' \int_{-L}^{L} \frac{\partial F_{\Lambda,e}}{\partial v_j}(x) \mu'(v_j) \, dv_j 
\]

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

\[
\int_{-L}^{L} \frac{\partial F_{\Lambda,e}}{\partial v_j}(x) \mu'(v_j) \, dv_j \leq \|\mu'\|_\infty \int_{-L}^{L} \frac{\partial F_{\Lambda,e}}{\partial v_j}(x) \, dv_j \\
= \|\mu'\|_\infty (F_{\Lambda,e}(v_j = -L) - F_{\Lambda,e}(v_j = L)) \\
= \|\mu'\|_\infty \int_{\mathbb{R}} (N_\Lambda(v_j = -L, x) - N_\Lambda(v_j = L, x))\varphi(x) \, dx \leq \|\mu'\|_\infty
\]

For the final estimate we use that passing from \(v_j = -L\) to \(v_j = L\) in \(H_\Lambda\) 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 estimate

\[
\mathbb{P}(\text{dist}(E, \text{spec}(H_\Lambda)) < \varepsilon) \leq \mathbb{E}\text{trace} 1_{[E-\varepsilon,E+\varepsilon]}(H_\Lambda) \leq 2\varepsilon \mathbb{E} \text{trace} \text{Im} (H_\Lambda - (E + i\varepsilon))^{-1} \\
\leq 2\varepsilon \mathbb{E} \sum_{n \in \Lambda} \text{Im} \langle (H_\Lambda - (E + i\varepsilon))^{-1}\delta_n, \delta_n \rangle
\]

where we used that

\[
1_{[E-\varepsilon,E+\varepsilon]}(x) \leq \frac{2\varepsilon^2}{\varepsilon^2 + (x - E)^2} = \text{Im} \frac{2\varepsilon}{x - (E + i\varepsilon)}
\]
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 $\varphi$ a unit vector, $\lambda$ a real scalar. From the resolvent identity, for any complex $z$ with $\text{Im} z > 0$,
\[
(A + \lambda \varphi \otimes \varphi - z)^{-1} - (A - z)^{-1} = -\lambda(A - z)^{-1}(\varphi \otimes \varphi)(A + \lambda \varphi \otimes \varphi - z)^{-1}
\]
\[
\langle (A + \lambda \varphi \otimes \varphi - z)^{-1} \varphi, \varphi \rangle - \langle (A - z)^{-1} \varphi, \varphi \rangle = -\lambda\langle (A + \lambda \varphi \otimes \varphi - z)^{-1} \varphi, \varphi \rangle \langle (A - z)^{-1} \varphi, \varphi \rangle
\]
\[
\langle (A + \lambda \varphi \otimes \varphi - z)^{-1} \varphi, \varphi \rangle = [\lambda + \langle (A - z)^{-1} \varphi, \varphi \rangle]^{-1}
\]
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
\[
\langle (H_\Lambda - (E + i \epsilon))^{-1} \delta_n, \delta_n \rangle = [V_\omega(n) + \langle (H_\Lambda^{(n)} - (E + i \epsilon))^{-1} \delta_n, \delta_n \rangle]^{-1}
\]
Writing
\[
\langle (H_\Lambda^{(n)} - (E + i \epsilon))^{-1} \delta_n, \delta_n \rangle^{-1} = -t - is, \quad s > 0
\]
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
\[
\mathbb{E} \text{Im} \langle (H_\Lambda - (E + i \epsilon))^{-1} \delta_n, \delta_n \rangle \leq \mathbb{E} \int_0^1 \frac{s \mu(dx)}{(x - t)^2 + s^2} \leq \pi \|\mu'\|_{\infty}
\]
which in combination with (3.8) yields
\[
\mathbb{P}(\text{dist}(E, \text{spec}(H_\Lambda)) < \epsilon) \leq 2\pi \epsilon \|\mu'\|_{\infty} |\Lambda|
\]
This is slightly worse than the previous proof but the precise constant is irrelevant.

The assumption of bounded density $\mu'$ 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 [Sto].

Anderson localization refers to the following statement.

**Theorem 3.5.** Let $H = \epsilon^2 \Delta_{Z^d} + V_\omega$ where $V_\omega$ is random i.i.d. potential with single site distribution $\mu$ of compact support and of bounded density with $\|\mu'\|_{\infty} \leq 1$. Then there exists $\epsilon_0 = \epsilon_0(d) > 0$ so that for all $0 < \epsilon \leq \epsilon_0$, almost surely $l^2(\mathbb{Z}^d)$ 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 $\epsilon$ 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 \cite{FroSpe} multi-scale analysis on the one hand, and the Aizenman-Molchanov \cite{AizMol} fractional moment method on the other hand. We will sketch the former and refer to \cite{Hun} for an introduction to the latter. A streamlined rendition of the induction-on-scales method of \cite{FroSpe} can be found in \cite{vDrKle}, which also does not require the use of the Simon-Wolff criterion \cite{SimWol}, 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 \cite{GerKle}.

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 \cite{DinSma} for a recent advance on this case in two dimensions and on localization for Anderson Bernoulli. The mobility of the eigenvalues of $H_{\Lambda}$ 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 \cite{CarKleMar}. While these authors rely on the original multi-scale methods of Fröhlich and Spencer, this is avoided in the recent papers \cite{BDFGVWZ}, \cite{GorKle}, and \cite{JitZhu}. The arguments there use the large deviation theorems and the methods of Bourgain, Goldstein \cite{BouGol}, see the final section of these notes.

Before getting into the details, some basic ideas and motivation. Suppose $H$ has an $\ell^2$-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 $\Lambda$ and write (heuristically)

\[
(H_{\Lambda} - (E + i\epsilon))^{-1} \approx \sum_j \frac{\phi_j \otimes \phi_j}{E_j - (E + i\epsilon)}
\]

where the sum extends over all eigenfunctions “supported” in the box $\Lambda$. It should be intuitively clear what this means. Then $|\langle \phi_j \otimes \phi_j, (x, y) \rangle| \lesssim \exp(-\gamma|x - y|)$ with $\gamma > 0$ for those $j$, for which either $x$ or $y$ are in the support of $\phi_j$. 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

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

(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 \( x \in Q_L(y) \) with \( L \geq L \), see Figure 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 \( \mathcal{O}(L^d (2L + 1)^d) \) 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 \( \Lambda \) be any box as in (3.9) and assume

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

Then

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

(3.10)

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

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

\[
G_\Lambda(E)(x, y) = (I + \varepsilon^2(V_\Lambda - E)^{-1}\Delta)^{-1}(V_\Lambda - E)^{-1}(x, y)
= \sum_{\ell=0}^{\infty} (-1)^\ell \varepsilon^{2\ell}\left[(V_\Lambda - E)^{-1}\Delta\right](x, y)(V_\Lambda - E)^{-1}(x)
\]

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 \( |\varepsilon^{2\ell}\left[(V_\Lambda - E)^{-1}\Delta\right](x, y)| \leq (4d\varepsilon^2\delta^{-1})^\ell \varepsilon^\ell \) 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). \( \Box \)

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. \( \Box \)

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 \( \Lambda = H_\Lambda, B = H_{\Lambda'} \oplus H_{\Lambda \setminus \Lambda'} \) viewed as operators on \( \ell^2(\Lambda) \). Then,

\[
(H_\Lambda - E)^{-1} = (H_\Lambda' \oplus H_{\Lambda \setminus \Lambda'} - E)^{-1} - \varepsilon^2(H_\Lambda - E)^{-1}\Gamma_{\Lambda,\Lambda'}(H_\Lambda \oplus H_{\Lambda \setminus \Lambda'} - E)^{-1}
\]

\[
G_\Lambda(E)(x, y) = G_\Lambda'(E)(x, y)1\{y \in \Lambda\} - \varepsilon^2 \sum_{(w', w) \in \delta\Lambda'} G_\Lambda(E)(w, y)G_\Lambda'(x, w')
\]  

(3.11)
We claim that $Q(x)$ are of the form

$$\text{maximal number of steps we can take from boundary}$$

provided $x \in \Lambda'$ 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 $\Lambda'$ inside of $\Lambda$, and $\Gamma_{\Lambda, \Lambda'} = 1_{\partial \Lambda'}$.

**Lemma 3.8.** Let $\Lambda$ 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

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

(3.12)

Suppose $8d(2L_0 + 1)^{d-1} \gamma_0 L_0 \leq \delta_0 \leq 1$. Then

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

(3.13)

provided

$$|\gamma_0 - \gamma_1(1 - \frac{1}{L_0 + 1} - \frac{8L_0}{L_1})^{-1} + \frac{2}{L_0}| \geq \frac{1}{\delta_0} \log \frac{L_1}{L_0} \log \frac{8d(2L_0 + 1)^{d-1}\gamma_0^{-1}}{\delta_0}$$

(3.14)

**Proof.** Pick $x, y \in \Lambda$ with $|x - y| \geq L_1/2$ and set $\Lambda_x = Q_{L_0}(x) \cap \Lambda$. If $\Lambda_x \cap \Lambda' = \emptyset$, 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 (3.12) leads to an expression of the form, with $w_0 = x$, $w_0 = y$,

$$G_{\Lambda}(E, x, y) = (-2^{d-1} e^{2g})^{s+t} \sum_{(w', w) \in \partial \Lambda'(x)} \sum_{(w_2, w) \in \partial \Lambda'(w_1)} \cdots \sum_{(w_j, w) \in \partial \Lambda'(w_{j-1})} \prod_{j=1}^{t} G_{\Lambda'(w_{j-1})}(w_{j-1}, w_j)$$

(3.15)

$$\sum_{(w', w) \in \partial \Lambda'(x)} \sum_{(w_2, w) \in \partial \Lambda'(w_1)} \cdots \sum_{(w_j, w) \in \partial \Lambda'(w_{j-1})} G_{\Lambda}(w_s, \bar{w}_1) \prod_{k=1}^{s} G_{\Lambda'(w_{k-1})}(\bar{w}_{k-1}, \bar{w}_k)$$

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 $\Lambda'$. 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 $s + t \geq 1$ is the minimal positive integer with

$$|x - y| < (s + t)(L_0 + 1) + 4L_0 + 1$$
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 $\Lambda'$. Thus,

$$\xi - 1 < s + t \leq \xi, \quad \xi = \frac{|x - y| - 3L_0}{L_0 + 1}$$  \hspace{1cm} (3.16)

To estimate (3.15), use $|G_{\Lambda}(w_s, \bar{w}_t)| \leq \|G_{\Lambda}\| \leq \delta_1^{-1}$ and $|G_{\Lambda'}(w_{j-1}, w_j')| \leq 4\delta_0^{-1}e^{\gamma_0L_0}$ 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

$$|G_{\Lambda}(E)(x, y)| \leq \delta_1^{-1} (8de^2(2L_0 + 1)^{d-1}\delta_0^{-1}e^{\gamma_0L_0})^{s+t} \leq \delta_1^{-1} (8de^2(2L_0 + 1)^{d-1}\delta_0^{-1}e^{\gamma_0L_0})^{\frac{|x-y|}{L_0+4}} \leq e^{\gamma_1|x-y|}$$  \hspace{1cm} (3.17)

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

$$(8de^2(2L_0 + 1)^{d-1}\delta_0^{-1}e^{\gamma_0L_0})^{\frac{|x-y|}{5n+4}} \leq e^{\gamma_1|x-y|}$$
which then implies \((3.13)\) via \((3.17)\). Taking logarithms, this reduces to
\[
\left[ \gamma_0 - \gamma_1 \left(1 - \frac{1}{L_0 + 1} - \frac{4L_0}{|x-y|}\right)^{-1} + \frac{2}{L_0} \right] \log \frac{1}{\varepsilon} \geq L_0^{-1} \log \left(8d(2L_0 + 1)^d - 1 \right)
\]
The worst case here is \(|x-y| = L_1/2\) which gives \((3.14)\).

\[ \square \]

**Definition 3.9.** Fix any \(x_0 \in \mathbb{R}^d\). Then we define an \(L\)-box \(\Lambda_L(x_0)\) to be \((\gamma, E)\)-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: \(|G_{\Lambda_L(x_0)}(E)(x,y)| \leq 4\delta(L)^{-1}e^{\beta|x-y|} \quad \forall \ x, y \in \Lambda_L(x_0)\) 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 \((\gamma, E)\)-singular if it is not \((\gamma, E)\)-regular.

At the initial scale of the induction, by Corollary [3.7]
\[
\sup_{E \in \mathbb{R}} \mathbb{P}(\exists \Lambda_{L_0}(x_0) \text{ which is } (1, E)\text{-singular}) \leq 4B(L_0)|Q_{L_0}(x_0)|\delta(L_0) =: p_0
\]
(3.18)
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
\[
p_0 = 4(2L_0 + 1)^d(2L_0 + 1)^d \exp(-L_0^\beta)
\]
where \(\beta\) is just chosen here to so that \(\exp(-L_0^\beta) = \delta\) and will in fact be in \((0, 1)\). Corollary [3.7] requires that, where \(\delta_0 := \delta(L_0)\),
\[
4d\varepsilon \leq \delta_0.
\]
(3.19)
Set \(L_1 = \lfloor L_0^\alpha \rfloor\) where \(\alpha > 1\) will also be specified later. By Lemma [3.8]
\[
\begin{cases}
\sup_{E \in \mathbb{R}} \mathbb{P}(\exists \Lambda_{L_1}(x_0) \text{ which is } (\gamma_1, E)\text{-singular}) \leq p_1 \\
p_1 := 4B(L_1)|Q_{L_1}(x_0)|\delta_1 + |Q_{L_1}(x_0)|^2 p_0^2 \\
\gamma_1 := (1 - \frac{1}{L_0 + 1} - \frac{8L_0}{L_0})^2(1 - L_0^{-\beta-1})
\end{cases}
\]
(3.20)
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 \(\delta_1\). 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 \(\delta_1 \leq p_0^2\) and thus \(p_1 \leq 5B(L_1)^2 p_0^2\). Finally, setting \(\gamma_0 = 1\) and \(\gamma_1\) as above in \((3.14)\) 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 (3.19) this holds for $L_0(d, \beta)$ sufficiently large, proving (3.20).

Inductively, define $L_{k+1} = [L_k^\alpha]$. In analogy with (3.20) one has with $\delta_k = \exp(-L_k^\beta)$,

$$\sup_{E \in R} \mathbb{P}(\exists \Lambda_{L_k}(x_0) \text{ which is } (\gamma_k, E)\text{-singular}) \leq p_k$$

$$\gamma_k := 4B(L_k)Q_{L_k}(x_0)\delta_k + |Q_{L_k}(x_0)|p_k^2$$

$$L_k := \gamma_k^{-1}(1 - \frac{1}{L_{k+1}^{\alpha}} - 8L_{k-1}^{-1}L_k^{-\beta - 1})$$

(3.21)

One has $p_k \leq B(L_k)^2(4\delta_k + p_k^2) \leq (2(L_k + 1))^{2d}(4\delta_k + p_k^2)$. On the one hand, of $L_0$ is large enough, then

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

since $L_k \geq (L_0)^{\alpha_k}$ and $\sum_{j \geq 0} (L_j^{-\alpha} + L_j^{-\beta - 1}) = \sigma(1)$ as $L_0 \to \infty$, 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 (3.21),

$$L_k^{m+1}p_k \leq 4L_k^{m+1}B(L_k)^2\delta_k + \{(L_k^{\alpha} + 1)\hat{\gamma}(2L_k + 1)^d p_k\}^2$$

$$\leq 4L_k^{m+1}(2(L_k + 1))^{2d}\delta_k + (L_k^m p_k)^2$$

(3.22)

where the second line holds provided $am/2 + d < m$ which requires $\alpha < 2$, and for $L_0$ large enough. We conclude from (3.22) that $L_k^{m+1}p_k \leq 1$ if $L_0$ is large. Moreover, due to

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

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 $L_{k+1} = [L_k^\alpha]$ for $k \geq 0$. Then for arbitrary $x_0 \in \mathbb{Z}^d$ and $E \in \mathbb{R}$,

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

(3.23)

with $0 < p_k \leq L_k^{-m} \leq L_0^{-md+1}$ 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 (3.23) holds as stated for $k \geq 1$ if we weaken the non-resonance condition in Definition 3.9 to the following one: dist(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 $\delta_1^{-1}$ in (3.13). 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 (3.23) 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 $E_* \in \mathbb{R}$,

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

(3.24)

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 $k - 1 \to k$ with $k \geq 2$, consists of the inequality (dropping $x_0$ for simplicity)
\begin{align*}
\mathbb{P}((\forall \Lambda_{L_k} \text{ which is } (\gamma_k, E)\text{-singular for some } E \in [E_\epsilon - \delta_k/2, E_\epsilon + \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_\epsilon - \delta_k/2, E_\epsilon + \delta_k/2])) \\
+ \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_\epsilon - \delta_k/2, E_\epsilon + \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_\epsilon) \leq \delta_k)) \\
+ \mathbb{P}((\exists \Lambda_{L_k} \text{ which contains two disjoint } L_{k-1}\text{-boxes which are both } \\
(\gamma_{k-1}, E)\text{-singular for the same } E \in [E_\epsilon - \delta_{k-1}/2, E_\epsilon + \delta_{k-1}/2]))
\end{align*}

The final two lines here follow from Lemma \ref{lem:3.8} see also Remark \ref{rem: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. □

This result cannot by itself establish localization, since it only controls the resonance of $H_{\Lambda}$ with a given energy $E$ on a single box $\Lambda$. 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 $k \geq 1$,
\begin{equation}
\mathbb{P}((\text{for some } E \text{ a box } \Lambda_{L_{k}}(x_0) \text{ is nonresonant at } E \text{ but } (\gamma_k, E)\text{-singular})) \leq q_k \tag{3.25}
\end{equation}
where for any $b > 1$ and all $k$, $q_k \leq L_{k+1}^{-b}$ provided $L_0$ is large (and thus $\epsilon$ is small) enough. Here nonresonant is as in Definition \ref{def:3.9} but with $\delta_k/2$. 

Proof. $E$-res stands for nonresonant at energy $E$, $E$-res for resonant at $E$, and sing for singular, Let $\Lambda_1 = \Lambda_{L_{k}}(x_0)$ be $E$-res, i.e., $\text{dist}(\text{spec}(H_{\Lambda_1}), E) \geq \delta_1/2$ but $(\gamma_1, E)$-sing. By Lemmas \ref{lem:3.8} and \ref{lem:3.6} there can be at most one resonant $L_0$-box inside of $\Lambda_{L_{k}}(x_0)$ (here but only here we measure resonance with $\delta_0$ and not $\delta_0/2$). Hence
\begin{align*}
\mathbb{P}((\Lambda_{L_{1}}(x_0) \text{ is } E\text{-res but } (\gamma_1, E)\text{-sing})) \\
\leq \mathbb{P}((\Lambda_{L_{1}}(x_0) \text{ contains two disjoint } L_0\text{-boxes, both } E\text{-res})) \\
\leq \sum_{\Lambda_{L_{1}}(x_0)} \sum_{\Lambda_{L_{0}}^c \subset \Lambda_{L_{1}}(x_0)} \sum_{\Lambda_{L_{0}}^c \subset \Lambda_{L_{1}}(x_0)} \mathbb{P}((\text{dist}(\text{spec}(H_{\Lambda_{L_{0}}}), E_j) \leq 2\delta_0 \text{ for some } E_j \in \text{spec}(H_{\Lambda_{L_{0}}}))) \\
\leq 8B(L_1)B(L_0)^2 |Q_{L_1}|^2 |Q_{L_0}|^2 \delta_0 =: q_1.
\end{align*}

In the third line the energy is eliminated by $\delta_0$-closeness of $E$ to some eigenvalue $E_j$ of $H_{\Lambda_{L_{0}}}$, and the fourth line is Wegner’s estimate. The sum over $\Lambda_{L_{1}}(x_0)$ 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, 
\[
\mathbb{P}(\{\text{for some } E \text{ a box } \Lambda_{L_k} \text{ is } E\text{-nres but } (\gamma_k, E)\text{-sing}\}) \\
\leq \mathbb{P}(\{\text{for some } E \text{ a box } \Lambda_{L_k} \text{ is } E\text{-nres, contains two disjoint } L_{k-1}\text{-boxes, both } (\gamma_{k-1}, E)\text{-sing}\}) \\
\leq \mathbb{P}(\{\text{for some } E \text{ a box } \Lambda_{L_k} \text{ contains two disjoint } L_{k-1}\text{-boxes, one } E\text{-res, the other } (\gamma_{k-1}, E)\text{-sing}\}) \\
+ \mathbb{P}(\{\text{for some } E \text{ a box } \Lambda_{L_k} \text{ contains two disjoint } L_{k-1}\text{-boxes, both } E\text{-nres, but } (\gamma_{k-1}, E)\text{-sing}\})
\]

In analogy with \( k = 1 \) we bound the third line by
\[
\leq \sum_{\Lambda_{L_k}} \sum_{\Lambda_{L_{k-1}} \subseteq \Lambda_{L_k}} \sum_{y_0 \in \Lambda_{L_k}} \mathbb{P}(\{\text{some box } \Lambda'_{L_{k-1}}(y_0) \subseteq \Lambda_{L_k} \setminus \Lambda'_{L_{k-1}} \text{ is } (\gamma_{k-1}, E)\text{-sing}\}) \\
\text{with } |E - E_j| \leq \delta_{k-1}/2, E_j \in \text{spec}(H_{\Lambda'_{L_{k-1}}})
\]

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 \( H_{\Lambda'_{L_{k-1}}(y_0)} \). Hence we may first condition on the random variables in \( H_{\Lambda'_{L_{k-1}}(y_0)} \). The fourth line above is bounded by the inductive assumption and independence, and so it is \( \leq B(L_k)|Q_{L_{k-1}}|^2|Q_{L_{k-1}}|p_{k-1} \). In summary, by Proposition \([3.10]\)
\[
q_k \leq B(L_k)B(L_{k-1})|Q_{L_{k-1}}|\frac{|Q_{L_{k-1}}|p_{k-1}}{L_k} + B(L_k)|Q_{L_{k-1}}|^2 q_{k-1}^2 \\
\leq (2(L_k + 1))^{3d} L_k^{-m} + (2(L_k + 1))^{3d} q_{k-1}^2
\]

and we conclude as for \([3.22]\) 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 \([3.25]\). We remove the 0-probability event
\[
\mathcal{B} = \cup_{x_0 \in \mathbb{Z}^d} \limsup_{k \to \infty} \mathcal{B}_k(x_0).
\]

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 \( |\psi(n)| \leq C(\sigma, \psi)|n|^m \) 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 Proposition \([3.13]\) for those \( k \),
\[
\max_{(y', y) \in \partial \Lambda_{L_k}(x_0)} |G_{\Lambda_{L_k}(x_0)}(E)(x_0, y', y) | \leq e^{\frac{1}{2} L_k}
\]

Then
\[
[H_{\Lambda_{L_k}(x_0)} - E]\psi(y') = -\sum_{(y', y) \in \partial \Lambda_{L_k}(x_0)} \psi(y),
\]
\[
|\psi(x_0)| \leq \sum_{(y', y) \in \partial \Lambda_{L_k}(x_0)} |G_{\Lambda_{L_k}(x_0)}(E)(x_0, y', y')||\psi(y)| \leq C(\sigma, \psi) \sum_{(y', y) \in \partial \Lambda_{L_k}(x_0)} e^{\frac{1}{2} L_k} L_k^\sigma
\]

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
The Fröhlich-Spencer-Wittwer theorem: even potentials.

4.1. The dynamics (rotation) $T_\theta \equiv \theta + \omega \mod 1$ 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 $SL(2, \mathbb{R})$ cocycles, cf. [Via] for a comprehensive exposition of this fundamental result as well as Lyapunov exponents in general. See the recent papers [BDFGVWZ], [GorKle], and [JitZhu] 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

$$H_\theta = \varepsilon^2 \Delta_Z + V_\theta, \quad V_\theta(n) = v(T^n_\theta) \quad \forall \, n \in \mathbb{Z}$$

(4.1)

where $\omega \in \mathbb{T}$ is Diophantine, viz. $\|n\omega\| \geq c_0 n^{-2}$ for all $n \geq 1$ with some $c_0 > 0$. There exists $\varepsilon_0(c_0, v)$ such that for all $0 < \varepsilon \leq \varepsilon_0$ the operators $H_{\theta, \varepsilon}$ exhibit Anderson localization for a.e. $\theta \in \mathbb{T}$.

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 $\theta$, see [JitSim]. As in the previous section, we shall drop the index $\varepsilon$ and simply write $H(\theta)$ for (4.1), and $H_\Lambda(\theta)$ for its finite volume version. It is important to keep track of $\theta$ so we include it in the notation (while in the random case we could drop the $\omega$, the variable in the probability space). Fix $\theta_* \in \mathbb{T}$ and $E_* \in \mathbb{R}$. The singular sites relative to $\theta_*, E_*$ are defined as

$$S_0 = S_0(\theta_*, E_*) := \{ n \in \mathbb{Z} \; | \; |V(\theta_* + n\omega) - E_*| \leq \delta_0 \}$$

$$= \{ n \in \mathbb{Z} \; | \; T^n\theta_* \in V^{-1}([E_* - \delta_0, E_* + \delta_0]) \}$$

(4.2)

Figure 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 $V$ Morse, $\max_{i=1,2} |J_i| \leq C_0(v)\delta_0^{-\frac{1}{2}}$ for all cases. We choose the resonance width $\delta_0 = A_0\varepsilon$ with a large constant $A_0$. We
investigate the structure of $S_0$ by means of the example $V(\theta) = \cos(2\pi \theta)$. If $k, \ell \in S_0$ are distinct, then
\[
|\sin(\pi(k - \ell)\omega)\sin(\pi(2\theta_\ast + (k + \ell)\omega))| \leq \delta_0
\]
which implies for small $\delta_0$ that
\[
m(k, \ell) := \min(||(k - \ell)\omega||, ||2\theta_\ast + (k + \ell)\omega||) \leq 2\sqrt{\delta_0}
\]
The first alternative here, viz. $||(k - \ell)\omega|| \leq 2\sqrt{\delta_0}$ occurs precisely if both $T^k_\omega \theta_\ast$ and $T^\ell_\omega \theta_\ast$ 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
\[
c_0|k - \ell|^2 \leq ||(k - \ell)\omega|| \leq 2\sqrt{\delta_0}, \quad |k - \ell| \geq \delta_0^{-\frac{1}{2}}
\]
Henceforth $\geq$ and $\leq$ will indicate multiplicative constants depending on $c_0, \nu$. On the other hand, $||2\theta_\ast + (k + \ell)\omega|| \leq 2\sqrt{\delta_0}$ 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 $\nu$ as in the theorem.

**Lemma 4.2.** Any two distinct $k, \ell \in S_0$ satisfy $m(k, \ell)^2 \leq \delta_0$, and any three distinct points $k, \ell, n \in S_0$ satisfy
\[
\max(|k - \ell|, |\ell - n|) \geq \delta_0^{-\frac{1}{2}}
\]

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

**Definition 4.3.** We label $S_0 = \{c_i^d\}_{i=\infty}^{\infty}$ in increasing order (assuming $S_0 \neq \emptyset$). These are the singular sites (or singular “intervals”) at level 0. Let $s_0 := \min\{|c_i^d - c_j^d| \mid i > j\}$. 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 $S_0$ with $\bar{S}_0 = S_0 \cup (S_0 + s_0)$, which we again label as $\{c_i^d\}_{i=\infty}^{\infty}$. 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 $c^i_1 := c^0_i$, in the double resonant case $I^1_i$ has length $\ell_1 := \lceil \log(1/\varepsilon) \rceil^4$, centered at $c^i_1 := (c^{2i}_0 + c^{2i+1}_0)/2 \in \frac{1}{2}\mathbb{Z}$. By construction, all of the $I^1_i$ are pairwise disjoint, and each $c^i_1$ is contained in a unique interval at level 1. We classify those intervals $I^1_i$ as singular provided
\[ \text{dist}(\text{spec}(H_i^1(\theta_*)), E_i) \leq \delta_1 := \varepsilon^{2/3} \quad (4.3) \]
and $S_1 := \{ c^i_1 \mid I^1_i \text{ is singular} \}$. 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. Figures 4 and 5 illustrate the two cases, with the blue dots being $\mathbb{Z} \setminus S_0$. The terminology simple/double resonance is derived from the structure of the eigenfunctions at level 1 associated with the operators $H_i^1(\theta_*)$ and the unique (as we shall see) eigenvalue $E^*_i(\theta_*)$ satisfying (4.3). In the simple resonance case, the eigenfunction has most (say 99%) of its $\ell^2$ mass at the center $c^i_0$, whereas in the double resonant case it may have significant mass at both sites $c^{2i}_0$ and $c^{2i+1}_0$.

Figure 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 $T^1_{\omega_1}\theta_*$ to $J_2$, whereas the red dot on the left a return to $J_1$, cf. Figure 5. 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 at least on the order of $\delta_0^{-\frac{3}{2}}$. 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 form $S_0$ to $\tilde{S}_0$ lies with the self-symmetry indicated in Figure 5 (i.e., $c^{2i+1}_0 - c^{2i}_0$ does not depend on $i$). To see this, note that by definition of $s_0$ there exist $k_1, k_2 \in S_0$ with $\theta_* + k_i \omega \in J_1$ 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 $\varepsilon$ needs to be small enough depending on $v$ and $c_0$ so as to guarantee this). Next, take any $k \in S_0$ with $\theta_* + k \omega \in J_1$ (everything modulo integers which will be henceforth understood tacitly). Then $\theta_* + (k + s_0) \omega \in J_2$, where $J_2$ 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 $S_0$ for the construction to work. In fact, Lemma 4.2 remains valid for $\tilde{S}_0$ and the defining inequality (4.2) is modified only slightly, viz. $|V(\theta_* + n \omega) - E_i| \leq \delta_0$ for all $n \in \tilde{S}_0$. 

![Figure 4. Simple resonances at level 1](image-url)
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 $c_1', c_1^j \in S_1$ one has $m(c_1', c_1^j)^2 \leq \delta_1$, with an absolute implied constant.

The idea is to carry out a similar argument as for Lemma 4.2, with the potential function $V$ replaced by the parametrizations of the eigenvalues of the operator $H_{\nu_1}(\theta)$ localized both near $\theta_*$ 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 $H_{\nu}(\theta)$. 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

$$s_n := \min(|c_n^i - c_n^j| \mid c_n^i, c_n^j \in S_n, \ i \neq j)$$

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 $c_{n+1}^i = (c_n^i + c_n^{i+1})/2 \in \frac{1}{2}\mathbb{Z}$, $\ell_{n+1} = \ell_n^4$, and also augment $S_n$ to $\bar{S}_n$ by including the mirror image of each $I_n^s$ if it was not already included in $S_n$. By mirror image we mean the reflection about 0, which is the same as the reflection about $1/2$ modulo $\mathbb{Z}$. By construction, the $I_{n+1}^s$ are pairwise disjoint and each $c_n^i$ is contained in a unique interval at level $n + 1$. An interval $I_{n+1}^s$ centered at $c_n^i$ is called *singular* if

$$\text{dist}(\text{spec}(H_{\nu_1}(\theta_*), E_*)) \leq \delta_{n+1} = \varepsilon_{n+1}^{1/3}$$

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

- An arbitrary interval $\Lambda \subset \mathbb{Z}$ is called $n$-regular provided every point in $\Lambda \cap S_0$ is contained in a regular interval $I_m^s \subset \Lambda$ for some $m \leq n$, cf. Figure 6. Note that every singular point at level 0 is either (i) contained in infinitely many singular intervals $I_n^s$ 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 $\Lambda$: \[ |G_{\Lambda}(\theta, E)(x, y)| \leq \varepsilon^{\frac{1}{2}|x-y|} \] for all $x, y \in \Lambda$, $|x - y| \geq \ell_n^{5/6}$, $|E - E_*| \leq \delta_n$, and $|\theta - \theta_*| \leq \delta_n$. 

**Figure 5.** Double resonances at level 1

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

$$s_n := \min(|c_n^i - c_n^j| \mid c_n^i, c_n^j \in S_n, \ i \neq j)$$

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 $c_{n+1}^i = (c_n^i + c_n^{i+1})/2 \in \frac{1}{2}\mathbb{Z}$, $\ell_{n+1} = \ell_n^4$, and also augment $S_n$ to $\bar{S}_n$ by including the mirror image of each $I_n^s$ if it was not already included in $S_n$. By mirror image we mean the reflection about 0, which is the same as the reflection about $1/2$ modulo $\mathbb{Z}$. By construction, the $I_{n+1}^s$ are pairwise disjoint and each $c_n^i$ is contained in a unique interval at level $n + 1$. An interval $I_{n+1}^s$ centered at $c_n^i$ is called *singular* if

$$\text{dist}(\text{spec}(H_{\nu_1}(\theta_*), E_*)) \leq \delta_{n+1} = \varepsilon_{n+1}^{1/3}$$

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

- An arbitrary interval $\Lambda \subset \mathbb{Z}$ is called $n$-regular provided every point in $\Lambda \cap S_0$ is contained in a regular interval $I_m^s \subset \Lambda$ for some $m \leq n$, cf. Figure 6. Note that every singular point at level 0 is either (i) contained in infinitely many singular intervals $I_n^s$ 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 $\Lambda$: \[ |G_{\Lambda}(\theta, E)(x, y)| \leq \varepsilon^{\frac{1}{2}|x-y|} \] for all $x, y \in \Lambda$, $|x - y| \geq \ell_n^{5/6}$, $|E - E_*| \leq \delta_n$, and $|\theta - \theta_*| \leq \delta_n$. 

**Figure 5.** Double resonances at level 1
• One has

\[ m(c_n^i, c_n^j)^2 \leq |E_n^i(\theta_n) - E_n^j(\theta_n)| \leq \delta_n \quad \forall c_n^i, c_n^j \in S_n \]  

for all \( n \geq 0 \). Here \( E_n^i(\theta_n) \) are the unique eigenvalues of \( H_n(\theta_n) \) in the interval \([E_n - c\delta_n, E_n + c\delta_n]\) with \( c \) small. This hinges crucially on the separation property of the eigenvalues, see Lemma 4.6. For simple resonances, we will use first order eigenvalue perturbation theory, and for double resonances, second order perturbation theory.

• 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 \( \theta \in \mathbb{T} \).

**Figure 6. Regular and singular intervals of 4 levels**

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

**Proof.** At \( n = 0 \) the interval \( \Lambda \) 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),

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

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 \( \Lambda \) is \( 1 \)-regular, then let \( \{I^1_i\}_{i=0}^{i_1} \) be a complete list of all level 1 intervals, in increasing order, which cover all points in \( \Lambda \cap S_0 \). 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 Figure 7; they are only depicted in this way to indicate that they are level 1 intervals. The line segment is supposed to depict \( \Lambda \) and it consists entirely of regular lattice points at level 0 apart from the red singular sites. For the double
resonance case shown in Figure 7, one red pairs is separated from another by \( \delta_0^{-1/4} \approx \epsilon^{-1/4} \), which is much larger than the \( I_1^1 \) which are of length \( (\log \epsilon)^4 \). On the other hand, in the single resonant case recall that the \( I_1^1 \) are of length \( [\log(1/\epsilon)]^2 \), and the separation between the singular sites in \( S_0 \) at least \( 4(\log \epsilon)^2 \) (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 [FroSpeWit] for the details.  

**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 \( E, \psi \) 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 \( (\theta, E) \) provided \( n \geq N \). If \( \Lambda_n \) is \( n \)-regular for infinitely many \( n \), then by the Poisson formula (3.27) for any \( j \) and large \( n \),

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

Taking the limit \( n \to \infty \) yields \( \psi \equiv 0 \), whence our claim. Next, we claim that

\[
\Lambda_n \cap I_1 \neq \emptyset \quad \text{for some singular } I_1
\]

for large \( n \). Since \( \Lambda_n \) is not \( n \)-regular, pick some \( c_0^j \in S_0 \cap \Lambda_n \) (this cannot be empty by Lemma 4.5).

![Figure 7. A 1-regular interval \( \Lambda \)](image)

By the recursive construction of singular intervals, see (4.4) there is the following dichotomy: either, there exists \( 1 \leq m \leq n \)

\[
c_0^j \in I_1^1 \subset \ldots \subset I_m^m
\]

and \( I_m^m \) is regular, or \( m = n \) with \( I_n^m \) singular. If the first alternative occurs for every \( c_0^j \in S_0 \cap \Lambda_n \), then we may slightly enlarge \( \Lambda_n \) to some \( \Lambda'_n \subset [-3\ell_n, 3\ell_n] \) which is \( n \)-regular. This is again impossible for large \( n \) and so (4.6) holds, see Figure 7. Then by (4.5) one has \( m(c_n^j, c_n^{j+1})^2 \leq \delta_n \). But \( \| (c_n - c_n^j) \omega \| \leq \delta_n^{1/2} \) is impossible by the Diophantine condition whence

\[
\| 2\theta + (c_n^j + c_n^{j+1}) \omega \| \leq \delta_n^{1/2}
\]

Given that there are at most \( \leq \ell_n^2 \) many choices of \( c_n^j, c_n^{j+1} \in \Lambda'_n \), it follows that the measure of \( \theta \) as in (4.7) is \( \leq \ell_n^2 \delta_n^{1/2} \). This can be summed, whence by Borel-Cantelli there is a set \( \mathcal{B} \) of measure 0 off of which for large \( n \), \( \Lambda'_n \) 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 $(\theta, E)$ with $\theta \in T \setminus \mathcal{B}$, Lemma 4.5 and (3.27) conclude the proof. It is essential here that $\mathcal{B}$ does not depend on $E$, as evidenced by (4.7).

\[ \square \]

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{Eigenvalue parameterizations for the cosine potential}
\end{figure}

**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 $\theta_* \in T$. Then $S_0 = \{c_i^0\}_{i=-\infty}^{\infty}$, and every $c_i^0$ is contained in a unique level 1 interval $I_i^1$, with $|I_i^1| = \ell_1 = \lceil \log(1/\varepsilon) \rceil^2$. These are pairwise disjoint by construction, and they may be regular or singular. We discard the regular ones and only consider those $c_i^1 = c_i^0$ for which $I_i^1$ is singular. By the definitions,

\begin{align}
|V(\theta_* + k\omega) - E_*| > \delta_0 \quad \forall k \in I_i^1 \setminus \{c_i^0\} \\
|V(\theta_* + c_i^0\omega) - E_*| \leq \delta_0
\end{align}

(4.8)

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

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

(4.9)

Figure 8 exhibits 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 $\varepsilon$. The figure demonstrates how we need to jump

\footnote{The graphs were produced by Yakir Forman at Yale.}
between different translates $V(\theta + k\omega)$ to approximate any given Rellich graph, hence $k(j, \theta)$, which
are not unique near crossing points of $V$ with its own translates by $\omega$. Since $\delta_1 \ll \varepsilon^2 \ll \varepsilon$, (4.3)
and (4.8), (4.9) imply that for all $|\theta - \theta_*| \ll \delta_0$
$$|V(\theta + c_0^1\omega) - E_*| \ll \delta_0$$
$$|V(\theta + k\omega) - E_*| \gtrsim \delta_0 \quad \forall \, k \in \mathbb{I}_1 \setminus \{c_0^1\}$$
(4.10)
with implied absolute constants (depending only on $\nu, \omega$). We now claim that a normalized eigenfunction
$\psi(\theta)$ associated with $H_1(\theta)\psi(\theta) = E_j^{\perp}(\theta)\psi(\theta)$ and $k(j, \theta) = c_j^1 = c_0^1$ satisfies
$$\|P^\perp \psi(\theta)\|_{\mathcal{L}^2(\mathbb{I}_1)} \lesssim \frac{\varepsilon^2}{\delta_0^\perp} \ll \varepsilon$$
(4.11)
where $P^\perp$ denotes the orthogonal projection onto all vectors perpendicular to $\delta_0^\perp$ in $\mathcal{L}^2(\mathbb{I}_1)$. Then
$$0 = \psi(\theta) P^\perp(H_1(\theta) - E_j^{\perp}(\theta))\delta_0 + P^\perp(H_1(\theta) - E_j^{\perp}(\theta))P^\perp \psi(\theta)$$
$$= \varepsilon^2 \psi(\theta)(\delta_{j-1}^0 + \delta_{j+1}^0) + P^\perp(H_1(\theta) - E_j^{\perp}(\theta))P^\perp \psi(\theta)$$
Here and below we use $\delta$ both for the resonance width and in the Dirac sense, without any danger
of confusion. By (4.9) and (4.10),
$$\|\{P^\perp(H_1(\theta) - E_j^{\perp}(\theta))P^\perp\}^{-1}\|_{\mathcal{L}^2(\mathbb{I}_1 \setminus \{c_0^1\})} \lesssim \frac{1}{\delta_0^\perp}$$
(4.12)
which implies (4.11) and
$$|E_j^{\perp}(\theta) - V(\theta + c_0^1\omega)| = |(H_1(\theta)\psi(\theta), \psi(\theta)) - V(\theta + c_0^1\omega)| \lesssim \frac{\varepsilon^2}{\delta_1^0}$$
(4.13)
By first order eigenvalue perturbation (Feynman formula), writing $E = E_j^{\perp}$ and $V(\theta)$ for the multi-
plcation operator by the potential,
$$|E_j^{\perp}(\theta) - V(\theta + c_0^1\omega)| = |(V'(\theta)\psi(\theta), \psi(\theta)) - V'(\theta + c_0^1\omega)| \lesssim \frac{\varepsilon^2}{\delta_1^0}$$
(4.14)
and by the second order perturbation formula, with $G^\perp$ being the resolvent on the left-hand side of (4.12),
$$|E^{(\perp)}'(\theta) - V'(\theta + c_0^1\omega)| = |(V''(\theta)\psi(\theta), \psi(\theta)) - V''(\theta + c_0^1\omega) - 2\psi(\theta), V'(\theta)G^\perp(E(\theta))V'(\theta)\psi(\theta))|$$
$$\lesssim \varepsilon^2 \delta_0^\perp + (\varepsilon^2 \delta_0^\perp)^2 \delta_0^\perp \lesssim \varepsilon^2 \delta_0^\perp \ll \varepsilon$$
(4.15)
The estimates (4.13), (4.15) hold for $|\theta - \theta_*| \ll \delta_0$. We conclude from (4.14), (4.15) that
$$\min_{|\theta - \theta_*| \ll \delta_0} |E^{(\perp)}'(\theta)| + |E^{(\perp)}''(\theta)| \gtrsim 1 \quad \forall \theta_* \in \mathbb{T}$$
(4.16)
Recall that $E(\theta) = E_j^{\perp}(\theta)$ depends on $\theta_*$ and $E_*$, where the latter is chosen so that $S_0 \neq \emptyset$. The
constant in (4.16) is uniform in $\theta_*, E_*$. The reader is invited to compare (4.16) to Figure 8. Now
suppose we have two distinct singular intervals $\mathbb{I}_1^1$ and $\mathbb{I}_1^2$ relative to $(\theta_*, E_*)$. Then the previous
analysis applies to both Rellich functions $E(\theta), E(\theta)$ defined for $|\theta - \theta_*| \ll \delta_0$ characterized by
$$\text{spec}(H_1(\theta)) \cap [E_* - \delta_0/2, E_* - \delta_0/2] = \{E(\theta)\},$$
$$\text{spec}(H_1(\theta)) \cap [E_* - \delta_0/2, E_* - \delta_0/2] = \{\tilde{E}(\theta)\}$$
(4.17)
By (4.3) we have $|E(\theta_\ast) - E_\ast| \leq \delta_1$, $|\bar{E}(\theta_\ast) - E_\ast| \leq \delta_1$. By (4.9) we have

$$|V(\theta_\ast + c_0^l \omega) - E(\theta_\ast)| \leq 2\varepsilon^2, \quad |V(\theta_\ast + c_0^l \omega) - \bar{E}(\theta_\ast)| \leq 2\varepsilon^2$$

whence

$$|V(\theta_\ast + c_0^l \omega) - V(\theta_\ast + c_0^l \omega)| \leq 4\varepsilon^2$$

We showed in Lemma 4.2 that this implies

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

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

$$H_{\bar{I}}(\theta) = H_{I}(\theta + (c_0^l - c_0^l) \omega) \quad (4.19)$$

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

$$|E(\theta_\ast) - E(\theta_\ast + (c_0^l - c_0^l) \omega)| \leq 2\delta_1$$

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

$$\bar{E}(\theta) = E(2\theta_\ast - \theta - \theta_\ast), \quad \theta_\ast := 2\theta_\ast + (c_0^l + c_0^l) \omega \quad (4.20)$$
for all $|\theta - \theta_*| \ll \delta_0$. In terms of Figure 5 this corresponds to $E(\theta)$ being approximated by $V$ over $J_1$, whereas $\bar{E}(\theta)$ 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. Figures 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 $\theta$ or trying rational $\omega$). The choice of $\theta$ was made as a crossing point of $V$ with one of its translates by a multiple of $\omega$, since double resonances occur near those points. The first eigenfunction shown in the upper left of Figures 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, Figure 10 exhibits double resonances 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 $I^I$ be singular (as the red interval on the right-hand side of Figure 5), centered at $c^I_1 = \frac{1}{2}(c^0_0 + c^0_0) \in \frac{1}{2}\mathbb{Z}$. As a side remark, suppose that $c^I_1 \in \frac{1}{2} + \mathbb{Z}$. Then all $c^I_j \in \frac{1}{2} + \mathbb{Z}$ due to $c^0_j - c^0_j = \text{const.}$ for all $j$ (since we passed to $\bar{S}_0$). At the next levels $n = 2, 3, \ldots, N$ we will encounter only simple resonances, and so all $c^k_n \in \frac{1}{2} + \mathbb{Z}$ for all these $n$. If we then encounter a double resonance at $N + 1$, it implies that $c^k_{N+1} \in \frac{1}{2}\mathbb{Z}$, and the pattern...
repeats itself. Continuing with the main argument, one then has

$$|V(\theta_s + c_0^{2i}(\omega) - E_s| \leq \delta_0, \quad |V(\theta_s + c_0^{2i+1}(\omega) - E_s| \leq \delta_0$$

$$|V(\theta_s + k\omega) - E_s| \gg \delta_0^{\frac{1}{2}} \quad \forall \ k \in I_1 \setminus \{c_0^{2i}, c_0^{2i+1}\}$$

where the second line follows from the Diophantine condition since $I_1 = [\log(1/\epsilon)]^4$ (one can choose a larger lower bound such as $\delta_0^a$ for any fixed $0 < a \leq \frac{1}{2}$ at the expense of making $\epsilon$ smaller). By (4.9),

$$\text{spec}(H_{11}(\theta)) \cap [E_\omega - \delta_0^{\frac{1}{2}}, E_\omega + \delta_0^{\frac{1}{2}}] = \{\bar{E}(\theta), \tilde{E}(\theta)\} \quad \forall |\theta - \theta_s| \leq \delta_0^{\frac{1}{2}}$$

(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 $H_{11}(\theta)$ associated with $E$, resp. $\tilde{E}$, are

$$\psi = A\delta_0^{2i} + B\delta_0^{2i+1} + O(\epsilon^2 \delta_0^{-\frac{1}{2}})$$

$$\tilde{\psi} = -B\delta_0^{2i} + A\delta_0^{2i+1} + O(\epsilon^2 \delta_0^{-\frac{1}{2}})$$

(4.22)

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

$$\|P^\perp(H_{11}(\theta) - E_s)P^\perp\|_{C(I_1[\{c_0^{2i}, c_0^{2i+1}\}])} \lesssim \delta_0^{-\frac{1}{2}}$$

(4.23)

where $P^\perp$ is the orthogonal projection perpendicular to $\text{span}(\delta_0^{2i}, \delta_0^{2i+1})$ in $L^2(I_1)$. 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

$$\langle H(\theta)\delta_0^{2i}, \delta_0^{2i} \rangle = E_0^{2i}(\theta), \quad \langle H(\theta)\delta_0^{2i+1}, \delta_0^{2i+1} \rangle = E_0^{2i+1}(\theta)$$

By Lemma 4.2, $\|\delta_0^{\frac{1}{2}}\| \lesssim |\theta_s| \leq \delta_0^{\frac{1}{2}}$ with $\theta_s = 2\theta_s + (c_0^{2i} + c_0^{2i+1})(\omega) = 2(\theta_s + c_1^{\omega})$. It follows that either (a) $|\theta_s + c_1^{\omega}| \leq \delta_0^{\frac{1}{2}}$ or (b) $|\theta_s + \frac{1}{2} + c_1^{\omega}| \leq \delta_0^{\frac{1}{2}}$. These relations show that the unique solution of $E_0^{2i}(\theta) = E_0^{2i+1}(\theta)$ on $|\theta - \theta_s| \leq \delta_0^{\frac{1}{2}}$ is either (a) $\theta_s = -c_1^{\omega}$ or (b) $\theta_s = \frac{1}{2} - c_1^{\omega}$ (henceforth, $\theta_s$ 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 Figure 11. Not only do the segments of the V-graphs (i.e., $E_0^{2i}$ and $E_0^{2i+1}$) intersect at $\theta_s$, but $\tilde{E}, \bar{E}$ have their critical point at $\theta_s$ within the interval $|\theta - \theta_s| \leq \delta_0^{\frac{1}{2}}$. Indeed,

$$H_{11}(\theta + \theta_s) = UH_{11}(-\theta + \theta_s)U$$
where $U$ is the reflection on $\mathbb{Z}$ about $c_1^j \in \frac{1}{2}\mathbb{Z}$. In particular, the eigenvalues are the same. In fact, using (4.21) one concludes that

$$E(\theta + \theta_s) = E(-\theta + \theta_s), \quad \tilde{E}(\theta + \theta_s) = \tilde{E}(-\theta + \theta_s) \quad \forall |\theta - \theta_s| \leq \frac{1}{2}$$

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

Next, we establish the lower bound

$$E(\theta) - \tilde{E}(\theta) > (c\varepsilon)^{5\ell_1/12} \gg \delta_1 = \varepsilon^{\ell_1/3} = \varepsilon^{\log \varepsilon^{1/3}}$$

which follows immediately from this separation lemma, see [FroSpeWit, Lemma 4.1]. This spectral gap is much larger than the resonance width $\delta_1$.

**Lemma 4.6.** Let $H \Lambda \psi_j = E_j \psi_j$, $j = 1, 2$ with nontrivial $\psi_j$. If $\|\psi_j\|_{L^2(\Lambda_0)} \geq \frac{1}{2}\|\psi_j\|_{L^2(\Lambda)}$ for $j = 1, 2$ with $\Lambda_0 \subset \Lambda$ and $|\Lambda_0| \geq 2$, then

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

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 \( \langle H_\Lambda \psi_1, \psi_2 \rangle = \langle \psi_1, H_\Lambda \psi_2 \rangle \) that
\[
2e^2|\bar{v}_1 \wedge \bar{v}_2| = 2e^2|\psi_1(n_0)\psi_2(n_0 - 1) - \psi_2(n_0)\psi_1(n_0 - 1)| = \\
\leq 2|E_1 - E_2|(|\psi_1||_2||\psi_2||_2|^3)(||\psi_1||_{C(\Lambda_0)}||\psi_2||_{C(\Lambda_0)})^\frac{1}{2} \\
\leq 2(Ce^{-2})^{\ell_0} |E_1 - E_2||\psi_1||_{C(\Lambda_0)}||\psi_2||_{C(\Lambda_0)}^\frac{1}{2}
\]
(4.27)
where \( \bar{v}_j = (\psi_{j(n_0)}) \) and \( C = C(v) \). The final estimate is obtained from the transfer matrix representation of the eigenfunctions, viz. for \( n \geq n_0 + 1 \)
\[
\begin{pmatrix}
\psi_j(n) \\
\psi_j(n - 1)
\end{pmatrix} = \prod_{k=n_0}^{n-1} \begin{pmatrix}
\varepsilon^{-2}(v_k - E_j) & -1 \\
1 & 0
\end{pmatrix} \bar{v}_j
\]
and for \( n \leq n_0 - 1 \)
\[
\begin{pmatrix}
\psi_j(n - 1) \\
\psi_j(n)
\end{pmatrix} = \prod_{k=n_0-1}^{n} \begin{pmatrix}
\varepsilon^{-2}(v_k - E_j) & -1 \\
1 & 0
\end{pmatrix}^{-1} \bar{v}_j
\]
On the one hand, with \( B = Ce^{-2} \), and using that \( ||\bar{v}_1 - \bar{v}_2||_2 = 2|\bar{v}_1 \wedge \bar{v}_2|, \)
\[
||\psi_1 - \psi_2||_{C(\Lambda_0)} \leq 2B^{\ell_0}||\bar{v}_1 \wedge \bar{v}_2| + \ell_0 B^{\ell_0}|E_1 - E_2| \\
\leq 2e^{-2}B^{\ell_0}|E_1 - E_2||\psi_1||_{C(\Lambda_0)}||\psi_2||_{C(\Lambda_0)}^\frac{1}{2} + \ell_0 B^{\ell_0}|E_1 - E_2|
\]
On the other hand, with \( a_j := ||\psi_j||_{C(\Lambda)} \),
\[
\sqrt{a_1^2 + a_2^2} \leq \frac{1}{2} (a_1 + a_2) + ||\psi_1 - \psi_2||_{C(\Lambda_0)} \\
\leq \frac{1}{2} (a_1 + a_2) + B^{\ell_0}(B \sqrt{a_1 a_2} + \ell_0)|E_1 - E_2| \\
\leq \frac{1}{2} (a_1 + a_2) + B^{\ell_0}(B + \ell_0)|E_1 - E_2| \sqrt{a_1 a_2}
\]
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). □

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 \( \varepsilon \) is small enough). This is immediate from the Diophantine condition due to \( s_0 \leq 4(\log \varepsilon)^2 \), cf. Figure 3. In particular, \( |V'(\theta)| \gg \delta_0^\frac{1}{2} \) on the interval \( |\theta - \theta_0| \leq \delta_0^\frac{1}{2} \). 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 \( \theta = \theta_0 \) it follows that \( A^2(\theta_0) - B^2(\theta_0) = O(\varepsilon^2 \delta_0^{-\frac{1}{2}}) = O(\varepsilon) \). Due to \( A^2 + B^2 = 1, |A(\theta)|^2 = 1/2 - O(\varepsilon) \) and \( |B(\theta)|^2 = \)
$1/2 - O(\varepsilon)$. In fact, the same argument shows that $|A(\theta)| \approx |B(\theta)| \approx 1$ for all $|\theta - \theta_*| \leq \delta_0^4$ with $|\partial_\theta E(\theta)| \leq \delta_0^4$.

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

$$E_0^{2i}(\theta) - E(\theta) = O(\varepsilon^2\delta_0^{-3}), \quad E_0^{2i+1}(\theta) - E(\theta) = O(\varepsilon^2\delta_0^{-2})$$

(4.28)

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

$$|E(\theta) - \bar{E}(\theta)| \leq \varepsilon^2\delta_0^{-\frac{3}{2}}$$

(4.29)

for all $|\theta - \theta_*| \leq \delta_0^4$ with $|\partial_\theta E(\theta)| \leq \delta_0^4$. The final step in our analysis is to establish a lower bound on $|\partial_\theta^2 E(\theta)|$ and $|\partial_\theta^2 \bar{E}(\theta)|$ for those $\theta$. This hinges on the second order perturbation formulas (suppressing $\theta$ as argument)

$$\partial_\theta^2 E = \langle \psi, V'' \psi \rangle - 2\langle V' \psi, G(E)^{1/2}V' \psi \rangle$$

$$\partial_\theta^2 \bar{E} = \langle \bar{\psi}, V'' \bar{\psi} \rangle - 2\langle V' \bar{\psi}, G(\bar{E})^{1/2}V' \bar{\psi} \rangle$$

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

$$\langle V' \psi, G(E)^{1/2}V' \psi \rangle = \frac{\langle \psi, V' \bar{\psi} \rangle^2}{E - \bar{E}} + \langle V' \psi, G(E)^{1 plus}V' \psi \rangle$$

where $G(E)^{1 plus} = P_{\psi}[P_{\psi}(H_{I_1} - E)P_{\psi}]^{-1}P_{\psi}$. By (4.21), $\|G(E)^{1 plus}\| \leq \delta_0^{-\frac{3}{2}}$. On the other hand, by (4.22)

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

whence from $|AB| \approx 1$,

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

Combining this with (4.29) we obtain

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

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

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

(4.30)

for all $|\theta - \theta_*| \leq \delta_0$ with $|\partial_\theta E(\theta)| \leq \delta_0^4$. The exact same analysis applies to $\bar{E}$. 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 $S_0 = \{e_i^{j}\}_{j=-\infty}^\infty$ with $e_i^{j+1} - e_i^j = s_0$ for all $i$. We have $|V(\theta + k\omega) - E_*| \leq \delta_0^{\frac{3}{2}}$ for $k \in \{e_i^{j}, e_i^{j+1}\}$, $|V(\theta + k\omega) - E_*| \gg \delta_0^{\frac{1}{2}}$ for all
k \in I_1 \setminus \{c_0^i, c_0^{2i+1}\}$, both for all $|\theta - \theta_i| \leq \delta_0^{1/2}$. Here $I_1 \subset \mathbb{Z}$, $|I_1| \approx (\log e)^4$, centered at $c_i \in \frac{1}{2}\mathbb{Z}$ and $\text{dist}(\text{spec}(H_{i1}(\theta_i)), E_\ast) \leq \delta_1 = \varepsilon^{1/3}$.

- $\text{spec}(H_{i1}(\theta)) \cap \{E_\ast - \delta_0^{1/3}, E_\ast + \delta_0^{1/3}\} = \{E(\theta), \tilde{E}(\theta)\}$ (with $E > \tilde{E}$) for these $\theta$, with all other eigenvalues being separated from $E_\ast$ by $\gg \delta_0^{1/3}$. From the level-0 estimate $m(c_i^2, c_0^{2i+1}) \leq \delta_0$, either $\theta_\ast = -c_1^i \omega$ or $\theta_\ast = \frac{1}{2} - c_1^i \omega$ satisfy $||\theta_\ast - \theta_i|| \leq \delta_0^{1/2}$ and the unique critical points of $E, \tilde{E}$ in this interval are at $\theta_\ast$. There is a spectral gap of size $E(\theta) - \tilde{E}(\theta) > e^{\delta_1^{1/2}}$.

- For every $|\theta - \theta_i| \leq \delta_0^{1/2}$ one has either both $|\partial_\theta E(\theta)| \gg \delta_0^{1/2}$ and $|\partial_\theta \tilde{E}(\theta)| \gg \delta_0^{1/2}$ (large slopes), or both $|\partial_\theta E(\theta)| \leq \delta_0^{1/2}$ and $|\partial_\theta \tilde{E}(\theta)| \leq \delta_0^{1/2}$ (small slopes). This follows from $|\partial_\theta E_\ast(\theta_\ast)| \gg \delta_0^{1/2}$, and the first order eigenvalue perturbation formulas

$$\partial E(\theta) = (A^2(\theta) - B^2(\theta))\partial_\theta E_\ast^2(\theta_\ast) + O(\delta_0^{2/3})$$

$$\partial \tilde{E}(\theta) = (-A^2(\theta) + B^2(\theta))\partial_\theta \tilde{E}_\ast^2(\theta_\ast) + O(\delta_0^{2/3})$$

for all $|\theta - \theta_i| \leq \delta_0^{1/2}$, cf. (4.22).

- If the small slope alternative occurs, then $|A(\theta)| \approx |B(\theta)| \approx 1$ 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 $\gg \delta_0^{-2}$, see (4.30). This means that the intervals of small slopes around the critical points at $\theta_\ast$ are of size $\ll \delta_0$.

- Figure 11 depicts the situation for a double resonance: $E$ reaches its minimum, resp. $\tilde{E}$ its maximum, at $\theta_\ast$. The spectral gap is the smallest at this point and the quantitative estimates above hold. In particular, this gap is much larger than $\delta_1$, whence exactly one of $E$ or $\tilde{E}$ achieve the resonance condition (4.3) at $\theta_\ast$.

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

$$||1/2 + c_1^i \omega|| + ||1/2 + c_1^i \omega|| \leq \delta_0^{1/2}$$

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

$$H_{i1}(\theta) = H_{i1}(\theta + (c_1^i - c_1^i)\omega) \text{ and } E_{i1}(\theta) = E_{i1}(\theta + (c_1^i - c_1^i)\omega), \quad \tilde{E}_{i1}(\theta) = \tilde{E}_{i1}(\theta + (c_1^i - c_1^i)\omega)$$

Finally, by (4.3), either

$$|E_{i1}(\theta_\ast) - E_{i1}(\theta_\ast + (c_1^i - c_1^i)\omega)| \leq \delta_1 \text{ or } |\tilde{E}_{i1}(\theta_\ast) - \tilde{E}_{i1}(\theta_\ast + (c_1^i - c_1^i)\omega)| \leq \delta_1$$

By the bounds derived above on the first and second derivatives on $E_{i1}$ etc. and elementary calculus, we finally conclude that $m(c_1^i, c_1^i) \leq \delta_1^{1/2}$. Indeed, in the large slopes case, $m(c_1^i, c_1^i) = ||(c_1^i -
c_1'(\omega) \leq \delta_0^{-\frac{1}{2}} \delta_1$, whereas in the small slopes case, $m(c_1', c_1') \leq \delta_0^0 \delta_1^2$. This is slightly better than what Lemma 4.4 claims, and we are done. □

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 [ProSpeWit]. Note, however, that Lemma 5.2 in loc. cit. erroneously sets $\theta_\ast = -c_m \alpha$ 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 $T$.

It seems very difficult to approach quasi-periodic localization in more general settings by relying on eigenvalue parametrization, 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 [ForVan]. We now sketch the proof of their result.

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

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

(4.31)

where $\omega \in T$ is Diophantine, viz. $||n\omega|| \geq c_0 n^{-2}$ for all $n \geq 1$ with some $c_0 > 0$. There exists $\varepsilon_0(c_0, V)$ such that for all $0 < \varepsilon \leq \varepsilon_0$ the operators $H_{\theta, \varepsilon}$ exhibit Anderson localization for a.e. $\theta \in T$.

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 $\theta_\ast, E_\ast$. However, the $m(k, \ell)$ function is no longer useful, as $||2\theta + (k + \ell)\omega||$ is no longer small if $T_\omega^k \theta_\ast$ and $T_\omega^\ell \theta_\ast$ fall into different connected components of $V^{-1}([E_\ast - \delta_0, E_\ast + \delta_0])$. Without symmetry, no such function $m$ can be defined to be independent of $E_\ast$.

Instead, we divide the energy axis into several overlapping intervals, and we construct a collection $E_1$ of well-separated Rellich functions of certain Dirichlet restrictions of $H$ whose domains cover the circle $T$ with the same structural properties as $E_0$, cf. Figure 12. We choose an initial interval length $t_1^{(1)}$ and consider energy regions of size $O((t_1^{(1)})^{-16})$. Each energy region can be characterized as double-resonant, if it contains some $E_\omega$ which satisfies $E_\omega = V(\theta_\omega) = V(\theta_\omega + n\omega)$ for some $\theta \in T$ and $|n| \leq t_1^{(1)}$, or simple-resonant if it does not. Each function $E_1 \in E_1$ is a Rellich function of $H_{\lambda_1}$, where $\Lambda_1 \subset \mathbb{Z}$ is an interval of length $t_1^{(1)}$ if the energy region is simple-resonant, or $t_1^{(2)} \approx (t_1^{(1)})^2$ if the energy region is double-resonant. The singular intervals are then characterized by

$$S_1 = \{\Lambda_1 + m | m \in \mathbb{Z}, |E_1(\theta_\ast + m\omega) - E_\ast| < \delta_1\}$$

where $E_1 \in E_1$ is the Rellich function defined in the energy region containing $E_\ast$, and $\delta_1$ is defined as Fröhlich, Spencer, and Wittwer define it.

---

\[3\]The remainder of this subsection was written by Forman and VandenBoom.
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 $\|P^\perp(H_{\Lambda,n} - E_1)P^\perp\|$, 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 $E_n$ at each scale. Since the energy regions at scale $s$ are of size at least $O(\delta_{s-2}^3)$, each energy region at scale $s-1$ gives rise to at most $O(\delta_{s-2}^3)$ Rellich functions at scale $s$; thus, we inductively bound $|E_n| \leq O(\delta_n^{-4})$. The bad set of $\theta$ at scale $n$ for a specific $E_n \in E_n$ is bounded in measure by $\ell_{n+1}^2 \delta_n^{1/4}$ 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 $E_{n+1}$ inherit the structural properties of those in $E_n$: namely, a Morse condition and a uniform separation estimate. By construction, simple resonant Rellich functions are well-separated from others, so they satisfy $\|E_{n+1} - E_n\|_{C^2} \ll \delta_n$ 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 $E_n$ of $H_{\Lambda,n}$ resolve as a pair of uniformly locally separated Morse Rellich functions $E_{n+1,\vee} > E_{n+1,\wedge}$ of $H_{\Lambda,n+1}$ with at most one
critical point, cf. Figure 13. The size of the gap is larger than the next resonance scale:

$$\inf E_{n+1,\wedge} - \sup E_{n+1,\vee} \gg \delta_{n+1}$$

This gap ensures that any Rellich function $E_n$ can resonate only with itself at future scales, which ultimately enables our induction.

**Figure 13.** The resolution of a double resonance of $E_0 = V$ into a pair of uniformly locally well-separated Rellich curves of a Dirichlet restriction $H_{\Lambda_1}$ of $H$. The curves $E_0(\theta)$ and $E_0(\theta + n\alpha)$ need not interlace the Rellich curves $E_1$, but the auxiliary curves $\lambda, \tilde{\lambda}$ (not pictured) must.

To prove Lemma 4.8, we interlace two auxiliary curves between the double-resonant Rellich pair. Specifically, let $E_{n}(\theta), \tilde{E}_{n}(\theta)$ 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 $P_{\psi}^t H_{\Lambda_{n+1}} P_{\psi}^\perp$ satisfying

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

Moreover, since we have projected away from one resonance, the arguments from the simple-resonance case can be used to show that $\|\tilde{\lambda} - \bar{E}_n\|_{C^1} \ll \delta_n$. As a consequence of the Morse condition, $|\partial_\theta \bar{E}_n| \gg \delta_n$, so $|\partial_\theta \tilde{\lambda}|$ is similarly bounded below. By repeating this process to construct an eigenvalue $\lambda$ of $P_{\tilde{\psi}}^t H_{\Lambda_{n+1}} P_{\tilde{\psi}}^\perp$ with $\|\lambda - E_n\|_{C^1} \ll \delta_n$, we construct two curves, with large opposite-signed first derivatives, which separate $E_{n+1,\vee}$ and $E_{n+1,\wedge}$. 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 section 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 $\Omega$. What sort of function is $\log |f(z)|$ for $f \in \mathcal{H}(\Omega)$ with $f \neq 0$? Recall that for $f \in \mathcal{H}(\Omega)$ if $f \neq 0$ in $\Omega$ 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 $f' = g'e^f$ 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 $\Omega$ is simply connected. The upshot of this is that for non-vanishing $f$, $\log |f| = \log e^{\Re g} = \Re g$ so that $\log |f|$ is harmonic. Notice that this is still true if $\Omega$ is not simply connected because being harmonic is a local property and we can always find the existence of such a $g$ in a disc 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 zeroes 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 $\Omega'$. 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 zeroes of $f$, and $-\infty$ there, so the value of the function should be lower than its average on a small disc. 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

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

$$u(x_0) \leq \frac{1}{\mathcal{H}(x_0, r)} \int_{\mathcal{H}(x_0, r)} u(y) \, dy$$

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

---

4Based on notes written and typed by Adam Black during a graduate class by the author at Yale.
One should think of subharmonic functions as lying below harmonic ones, see Corollary 5.9 below. 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 $f : K \to [-\infty, \infty)$ be usc with $K$ compact. Then $f$ attains its maximum.

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

### 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 $u(z_0) \leq \frac{1}{\partial D(z_0, r)} u(z) \, dz$ for all $D(z_0, r) \subset \Omega$.

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

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

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

**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 $D(z_0, r) \subset \Omega$

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

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

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

so taking limsups 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. \hfill \Box

**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

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

where $\sigma$ is the surface measure on the circle. We compute

$$\partial_r(Mu)_{x_0}(r) = \int_{|w|=1} \nabla u(x_0 + rw) \cdot w \, \sigma(dw) = \frac{1}{|\partial D(0, 1)|r} \int_{\partial D(x_0, r)} \nabla u(y) \cdot \hat{n} \, \sigma(dw)$$
which by the divergence theorem is equal to
\[ \frac{1}{|\partial \mathbb{D}(0,1)|} \int_{\mathbb{D}(x_0, r)} \Delta u(y) \, dy \]

Thus, we see that if \( \Delta u \geq 0 \) then \((Mu)_{x_0}(r)\) is non-decreasing with \( r \), and as its limit as \( r \to 0 \) 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
\[ \lim_{n \to \infty} \int_{\partial \mathbb{D}(z_0, r)} f_n(z) \, dz = \int_{\partial \mathbb{D}(z_0, r)} f(z) \, dz \]

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 \( u \in SH(\Omega) \) with \( \Omega \) 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 \( f(z) \leq \int_{\partial \mathbb{D}(z, r)} f(w) \, dw \) 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 \( u \in SH(\Omega) \). If \( v \) is harmonic on \( \overline{\Omega'} \subset \Omega \) for \( \Omega' \) bounded and \( v \geq u \) on \( \partial \Omega' \) then \( v \geq u \) in \( \Omega' \).

**Proof.** The function \( u - v \) is subharmonic so that if \( u - v > 0 \) in \( \Omega' \) 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 [Joh]. 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 \( u, v \in C^2(\Omega) \):
\[
\int_{\Omega} u(\zeta) \Delta v(\zeta) \, d\zeta = \int_{\Omega} \Delta u(\zeta) v(\zeta) \, d\zeta + \int_{\partial \Omega} (u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n}) \, d\sigma \tag{5.1}
\]
If \( v = G(z, \zeta) \) is such that (in the sense of distributions) \( \Delta G(z, \zeta) = \delta_z(\zeta) \) and \( G(z, \zeta) = 0 \) for \( z \in \partial \Omega \) then
\[
 u(z) = \int_\Omega G(z, \zeta) f(\zeta) m(d\zeta) + \int_{\partial \Omega} \frac{\partial G}{\partial n}(z, \zeta) g(\zeta) \sigma(d\zeta) \tag{5.2}
\]
with \( m \) Lebesgue measure in the plane and \( \sigma \) surface measure on the boundary. Such a Green function \( G(z, \zeta) \) exists for any bounded domain \( \Omega \) for which \( \partial \Omega \) satisfies an exterior cone condition. This is a standard application of Perron’s method, see [Joh] (this method applies to any dimension).

For the case of a disk \( D(0, R) \subset \mathbb{C} \), there is the explicit formula given by the logarithm of the absolute value of the conformal automorphism of the disk:
\[
 G(z, \zeta) = \frac{1}{2\pi} \log |z - \zeta| + \frac{1}{2\pi} \log \left| \frac{R}{R^2 - \zeta^2} \right| \tag{5.3}
\]
In particular, by (5.1) a harmonic function on \( \Omega \) which is \( C^2(\bar{\Omega}) \) with boundary values \( g \) is given by
\[
 u(z) = \int_{\partial \Omega} \frac{\partial G}{\partial n}(z, \zeta) g(\zeta) \sigma(d\zeta) \tag{5.4}
\]
This is Poisson’s formula and \( P_\zeta(z) = \frac{\partial G}{\partial n}(z, \zeta) \) is the Poisson kernel of \( \Omega \). If \( g \in C(\partial \Omega) \), then (5.4) defines a harmonic function in \( \Omega \) which is the unique solution of the boundary value problem (uniqueness by the maximum principle). For the disc of radius \( r \) in the plane we have
\[
 P_\zeta(z) = \frac{1}{2\pi} \frac{r^2 - |z|^2}{|z - \zeta|^2}
\]
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 disc by its value at the center.

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

**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 \( \Omega \) has a subsequence which converges to some harmonic \( u \) uniformly on each compact subset.

**Proof.** If \( u(z) \) is harmonic on \( D(a, r) \) then taking derivatives of (5.4) shows that \( |D^p u(a)| \leq \frac{C_p}{r^p} \) for some universal constant \( C_p \). 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 $f \in H(\Omega)$ 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 $\Omega'$ 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 $\Omega'$. 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 $\mu$. We begin with some basic properties of logarithmic potentials of such measures.

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

- $u \in SH(\Omega)$
- $u > -\infty$ (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 $D \Subset \Omega$ a disc of radius $R$. Then with $m$ the Lebesgue measure in $\mathbb{R}^2$,

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

by Fubini-Tonelli because the integrands are bounded from above. In fact, $\log |z - \zeta|$ is Lebesgue integrable on $D$:

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

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

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

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

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

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

**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 \( u \in \mathcal{S}H(\Omega) \) where \( \Omega \) is a bounded domain. Then there exists a sequence \( u_n \in \mathcal{S}H(\Omega_{1/n}) \cap C^\infty(\Omega_{1/n}) \) where \( \Omega_{1/n} := \{ z \in \Omega \mid \text{dist}(z, \partial \Omega) > 1/n \} \) such that \( u_n \to u \) pointwise and monotone decreasing (in \( \Omega_{1/n} \) 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 \( \int_{\mathbb{R}^2} \varphi(x) \, dx = 1 \). 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

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

with \( e(\theta) = e^{2\pi i \theta} \), the final inequality implied by the smvp. First, \( \nu(\zeta) := \int_0^1 u(z - \zeta e(\theta)) \, d\theta \) 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, \( u_n(z) \leq \max_{|z - \zeta| \leq 1/n} u(w) \) for \( z \in \Omega_{1/n} \) so that by use \( u_n(z) \to u(z) \) as \( n \to \infty \).

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

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

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

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

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

\[
u_n \leq M_n \text{ on } \overline{D}(0, 4), \quad \nu_n(0) \geq m
\]

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

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

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

\[
\langle u_n, \varphi \rangle \to \langle u, \varphi \rangle = \int_{\overline{D}(0, 3)} u(x + iy)\varphi(x + iy) \, dx \, dy
\]

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

\[
\mu(\overline{D}(0, 3)) \leq C_0(M - m)
\]
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
\[
\lim_{n \to \infty} \int_{\mathbb{D}(0,3)} \psi(\zeta) \mu_n(d\zeta) = \int_{\mathbb{D}(0,3)} \psi(\zeta) \mu(d\zeta)
\]
which implies that
\[
\left\langle \int_{\mathbb{D}(0,3)} \log |z - \zeta| \mu_n(d\zeta), \varphi \right\rangle \to \left\langle \int_{\mathbb{D}(0,3)} \log |z - \zeta| \mu(d\zeta), \varphi \right\rangle.
\]
(5.6)

By the theorem in the smooth case,
\[
\limsup_{n \to \infty} \|h_n - M\|_{L^\infty(\mathbb{D}(0,3 - \delta))} \leq C_0(\delta)(M - m)
\]
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 \( \varphi \in C(\mathbb{D}(0,3)) \) of compact support, \( \langle h_n, \varphi \rangle \to \langle h, \varphi \rangle \) 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
\[
u(z) = \int_{\mathbb{D}(0,3)} \log |z - \zeta| \mu(d\zeta) + h(z) \quad \text{almost everywhere in } \mathbb{D}(0,3),
\]
which in turn implies equality everywhere by Corollary 5.4. Finally, to obtain the desired form we write
\[
\int_{\mathbb{D}(0,3)} \log |z - \zeta| \mu(d\zeta) = \int_{\partial \mathbb{D}(0,3)} \log |z - \zeta| \, d\sigma(\zeta) + \int_{\mathbb{D}(0,3)} \log |z - \zeta| \, \mu(d\zeta)
\]
and notice that \( \int_{\partial \mathbb{D}(0,3)} \log |z - \zeta| \, d\sigma(\zeta) \) 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 \( L^\infty \) 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)
\[
u(z) = \int_{\mathbb{D}(0,4)} G(z, \zeta) \Delta u(\zeta) \, m(d\zeta) + \int_{\partial \mathbb{D}(0,4)} \frac{\partial G}{\partial n}(z, \zeta) u(\zeta) \, d\sigma(\zeta)
\]
(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
\[
u(z) = \int_{\mathbb{D}(0,3)} \log |z - \zeta| \mu(d\zeta) + \int_{\mathbb{D}(0,3)} \frac{4}{16 - |z|^2} \mu(d\zeta) + \int_{\mathbb{D}(0,4) \setminus \mathbb{D}(0,3)} G(z, \zeta) \Delta u(\zeta) \, d\zeta + h_0(z)
\]
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 $16 - |z|^2 \neq 0$ and the third term because $\zeta \in \mathbb{D}(0,4) \setminus \mathbb{D}(0,3)$ and thus

$$h(z) = \int_{\mathbb{D}(0,3)} \log \frac{4}{16 - |z|^2} \mu(d\zeta) + \int_{\mathbb{D}(0,4) \setminus \mathbb{D}(0,3)} G(z, \zeta) \Delta u(\zeta) d\zeta + h_0(z)$$

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

$$u(0) = \int_{\mathbb{D}(0,4)} G(0, \zeta) \mu(d\zeta) + h_0(0) = \int_{\mathbb{D}(0,4)} \log \frac{|\zeta|}{4} \mu(d\zeta) + h_0(0)$$

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

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 $|z| \leq 3 - \delta$. \hfill \Box

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 $u \in SH(\overline{\mathbb{D}(0,R_1)})$ 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 $\mu \in M^+(\mathbb{D}(0,R_2))$ 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 $\|h - M\|_L^\infty(\mathbb{D}(0,R_1)) \leq C_1 (M - m)$.

See Theorem 2.2 in [HanLemSch] 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 [Lev] has much more on this topic, see page 76.

**Theorem 5.17.** Let \( \mu \) 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 \( \{D(z_j, r_j)\}_{j=1}^J \), for \( 1 \leq J \leq \infty \) with \( \sum_{j=1}^J r_j \leq 5H \) and

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

**Proof.** Let \( z \in \mathbb{C} \) be a good point if \( n(z, r) := \mu(D(z, r)) \leq pr \) 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 \( \{D(z_j, r(z_j))\}_{j=1}^J \) are pairwise disjoint and

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

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

\[
  u(z) \geq \int_0^1 \log r \, d(n(z, r)) = - \int_0^1 \frac{n(z, r)}{r} \, dr \geq - \int_0^H p \, dr + \|\mu\| \log H
\]

\[
  = -pH + \|\mu\| \log H = \|\mu\| \log(H/e)
\]

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 \( \{D(z_j, r_j)\}_{j=1}^J \), for \( 1 \leq J \leq \infty \) with \( \sum_{j=1}^J r_j^\delta \leq 5^\delta H \) and

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

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 \( \dim(\{z \in \mathbb{C} \mid u(z) = -\infty\}) = 0 \) 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 \( \Omega \) for any subharmonic function on \( \Omega \) 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

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

Given $0 < H < 1$, there exist disks $D(z_j, r_j)$, $1 \leq j \leq J$, with $\sum_{j} r_j \leq 5H$ and $|P(z)| \geq (H/e)^N$ for all $z \in \mathbb{C} \setminus \bigcup D(z_j, r_j)$. By the maximum principle, each disk contains a zero of $P$. 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 [Lev, 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 $\zeta_j$ of radius $\rho = 1/N$. Then for any $z$ with $z = \zeta_j + \rho e(\theta)$ we have

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

(5.11)

It follows from the maximum (minimum) principle for analytic functions that $|P(z)| \geq 3 - e$ for all $z \in \mathbb{C} \setminus \bigcup_{j=1}^{N} D(\zeta, 1/N)$. 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 from the one for discs 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 $\Omega$ with $\sup_{\Omega} u \leq M$. Let $K \subset \Omega$ be compact and suppose $\sup_{K} u \geq m > -\infty$. For any $\Omega_2 \Subset \Omega_1 \Subset \Omega$, there exist a positive measure $\mu$ on $\Omega_1$ and a harmonic function $h$ on $\Omega_1$ such that

$$u(z) = \int_{\Omega_1} \log |z - \zeta| \mu(d\zeta) + h(z) \quad \forall z \in \Omega_1$$

(5.12)

$$\mu(\Omega_1) \leq C_1(\Omega, K, \Omega_2)(M - m)$$

$$||h - M||_{L^\infty(\Omega_2)} \leq C_2(\Omega, K, \Omega_1, \Omega_2)(M - m)$$

**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 $\Omega_1$ 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 \( D(z_j, \delta) \subset \Omega, 0 \leq j \leq N, \) with \( z_N = z \), and \( z_j \in D(z_{j-1}, \delta/2) \) for all \( j \geq 1 \). Moreover, we may assume that \( \Omega_2 \subset \bigcup_{z \in \Omega_1} D(z, \delta/2) \) and by compactness this can be chosen as a finite union. By Riesz’s representation as in Theorem 5.16 we have
\[
    u(z) = \int_{D(z_0, \delta/2)} \log |z - \zeta| \mu(d\zeta) + h_0(z) \quad \forall \ z \in D(z_0, \delta/2)
\]
(5.13)
\[
    \mu(D(z_0, \delta/2)) \leq C_0(\delta)(M - m), \quad \|h_0 - M\|_{L^\infty(D(z_0, \delta/2))} \leq C_0(\delta)(M - m)
\]
By Harnack’s inequality, (5.15) and (5.16) imply that
\[
    \int_{D(z_0, \delta/2)} \log |z - \zeta| \mu(d\zeta) + h_0(z) \geq h(z) \quad \forall \ z \in D(z_0, \delta/2)
\]
(5.17)

Next, apply Theorem 5.17 to the logarithmic potential in (5.13) with \( h \). We may repeat this procedure finitely many times to cover all of \( \Omega_1 \) and hence everywhere on \( \Omega \). Thus,
\[
    h(z) \geq u(z) - \log(\text{diam}(\Omega_1)) \mu(\Omega_1) \geq u(z) - C(M - m)
\]
with the same type of constant as before. By the previous Cartan estimate and chaining argument, we can find \( z_{\ast \ast} \in D(z_0, \varepsilon_0/4) \) which satisfies a bound (5.14) with a purely geometric constant. Hence
\[
    h(z_{\ast \ast}) \geq m - C(M - m)
\]
(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 \( |z - z_\ast| = \varepsilon_1 \) one has
\[
    \int_{\Omega_1} \log |z - \zeta| \mu(d\zeta) \geq -C(M - m)
\]
whence
\[
    h(z) \leq M + C(M - m) \quad \forall \ |z - z_\ast| = \varepsilon_1
\]
(5.16)

By Harnack’s inequality, (5.15) and (5.16) imply that \( h \) satisfies the desired bound on \( D(z_\ast, \varepsilon_0/2) \) and hence everywhere on \( \Omega_2 \).

Alternatively, one can rely the proof strategy of Theorem 5.15 and use the Green function on general subdomains of \( \Omega \) 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 [BouGol]. Similar to Theorem 4.1 it addresses Anderson localization for the operators
\[
    (H_{x, \omega} \psi)_n = \psi_{n-1} + \psi_{n+1} + V(T^n_{\omega} x) \psi_n
\]
(6.1)
on \( \ell^2(\mathbb{Z}) \), where \( T_\omega : \mathbb{T} \to \mathbb{T} \) is the rotation \( x \mapsto x + \omega \mod 1 \) and \( V : \mathbb{T} \to \mathbb{R} \) 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 [BouGol, Bou1] 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 $\Delta_{Z^d}$ with $d \geq 2$, and less as a review of [BouGol] itself. We will often drop $\omega$ 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 $\omega$. For $V(x) = \cos(2\pi x)$, Jitomirskaya proved [Jit] 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” $\omega$ in addition to a measure 0 set of non-Diophantine $\omega$. The smallness condition on $\varepsilon$ 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) $T : X \rightarrow X$ with an ergodic transformation on a probability space $(X, \nu)$, and $V$ is a real-valued measurable function. Define

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

where $M_n$ are the transfer matrices

$$M_n(x, E) = \prod_{k=n}^{1} \begin{bmatrix} E - V(T^k x) & -1 \\ 1 & 0 \end{bmatrix}$$  \hspace{1cm} (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 $\lim_{n \to \infty} \frac{1}{n} a_n = \inf_{n \geq 1} \frac{1}{n} a_n$ 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 [Her, HanLemSch] for an example of this. But this circle of problems will not concern us here. It was shown by Fürstenberg and Kesten [FurKes], later generalized in Kingman’s subadditive ergodic theorem, that

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

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

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

(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 \( N \to \infty \). In other words, there exists a deterministic nondecreasing function \( N \) so that for \( \omega \) a.e. \( x \in X \) one has

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

where \( E_j^{(\Lambda)}(x) \) are the eigenvalues of \( H_j^\Lambda \), the restriction of (6.1) to \( \Lambda \) with Dirichlet boundary conditions. The existence of this limit holds in great generality, see [FigPas]. 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 \( \frac{dN}{dx} \) 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 [Avi].

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 [GolSch1] 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 \( \omega \) so that \( \|n\omega\| \geq b \, n^{-a} \) for all \( n \geq 1 \).

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

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

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

(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:

\[ u(x) = \sum_{k=1}^q \log |e(x) - e(k\omega)| \]  

(6.7)

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

\[ \| \{ x \in \mathbb{T} : u(x) < \lambda \} \| = \| \{ x \in \mathbb{T} : |e(x) - 1| < e^\lambda \} \| \]  

(6.8)

which is of size \( e^\lambda \). 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

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

(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

\[
u(z) = \sup_{||v||=1} \log |\langle M_n(z, E)v, \bar{w} \rangle|\]

First, \( \log |\langle M_n(z, E)v, \bar{w} \rangle| \) is subharmonic by analyticity of \( \langle M_n(z, E)v, \bar{w} \rangle \). 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 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 \( \mu \) on \( R \) and a harmonic function on \( R \) such that

\[
\nu(z) = \log |M_n(z, E)| = \int_R \log |z - \zeta| \mu(d\zeta) + h(z) \tag{6.10}
\]

Since \( |M_n(z)| \leq e^{cn}, 0 \leq \nu(z) \leq n \) on \( R \) (with a constant that depends on \( V, R \) and \( E_0 \)) and thus \( ||\mu|| \leq n \) as well as \( ||h||_{L^{\infty}(R')} \leq n \), 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 \( D_0 = D(x_0, n^{-2\delta}) \), \( x_0 \in [0, 1] \) with the property that \( \mu(D_0) \leq n^{1-2\delta} \). Write

\[
\int_R \log |z - \zeta| \mu(d\zeta) = \nu_1(z) + \nu_2(z) = \int_{D_0} \log |z - \zeta| \mu(d\zeta) + \int_{\mathbb{C} \setminus D_0} \log |z - \zeta| \mu(d\zeta)
\]

Set \( D_1 = D(x_0, n^{-3\delta}) \). Then

\[
|\nu_2(z) - \nu_2(z')| \leq n^{-1-3\delta} \quad \forall z, z' \in D_1
\]

since

\[
|\nu_2(z) - \nu_2(z')| = \left| \int_{\mathbb{C} \setminus D_0} \log \left| 1 + \frac{z' - \zeta}{z - \zeta} \right| \mu(d\zeta) \right| \leq \frac{n^{-3\delta}}{n^{-2\delta}} \mu(\mathbb{C}) \leq n^{1-\delta}
\]

Cartan’s theorem applied to \( \nu_1(z) \) yields disks \( \{D(z_j, r_j)\} \) so that \( \sum r_j \lesssim \exp(-2n^\delta) \) and with the property that

\[
\nu_1(z) \gtrsim -n^{1-\delta} \quad \forall z \in \mathbb{C} \setminus \bigcup_j D(z_j, r_j)
\]

From \( \nu_1 \leq 0 \) on \( D_1 \) and \( |h(z) - h(z')| \leq n|z - z'| \) on \( R' \), it follows that

\[
|\nu(z) - \nu(z')| \lesssim n^{1-\delta} \quad \forall z, z' \in D_1 \setminus \bigcup_j D(z_j, r_j) \tag{6.11}
\]

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

\[
x + k\omega, x' + k'\omega \in D_1 \quad \mod \mathbb{Z}
\]

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
$\bigcup_j \mathcal{D}(z_j, r_j)$ we need to remove a set $\mathcal{B} \subset \mathbb{T}$ of measure $\leq \exp(-n^\delta)$. For this step is is important that Cartan controls the sum of the radii, i.e., $\sum_j r_j \leq \exp(-2n^\delta)$ since then the disks remove at most measure $\leq \exp(-2n^\delta)$ from the real line. Then from the almost invariance (6.9), for any $x, x' \in \mathbb{T} \setminus \mathcal{B}$,

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

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

This proof generalizes to other types of dynamics such as higher-dimensional shifts $T x = x + \omega \mod \mathbb{Z}^d$, 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 $\mathcal{B} \subset \bigcup_j \mathcal{D}(z_j, r_j)$ with

$$\sum_j r_j \leq C_0 H.$$  

(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

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

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

The following theorem from [GolSch1] furnishes they 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 [Sch]).

![Figure 14. Cartan-2 sets in $\mathbb{C}^2$](image)
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
\[
\begin{align*}
\{ z_1 \mapsto u(z_1, z_2) \text{ is subharmonic for each } z_2 \in \mathbb{D}(0, 2) \\
\{ z_2 \mapsto u(z_1, z_2) \text{ is subharmonic for each } z_1 \in \mathbb{D}(0, 2).
\end{align*}
\]

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
\[
|u(z_1, z_2) - u(z_1', z_2')| \leq C_r r^{1-2\gamma} \log \frac{1}{r} \text{ for all } (z_1, z_2), (z_1', z_2') \in \Pi \setminus \mathcal{B}.
\]
(6.13)

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 [Bou1, BouGol].

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

with $\mu$ a positive measure on $\mathcal{A}$ with $\mu(\mathcal{A}) \leq n$ and $\|h\|_{L^\infty(\mathcal{A})} \leq n$ for a slightly thinner annulus $\mathcal{A}'$. Note that $u \geq 0$ on $|z| = 1$. In particular,
\[
u x \in \mathbb{T}
\]

Next, we claim that
\[
f_{\zeta}(x) := \log |e(x) - \zeta| \text{ satisfies } \sup_{\zeta \in \mathbb{C}} |f_{\zeta}(k)| \leq C|k|^{-1} \quad \forall k \neq 0
\]
(6.15)

with an absolute constant. First, it suffices to prove this $|\zeta| \leq 1$ by pulling out $\log |\zeta|$ otherwise. By translation in $x$ we may further assume that $0 \leq \zeta \leq 1$. One checks that
\[
\partial_x \log |e(x) - \zeta| = \pi \cot(\pi x), \quad \partial_x \log |e(x) - r| = \pi \frac{2r \sin(2\pi x)}{1 + r^2 - 2r \cos(2\pi x)}
\]
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).

We conclude that $|\hat{u}(k)| \leq C n|k|^{-1}$ for all $k \neq 0$ by integrating over the Riesz mass. For the harmonic function we simply use that $|\partial_h h(e(x))| \leq n$ and the decay of the Fourier coefficients follows. By the almost invariance property (6.9),
\[
u x \in \mathbb{T}
\]
Then one has that
\[
\left| \frac{1}{k} \sum_{j=1}^{k} e(j \nu \omega) \right| \leq \min(1, k^{-1} \| \nu \omega \|^{-1})
\]
for all \( \nu \geq 1 \). Also, it follows from (6.10) that \( |\hat{u}(\nu)| \leq n |\nu|^{-1} \) which in turn implies that
\[
|u(x) - \langle u \rangle| \leq \frac{1}{k} \sum_{j=1}^{k} \sum_{|\nu| > K} |\hat{u}(\nu)e(\nu(x + j \omega))| + \sum_{0 < |\nu| \leq K} n |\nu|^{-1} \min(1, k^{-1} \| \nu \omega \|^{-1})
\]
On the one hand, by Plancherel and the decay of the Fourier coefficients,
\[
\left\| \frac{1}{k} \sum_{j=1}^{k} \sum_{|\nu| > K} \hat{u}(\nu)e(\nu(x + j \omega)) \right\|_{L^2} \leq \left\| \sum_{|\nu| > K} \hat{u}(\nu)e(\nu x) \right\|_{L^2} \leq n K^{-1/2}
\]
On the other hand, setting \( K = e^{n^\tau} \) it follows from the Diophantine condition (with \( a = 2 \) for simplicity) that
\[
\sum_{0 < |\nu| \leq K} n |\nu|^{-1} \min(1, k^{-1} \| \nu \omega \|^{-1}) \leq n k^{-1} \log K \leq n^{1+\tau} k^{-1/2}
\]
Choosing \( \tau > 0 \) small and \( k = n^{1/2} \), 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 \( k^{-1/2} < 1 \). 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 [GolSch1] by a more involved argument. The book [Bou1] 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 [GolSch3, 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 Section 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 \( a \leq j \leq k \leq b \),

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

(6.17)

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

\[
H_{[1,n]}(x) - E = \begin{bmatrix}
v_1(x) - E & 1 & 0 & 0 & \cdots & 0 \\
1 & v_2(x) - E & 1 & 0 & \cdots & 0 \\
0 & 1 & v_3(x) - E & 1 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 & v_n(x) - E
\end{bmatrix}
\]

(6.18)

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

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

(6.19)

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

\[
M_n(x, E) = \begin{bmatrix}
(-1)^n f_n(x, E) & (-1)^n f_{n-1}(T x, E) \\
(-1)^n f_{n-1}(x, E) & (-1)^n f_{n-2}(T x, E)
\end{bmatrix}
\]

(6.20)

where we set \( f_{-1} = 0 \). The following uniform upper bound from [GolSch2, 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 [GolSch1]. Moreover, [GolSch2] 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. \( \omega \). 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

\[
\sup_{x \in \mathbb{T}} \log ||M_n(x, E)|| \leq nL_\gamma(E) + C(\log n)^B,
\]

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

In view of (6.20) and the Thouless formula (5.5) 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 [GolSch2, Section2].

**Proposition 6.8.** There exists \( \sigma > 0 \) so that for large \( n \)

\[
||x \in \mathbb{T} | \log |f_n(x, E)| < nL_\gamma(E) - n^\sigma|| \leq e^{-n^\sigma}
\]
A stronger statement is possible if we assume positive Lyapunov exponents. See [GolSch3, Lemma 5.1].

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

\[
\left| \int_0^1 \log |\det(H_{1,n}(x) - E)| \, dx - n L_n(E) \right| \leq C \tag{6.21}
\]

Thus,

\[
|| \log |\det(H_{1,n}(x) - E)| - n L_n(E) ||_{BMO} \leq C(\log n)^A.
\]

Thus,

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

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 [GolSch2]. However, they help to streamline some of the technical aspects of [BouGol]. 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)

\[
|G_{1,n}(x, E)(j, k)| \leq \exp\left( (j - 1) L_{j-1}(E) + (n - k) L_{n-k}(E) - n L_n(E) + (\log n)^{2A} \right) \tag{6.23}
\]

provided

\[
|f_n(x, E)| \geq n L_n(E) - (\log n)^{2A}
\]

which therefore holds up to a set of measure $\leq \exp\left( -(\log n)^A \right)$ (assuming $A \geq B$). It was proved in [GolSch1] that $L_n(E) - L(E) \leq Cn^{-1}$ whence it follows from (6.23) that

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

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

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

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

\[
|E| \leq 2 + ||V||_\infty \tag{6.26}
\]

since this range contains $\text{spec}(H_{1,n})$. 

Corollary 6.10. Under the same assumptions as Proposition 6.9 one has
\[ |\{ x \in \mathbb{T} | \text{dist}(\text{spec}(H_{[1,n]}(x)), E) < \exp(-n^{1/3}) \}| \leq e^{-n^{1/2}} \]  
(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
\[ \text{dist}(\text{spec}(H_{[-m,m]}(0)), E) \leq e^{-m^{1/4}} \]  
(6.28)
Indeed, denote the set in Proposition 6.8 by \( B_n \). It consists of \( O(n) \) intervals of length \( e^{-n^r} \). Therefore, by the Diophantine condition the set
\[ \{ n \leq m \leq n^2 | m\omega \in B_n(E) \cup (-B_n(E)) \} \]
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 (3.27) 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
\[ |k\omega| \geq \frac{b}{k(1 + \log k)^2} \quad \forall 1 \leq k \leq n \]  
(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 [GolSchVod, 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),
\[ |\{ x \in \mathbb{T} | \log |f_n(x, \omega, E)| < nL_n(E, \omega) - n^{\frac{1}{2}} \text{ for some } |E - E_0| \leq e^{-n} \}| \leq e^{-n^{\frac{1}{2}}} \]
(6.30)
if \( n \geq n_0(V, b, \gamma) \) is large.

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 [GolSchVod, Corollary 2.15] provides. The statement in loc. cit. is slightly weaker, but replacing the upper bound of [GolSchVod, 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 $b > 0$

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

(6.31)

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

$$|S_n(b)| \leq 3n^5 e^{-n^{1/12}}.$$  

(6.32)

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

$$B_n(b) := \{ \omega \in \mathbb{T} \mid (\omega, \ell \omega) \in S_n(b) \text{ for some } \pm \ell \in [n^{2s}, 2n^{2s}] \}$$

(6.33)

Here $s \geq 2$ is an absolute constant, which we will specify later. The following lemma on steep lines from [BouGol] guarantees that $B_n(b)$ 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 $S_n(b)$. Specifically, the fact that the horizontal slices

$$(S_n(b))_x := \{ \omega \in \mathbb{T} \mid (\omega, x) \in S_n(b) \}$$

(6.34)

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

---

**Figure 15.** The lemma on steep lines
Lemma 6.12. Suppose the Borel set \( S \subset \mathbb{T}^2 \) has the property that for every \( x \in \mathbb{T} \) the horizontal slice \( S_x \), viewed as a subset of \([0, 1]\), consists of no more than \( M \) intervals. Then
\[
\| \omega \in \mathbb{T} \mid (\omega, \ell \omega) \in S \mod \mathbb{Z}^2 \text{ for some } \ell \in [N, 2N] \| \leq \frac{M}{N} + 8N^2 \| S \|^{\frac{1}{2}} \tag{6.35}
\]

Proof. By Fubini, for each \( \gamma > 0 \),
\[
|\{ x \in \mathbb{T} \mid |S_x| > \gamma \}| \leq |S| \gamma^{-1}
\]
We define the set of good \( x \in \mathbb{T} \) as
\[
G := \{ x \in \mathbb{T} \mid |S_x| \leq \gamma \text{ and for all } j \in [1, N] \text{ one has } \|x_j\| > 4N^2 \gamma \} \tag{6.36}
\]
with \( \| \cdot \| \) the norm of \( \mathbb{T} \). Then
\[
|\mathbb{T} \setminus G| \leq |S| \gamma^{-1} + 4N^3 \gamma
\]
We optimize here by setting \( \gamma = (4N^3)^{\frac{1}{2}} |S|^{\frac{1}{2}} \) whence
\[
|\mathbb{T} \setminus G| \leq 4N^2 |S|^{\frac{1}{2}} \tag{6.37}
\]
Correspondingly,
\[
S = S_x \cup S_{xx} := (S \cap G) \cup (S \cap (\mathbb{T} \setminus G))
\]
We eliminate \( S_{xx} \) as follows:
\[
\| \{ \omega \in \mathbb{T} \mid (\omega, \ell \omega) \in S_{xx} \mod \mathbb{Z}^2 \text{ for some } \ell \in [N, 2N] \| \leq \sum_{\ell = N}^{2N} |\{ \omega \in \mathbb{T} \mid \ell \omega \in \mathbb{T} \setminus G \}| \leq (N + 1)|\mathbb{T} \setminus G| \leq 8N^2 |S|^{\frac{1}{2}} \tag{6.38}
\]
On the other hand, where \( \{x\} = x - \lfloor x \rfloor \) for \( x > 0 \) denotes the fractional part,
\[
\| \{ \omega \in \mathbb{T} \mid (\omega, \ell \omega) \in S_x \mod \mathbb{Z}^2 \text{ for some } \ell \in [N, 2N] \| \]
\[
\leq \sum_{\ell = N}^{2N} \int_{0}^{1} 1_{S_x}(x, \{\ell \omega\}) \, d\omega = \sum_{\ell = N}^{2N} \frac{1}{\ell} \sum_{k = 0}^{\ell - 1} \int_{0}^{1} 1_{S_x}(x + k/\ell, x) \, dx \tag{6.39}
\]
\[
\leq \sum_{a = 1}^{M} \frac{1}{N} \int_{G} \sum_{\ell = N}^{2N} \sum_{k = 0}^{\ell - 1} 1_{I_a}(x + k/\ell) \, dx
\]
Here, for \( x \in G \), \( S_x = \bigcup_{a = 1}^{M} I_a(x) \) with \( I_a(x) \) intervals of length \( |I_a(x)| \leq \gamma \), possibly empty.
We claim that for all \( x \in G \) one has
\[
\sum_{\ell = N}^{2N} \sum_{k = 0}^{\ell - 1} 1_{I_a}(x + k/\ell) \leq 1 \tag{6.40}
\]
Indeed, suppose \( \ell \neq \ell' \) both in \([N, 2N]\) and \( \frac{x + k}{\ell}, \frac{x + k'}{\ell'} \in I_a(x) \). Then
\[
\left| \frac{x + k}{\ell} - \frac{x + k'}{\ell'} \right| \leq |I_a(x)|
\]
whence \(|x(\ell \ell') + k\ell' - k\ell| \leq \ell\ell'|I_n(x)|\) and thus \(||x|| \leq 4N^2\gamma\) 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 \(k \neq k'\)
\[
\gamma \geq |I_n(x)| \geq \left|\frac{x + k}{\ell} - \frac{x + \ell'}{\ell'}\right| \geq \frac{1}{\ell} \geq \frac{1}{N}
\]
which contradicts that \(N^2\gamma^2 = (4N)^{-1}|S| < 1\). So Claim \((6.40)\) is correct, and the entire contribution to \((6.39)\) is at most \(M/N\). 

\[\Lambda_0 \quad \text{all intervals of size } \geq m^{1/4} \text{ are regular for } E \]

\[0 \quad m \quad \approx m^{1/4} \quad \approx m^{1/4} \]

**Figure 16.** Absence of double resonances

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

\[S = \bigcup_k \bigcap_{j \in F_k} \{P_j \sigma_{kj} 0\}\]

with \(\sigma_{kj} \in \{\leq, \geq, 0\}\) and \(F_k \subset \{1, 2, \ldots, s\}\). The degree of \(S\) 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 \(S_n(b)\) 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 \(S\) 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 \(\text{[BasPolRoy1]}\) and \(\text{[BasPolRoy2]}\).

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 \(\text{[BouGo]}\). 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 \(S_n(b)\) is contained in

\[\overline{S}_n(b) := \Pi_{\mathbb{R}^2}: (\omega, x, E) \in \text{Dioph}_n(b) \times \mathbb{T} \times \mathbb{R} \mid \log |f_{2n+1}(n\omega, \omega, E)| \leq (2n + 1)L_{2n+1}(E, \omega) - n^{1/4}/2, \]

and

\[\log |f_m(x, \omega, E)| \leq mL_m(E, \omega) - m^{1/2} \text{ for some } m \in \{n^{1/4}/2, n^{1/4}\}\]

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

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


These relations follow from noting that
\[ \text{dist}(\text{spec}(H_{[-n,n]}(0, \omega)), E) = \|(H_{[-n,n]}(0, \omega)) - E\|^{-1} \] for any $d \times d$ matrix $A$, and using the relation \eqref{eq:dist-spec}. In particular, we obtain essentially the same estimates on their two-dimensional measure. The sets $\mathcal{S}_n(b) := \prod_{n \geq 2} \mathcal{D}_n(b)$ with
\[ \mathcal{D}_n(b) = \{(\omega, x, E) \in \text{Dioph}_n(b) \times \mathbb{T} \times \mathbb{R} \mid f_{m+1}(n \omega, \omega, E) \leq \exp((2n + 1) L_{2n+1}(E, \omega) - n^{1/4}/2) \}
\] 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 \cite{GolSch1} rate of convergence estimate $L_m(E, \omega) - L(E, \omega) \leq C m^{-1}$, it would be counter productive to do so at this point. Rather, we will use that uniformly in $x$,
\[ nL_m(E, \omega) = \frac{1}{n^2} \sum_{j=1}^{n^2} \log \|M_n(x + j \omega, \omega, E)\| + O((\log n)^4) \]
This follows from $\|M_n\| \geq 1$, \eqref{eq:Mn-bound}, and the same arguments which we used in the proof of \eqref{eq:Lyap-exponent}. Therefore, we can replace
\[ |f_m(x, \omega, E)| \leq \exp(mL_m(E, \omega) - m^{1/2}) \]
with
\[ |f_m(x, \omega, E)|^{2n^2} \leq e^{-m^{3/2}} \prod_{j=1}^{n^2} \|M_n(j \omega, \omega, E)\|_{HS}^2 \]
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 $\leq n^4$ polynomials involved in the description of the semi-algebraic set $\mathcal{D}_n(b)$ above, it is of degree $\leq n^4$. Projecting out $E$, we conclude that $\mathcal{S}_n(b)$ has degree $O(n^4)$ for some finite $s$ as claimed. Finally, each horizontal slice consists of at most $O(n^4)$ many connected components, i.e., intervals.

**Proof of Theorem 6.1** The set of admissible $\omega$ for the theorem is
\[ \Omega := \text{Dioph} \setminus \bigcup_{j=1}^{\infty} \limsup_{n \to \infty} \mathcal{B}_n(1/j), \quad \text{Dioph} := \bigcup_{j=1}^{\infty} \text{Dioph}(1/j) \]
where $\mathcal{B}_n(b)$ is defined in \eqref{eq:Dioph-Bn}. By Lemma 6.12
\[ |\mathcal{B}_n(1/j)| \leq C(j)(n^{-s} + n^C e^{-n^{1/2}}), \quad \sum_{n=1}^{\infty} |\mathcal{B}_n(1/j)| < \infty \]
whence by Borel-Cantelli $\limsup_{n \to \infty} \mathcal{B}_n(1/j) = 0$. Since Dioph has full measure in $\mathbb{T}$, so does $\Omega$. 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 \( B_n(b) \), see Figure 16 we conclude that all Green functions \( G_\Lambda(0, \omega, E) \) with \( \Lambda \subset [m^s, 2m^s] \) and \(|\Lambda| \sim m^{\frac{1}{4}}\) satisfy
\[
\|G_\Lambda(0, \omega, E)\| \leq e^{\frac{1}{2}|\Lambda|}, \quad |G_\Lambda(0, \omega, E)(x, y)| \leq e^{-\gamma|x-y|+|\Lambda|^\frac{1}{2}} \quad \forall \ x, y \in \Lambda.
\]
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 \( E \):
\[
\|G_{\pm[ms, 2ms]}(0, \omega, E)\| \leq e^{\frac{1}{8}m^s}, \quad |G_{\pm[ms, 2ms]}(0, \omega, E)(x, y)| \leq e^{-\gamma|x-y|+m^s} \quad \forall \ x, y \in \pm[m^s, 2m^s]
\]
from which the exponential decay of \( \psi \) immediately follows. □

```
\( n = 200; \) \( \epsilon = 0.3; \) \( \omega = \text{Sqrt}[2]; \) \( \theta = -17 \times \omega / 2; \)
\( \text{Laplace} = \epsilon \times \text{Total[]} \)
\( \{\text{DiagonalMatrix}[\text{Array}[-1 &, n-1, 1]], \text{DiagonalMatrix}[\text{Array}[-1 &, n-1, -1]]\} \)
\( f = \text{Cos}[2 \times \pi \times (\theta + \omega \times \omega)]; \)
\( V = \text{DiagonalMatrix}[\text{Array}[f, n]]; \)
\( H = \text{Laplace} \times V; \)
\( \text{With}[\{\text{eigs} = \text{Eigenvectors}[H]\}, \)
\( \text{Manipulate}[, \text{ListLinePlot}[\text{eigs}[k]], \{\text{ImageSize} \to 4 \times 72, \text{PlotRange} \to (-1, 1)], \{\{k, 1, 1, \text{Length}[	ext{eigs}, 1]\}]
```

**Figure 17. Mathematica code for Figures (9), (10)**

**References**

[AizMol] Aizenman, M., Molchanov, S. *Localization at large disorder and at extreme energies: an elementary derivation.* Comm. Math. Phys. 157 (1993), no. 2, 245–278.

[AizWar] Aizenman, M., Warzel, S. *Random operators. Disorder effects on quantum spectra and dynamics.* Graduate Studies in Mathematics, 168. American Mathematical Society, Providence, RI, 2015.
ANDERSON LOCALIZATION, MULTISCALE TECHNIQUES: I

[Avi] Avila, A. Global theory of one-frequency Schrödinger operators. Acta Math. 215 (2015), no. 1, 1–54.

[BasPolRoy1] Basu, S., Pollack, R., Roy, M. On the combinatorial and algebraic complexity of quantifier elimination, Journal of the ACM, 43 (1996), no. 6, 1002–1045.

[BasPolRoy2] Basu, S., Pollack, R., Roy, M. Algorithms in real algebraic geometry, Springer-Verlag, Berlin, 2003.

[Ber] Berezanskii, J. Expansions in Eigenfunctions of Selfadjoint Operators, Translations of Mathematical Monographs, Vol. 17, Amer. Math. Soc., Providence, R.I., 1968.

[BouLac] Bougerol, P., Lacroix, J. Products of random matrices with applications to Schrödinger operators. Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, 1985.

[Bou1] Bourgain, J. Green’s function estimates for lattice Schrödinger operators and applications. Annals of Mathematics Studies, 158, Princeton University Press, Princeton, NJ, 2005.

[Bou2] Bourgain, J. Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, $d$ arbitrary. Geom. Funct. Anal. 17 (2007), no. 3, 682–706.

[BouGol] Bourgain, J.; Goldstein, M. On nonperturbative localization with quasi-periodic potential. Ann. of Math. (2) 152 (2000), no. 3, 835–879.

[BouGolSch1] Bourgain, J., Goldstein, M., Schlag, W. Anderson localization for Schrödinger operators on $\mathbb{Z}$ with potentials given by the skew-shift. Comm. Math. Phys. 220 (2001), no. 3, 583–621.

[BouGolSch2] Bourgain, J., Goldstein, M., Schlag, W. Anderson localization for Schrödinger operators on $\mathbb{Z}^2$ with quasi-periodic potential. Acta Math. 188 (2002), no. 1, 41–86.

[BDFGVWZ] Bucaj, V., Damanik, D., Fillman, J., Gerbuz, V., VandenBoom, T., Wang, F., Zhang, Z. Localization for the one-dimensional Anderson model via positivity and large deviations for the Lyapunov exponent. Trans. Amer. Math. Soc. 372 (2019), no. 5, 3619–3667.

[CarKleMar] Carmona, R., Klein, A., Martinelli, F. Anderson localization for Bernoulli and other singular potentials. Comm. Math. Phys. 108 (1987), no. 1, 41–66.

[CarLac] Carmona, R., Lacroix, J. Spectral theory of random Schrödinger operators. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1990.

[CraSim] Craig, W., Simon, B. Subharmonicity of the Lyaponov index. Duke Math. J. 50 (1983), no. 2, 551–560.

[DamFil1] Damanik, D., Fillman, J. One-Dimensional Ergodic Schrödinger Operators I. General Theory, Graduate Studies in Mathematics, AMS, to appear.

[DamFil2] Damanik, D., Fillman, J. One-Dimensional Ergodic Schrödinger Operators II. Special Cases, Graduate Studies in Mathematics, AMS, to appear.

[DinSma] Ding, J., Smart, C. Localization near the edge for the Anderson Bernoulli model on the two dimensional lattice. Invent. Math. 219 (2020), no. 1, 467–506.

[Far] Faris, W. Perturbations and non-normalizable Eigenvectors, Helvetica Physica Acta, 44 (1971), 930–936.

[FigPas] Figotin, A., Pastur, L. Spectra of Random and Almost-Periodic Operators, Springer, 1992.

[ForVan] Forman, Y., VandenBoom, T. Localization and Cantor spectrum for quasiperiodic discrete Schrödinger operators with asymmetric, smooth, cosine-like sampling functions, preprint 2021.

[FroSpe] Fröhlich, J., Spencer, T. Absence of diffusion in the Anderson tight binding model for large disorder or low energy. Comm. Math. Phys. 88 (1983), no. 2, 151–184.

[FroSpeWit] Fröhlich, J., Spencer, T., Wittwer, P. Localization for a class of one-dimensional quasi-periodic Schrödinger operators. Comm. Math. Phys. 132 (1990), no. 1, 5–25.

[FurKes] Fürstenberg, H., Kesten, H. Products of random matrices. Ann. Math. Statist 31 (1960), 457–469.

[GerKle] Germinet, F., Klein, A. Bootstrap multiscale analysis and localization in random media. Comm. Math. Phys. 222 (2001), no. 2, 415–448.

[GolSch1] Goldstein, M., Schlag, W. Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. Ann. of Math. (2) 154 (2001), no. 1, 155–203.

[GolSch2] Goldstein, M., Schlag, W. Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues, Geom. Funct. Anal. 18 (2008), no. 3, 755–869.

[GolSch3] Goldstein, M., Schlag, W. On Schrödinger operators with dynamically defined potentials. Mosc. Math. J. 5 (2005), no. 3, 577–612.
On localization and the spectrum of multi-frequency quasi-periodic operators, [https://arxiv.org/abs/1610.00380](https://arxiv.org/abs/1610.00380)

Effective multi-scale approach to the Schrödinger cocycle over a skew shift base, to appear in Ergodic Theory and Dyn. Systems, preprint 2018.

Department of Mathematics, Yale University, New Haven, CT 06511, USA
Email address: wilhelm.schlag@yale.edu