LOCALIZATION FOR ONE DIMENSIONAL, CONTINUUM, BERNOULLI-ANDERSON MODELS

DAVID DAMANIK, ROBERT SIMS, AND GÜNTER STOLZ

Abstract. We use scattering theoretic methods to prove strong dynamical and exponential localization for one dimensional, continuum, Anderson-type models with singular distributions; in particular the case of a Bernoulli distribution is covered. The operators we consider model alloys composed of at least two distinct types of randomly dispersed atoms. Our main tools are the reflection and transmission coefficients for compactly supported single site perturbations of a periodic background which we use to verify the necessary hypotheses of multi-scale analysis. We show that non-reflectionless single sites lead to a discrete set of exceptional energies away from which localization occurs.

1. Introduction

Perhaps the most studied and best understood type of random operators which are used to describe spectral and transport properties of disordered media are the Anderson models, whether one considers their original discrete version in $\ell^2(\mathbb{Z}^d)$ or their continuum analogs. They describe materials of alloy type, i.e. a structure in which single site potentials are centered at the points of a regular lattice and then multiplied by random coupling constants, modelling the differing nuclear charges of the alloy's component materials. It is therefore physically most relevant to study the case where the coupling constants take on only finitely many values. The special case of a two-component alloy, i.e. coupling constants given by Bernoulli random variables, has been dubbed the Bernoulli-Anderson model. Unfortunately, most of the rigorous results on Anderson models require more regularity of the probability distributions governing the coupling constants. Results on exponential localization and, more recently, dynamical localization have usually been proven under the assumption that the distribution is absolutely continuous or, at least, has an absolutely continuous component, see [13] and [38] for results from the 80s, and [1, 23, 15, 22, 30, 31, 46] for some of the more recent contributions. In dimension $d > 1$ the best results with respect to weak regularity are due to Carmona, Klein, and Martinelli [12] for the discrete case and to Stollmann [43] and Damanik and Stollmann [16] for continuum models, where localization properties are shown under merely assuming hölder-continuity of the distribution.

Discrete distributions and, in particular, the Bernoulli-Anderson model, are currently only accessible in $d = 1$. For the discrete, one-dimensional Anderson model

$$(1.1) \qquad (h_\omega u)(n) = u(n+1) + u(n-1) + q_n(\omega)u(n), \quad n \in \mathbb{Z},$$

exponential localization has been proven at all energies for arbitrary, non-trivial distribution (i.e. support containing more than one point) of the independent,

Date: October 13, 2000.

identically distributed (i.i.d.) random variables q_n. This was first proven by Carmona, Klein, and Martinelli [12] and later by Shubin, Vakilian, and Wolff [39] with a different approach which, in turn, is based on results from [9]. An interesting recent paper by De Bièvre and Germinet [6] studies the so-called *random dimer model*. This is the case where in (1.1) one chooses $q_{2n+1}(\omega) = q_{2n}(\omega)$ and the q_{2n}, $n \in \mathbb{Z}$ as i.i.d. random variables with only two values $\pm q$ for some $q > 0$. [6] establishes dynamical localization for the random dimer model in compact intervals away from certain critical energies, where the Lyapunov exponent vanishes and delocalized states exist.

In this paper we will prove localization, exponential and dynamical, for *continuous* one-dimensional Anderson models with arbitrary non-trivial distribution. Let

$$H_\omega = -\frac{d^2}{dx^2} + V_{\mathrm{per}} + V_\omega, \quad \omega \in \Omega, \tag{1.2}$$

with

$$V_\omega(x) = \sum_{n\in\mathbb{Z}} q_n(\omega) f(x-n). \tag{1.3}$$

We assume that the background potential V_{per} has period 1, is real-valued and locally in L^1. The single site potential $f \in L^1$ is real-valued, supported in $[-1/2, 1/2]$, and not 0 (in L^1-sense). The coupling constants q_n are i.i.d. random variables on a complete probability space Ω. We assume that the support of their common distribution μ is bounded and non-trivial, i.e. contains at least two points. Note that without restriction we may assume that $\mathrm{supp}(\mu)$ contains 0 and 1 (if $\{a,b\} \subset \mathrm{supp}(\mu)$, then replace V_{per} by $V_{\mathrm{per}} + a\sum_n f(\cdot - n)$ and f by $(b-a)f$). We will do this in all our proofs.

Under these assumptions the operators H_ω can be defined either by form methods or via Sturm-Liouville theory and are self-adjoint for every ω. Representing, as usual, Ω as an infinite product and $q_n(\omega)$ as ω_n, one easily sees that

$$H_{\theta_k\omega} f = \tau_k^* H_\omega \tau_k f \quad \text{for } f \in D(H_{\theta_k\omega}),$$

with $\tau_k f = f(\cdot - k)$ and the ergodic shifts $(\theta_k\omega)_n = \omega_{n+k}$. Thus, e.g. [13], H_ω has non-random almost sure spectrum Σ and spectral types Σ_{ac}, Σ_{sc}, and Σ_{pp}. Our first main result is

Theorem 1.1 (Exponential Localization). *Almost surely, the operator H_ω has pure point spectrum, i.e. $\Sigma_{ac} = \Sigma_{sc} = \emptyset$, and all eigenfunctions decay exponentially at $\pm\infty$.*

Our method to prove Theorem 1.1 is basically to adapt the approach of [12] to continuum models. We will, however, use a variable-energy multiscale analysis, which was introduced in [19] for discrete models and later adapted to the continuum, e.g. [44, 23, 22].

Recently, it has been demonstrated by Damanik and Stollmann [16] that the variable-energy multiscale analysis implies *strong dynamical localization*. Thus our method of proof will also yield

Theorem 1.2 (Strong Dynamical Localization). *Let H_ω be defined as above. There exists a discrete set $M \subset \mathbb{R}$ such that for every compact interval $I \subset \mathbb{R} \setminus M$, and every compact set $K \subset \mathbb{R}$, and every $p > 0$,*

$$\mathbb{E}\left\{\sup_{t>0} \| |X|^p e^{-itH_\omega} P_I(H_\omega)\chi_K \| \right\} < \infty, \tag{1.4}$$

where P_I is the spectral projection onto I.

We note that the "strong" in Theorem 1.2 refers to the ability to show finite expectation in (1.4). This is stronger than showing that the supremum in (1.4) is almost surely finite, which has also been used to describe dynamical localization. For more details on dynamical localization we refer to [18, 23, 2, 4, 16].

A crucial tool in [12] as well as in our work is positivity of the Lyapunov exponent. To define it in the continuous case, let $g_\lambda(n,\omega)$ denote the transfer matrix from $n-1/2$ to $n+1/2$ of

$$-u'' + V_\omega u = \lambda u \tag{1.5}$$

i.e. for any solution u of (1.5)

$$\begin{pmatrix} u(n+1/2) \\ u'(n+1/2) \end{pmatrix} = g_\lambda(n,\omega) \begin{pmatrix} u(n-1/2) \\ u'(n-1/2) \end{pmatrix}. \tag{1.6}$$

For $n \in \mathbb{N}$, let

$$U_\lambda(n,\omega) = g_\lambda(n,\omega) \cdot \ldots \cdot g_\lambda(1,\omega).$$

The subadditive ergodic theorem (e.g. [13]) guarantees the existence of the Lyapunov exponent, i.e. the limit

$$\gamma(\lambda) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left(\|U_\lambda(n,\omega)\|\right).$$

Kotani Theory implies that $\gamma(\lambda)$ is positive for almost every $\lambda \in \mathbb{R}$ since H_ω is non-deterministic in the sense of Kotani, e.g. [34]. But this is insufficient for a proof of localization, due to the fact that the singular distribution of μ prevents spectral averaging techniques from being used to prove absence of singular continuous spectrum. For the discrete model (1.1), one can instead use Furstenberg's Theorem on products of independent random matrices to show that the Lyapunov exponent is positive for all energies, which is the starting point of the approach in [12].

We will use the same idea, but there is an additional, significant difficulty to overcome. For the continuous model (1.2), (1.3) there may exist a set of critical energies at which Furstenberg's Theorem does not apply, and, in most cases, $\gamma(\lambda)$ vanishes. This gives rise to the exceptional set M in Theorem 1.2 and is a continuum analog of the critical energies observed in [6] for the random dimer model.

We manage to prove discreteness of the set of critical energies M for arbitrary single site potentials f with support in $[-1/2, 1/2]$ and arbitrary periodic background V_{per} by using methods from scattering theory. More precisely, we will consider scattering at f with respect to the periodic background V_{per}. For energies λ in the interior of spectral bands of the periodic operator $H_0 = -d^2/dx^2 + V_{\text{per}}$, we will introduce the reflection and transmission coefficients $b(\lambda)$ and $a(\lambda)$ relative to V_{per}. The roots of $b(\lambda)$ give rise to critical energies where the Lyapunov exponent vanishes and extended states exist. Since $f \neq 0$ it can be shown that this set of energies is discrete. Other discrete exceptional sets will be included in M, among them the band edges of H_0 and roots of the analytic continuations of $b(\lambda)$ and

$a(\lambda)$ to the spectral gaps of H_0. Away from M, positivity of γ will follow from Furstenberg's Theorem.

The fact that extended states exist at some critical energies does not affect Theorem 1.1 since the discrete exceptional set can not support continuous spectrum. It does, however, affect dynamical localization in that the interval I in Theorem 1.2 specifically excludes these critical values.

The existence of extended states at critical energies in certain one-dimensional tight binding models or, equivalently, disordered harmonic chains, is well known in the physics literature. Various aspects of this are discussed in Section 10 of [35], which also includes continuum models. Additional references can be found through [17] and [6]. These works also note the vanishing of single site reflection coefficients as the basic mechanism for obtaining critical energies and discuss the anomalous transport behavior near these energies. In several of these models the discrete random potential is of the form

$$V_\omega(n) = \sum_{k\in\mathbb{Z}} q_k(\omega) f(n-k\ell), \tag{1.7}$$

where the single site potential $f:\mathbb{Z}\to\mathbb{R}$ is supported on $\{0,\dots,\ell-1\}$. For example, in the dimer model as studied in [6] one sets $\ell=2$, $f=\chi_{\{0,1\}}$. While in our work we focus on a systematic study of exceptional energies in continuum Anderson models, we expect that our methods can be adapted to yield exponential and dynamical localization for general discrete Anderson models of the form (1.7). While we exclude neighborhoods of critical energies in our discussion of dynamical localization, as does [6], it is an interesting open problem to describe the transport behavior near critical energies rigorously.

A review of results on positivity of the Lyapunov exponent which have been obtained from Furstenberg's Theorem is given in Section 14.A of [38]. Most of these results were obtained for discrete Schrödinger operators and Jacobi matrices, see e.g. [26, 25, 21]. Results for continuum Anderson models (1.2), (1.3) had been restricted to special exactly solvable cases which could be reduced to discrete models. For example, the set where γ vanishes was characterized for the cases where $V_{\mathrm{per}}=0$ and either f is a δ-point-interaction [25] or $f=\chi_{[-1/2,1/2]}$ [5]. For the latter, if it is assumed that $\mathrm{supp}(\mu)=\{0,\xi\}$ for some $\xi>0$, then one gets the critical energies $\{\lambda:\gamma(\lambda)=0\}=\{\xi+\pi^2n^2, n\in\mathbb{Z}\}$.

The first paper where reflection coefficients are used rigorously in combination with Furstenberg's Theorem to characterize the exceptional set for general f is [32]. Many of the ideas used in Section 2 below can be found there. They assume $V_{\mathrm{per}}=0$ and thus can work with the "classical" reflection and transmission coefficients for scattering at f. Also, [32] assumes that the distribution μ of q_n has continuous density, but the methods used can be modified to include some examples with singular distributions. We also mention the recent paper [33] which contains bounds for the Lyapunov exponent, which might be useful to study transport properties near critical energies.

That energies with vanishing reflection coefficient at a single site have zero Lyapunov exponent was also observed for random displacement models in [42]. That the latter model shows similar properties as the Bernoulli-Anderson model is not surprising: Both models display identical single site potentials at random distances.

To our knowledge, the only previous work on localization for *continuum* Bernoulli-Anderson models is [10]. This work, in combination with remarks in [11], looks at the example $V_{\text{per}} = 0$, $f = \chi_{[-1/2,1/2]}$, $\text{supp}(\mu) = \{0,1\}$, and states exponential localization for all energies. Unfortunately, it seems that some details of the proof, which is sketched in [10] and quite different from ours, have never appeared in print.

We now outline the contents of the remaining sections. In Section 2 we prove positivity of the Lyapunov exponent away from a discrete set using Furstenberg's Theorem. A key tool is to show that the identical vanishing of $b(\lambda)$ on a band of H_0 implies that $f = 0$. This generalizes the well known fact that there are no compactly supported solitons to scattering at periodic background. Sections 3 and 4 establish that the Lyapunov exponent and the integrated density of states are Hölder continuous away from the exceptional set. Both sections rely heavily on known results for the discrete case, which can all be found in [13]. We outline these results and discuss the necessary changes for continuum models. In particular, we make use of a version of the Thouless formula by Kotani [34].

The next two sections provide the main ingredients for the start of a multiscale analysis. In Section 5 a Wegner estimate is proven, again away from the exceptional set. Here we can closely follow the argument of [12]. Section 6 provides an initial length scale estimate (ILSE) for the same energies. Here our argument differs somewhat from the approach in [12]. While the latter uses both, positivity of the Lyapunov exponent and the Wegner estimate, to get an ILSE, we prefer to use an approach which shows that ILSE follows directly from positivity of γ. This uses some results on large deviations for γ, which can all be found in [8]. We note that this also gives a new, more natural proof of an ILSE for discrete models. Having established all what is necessary for a multiscale analysis, we can then prove Theorems 1.1 and 1.2 in Section 7 by merely referring to well known results, e.g. [44] and [16].

Appendix A contains some a priori estimates for solutions of the Schrödinger equation which we use frequently. In Appendix B some basic facts about cocycles and the existence and uniqueness of invariant measures for group actions are listed. They are used in Section 3.

Acknowledgement: We are grateful to B. Simon for pointing us to the simple proof of Lemma 2.2 below, i.e. that there are no compactly supported “solitons” for scattering at periodic background. After this we learned a different proof of this fact from E. Korotyaev. We also acknowledge useful discussions with L. Pastur and R. Schrader.

2. Positivity of the Lyapunov exponent

2.1. **The Floquet Solutions.** We start by collecting some facts from Floquet theory for the periodic operator $H_0 := -\frac{d^2}{dx^2} + V_{\text{per}}$, see e.g. [20]. For any $z \in \mathbb{C}$, let $u_N(\cdot, z)$ and $u_D(\cdot, z)$ denote the solutions of

$$-u'' + V_{\text{per}}u = zu \tag{2.1}$$

with $u_N(-1/2) = u_D'(-1/2) = 1$ and $u_N'(-1/2) = u_D(-1/2) = 0$. The transfer matrix of (2.1) from $-1/2$ to $1/2$ is the matrix

$$g_0(z) = \begin{pmatrix} u_N(1/2, z) & u_D(1/2, z) \\ u_N'(1/2, z) & u_D'(1/2, z) \end{pmatrix},$$

which is entire in z as solutions of (2.1) (resp. their derivatives) are, for each fixed x, entire in z. The eigenvalues of $g_0(z)$ are the roots of

$$\rho^2 - D(z)\rho + 1 = 0, \tag{2.2}$$

i.e.,

$$\rho_\pm(z) = \frac{D(z) \pm \sqrt{D(z)^2 - 4}}{2}, \tag{2.3}$$

where $D(z) = \mathrm{Tr}[g_0(z)]$. As roots of (2.3), the functions $\rho_\pm$ are algebraic with singularities at points with $D(z) = \pm 2$.

The spectrum of H_0, $\sigma(H_0)$, consists of bands which are given by the sets of real energies λ for which $D(\lambda) \leq 2$. Let (a, b) be a stability interval of H_0; i.e. a maximal interval such that $D(\lambda) < 2$ for every $\lambda \in (a, b)$. As both $u_N(\cdot, \lambda)$ and $u_D(\cdot, \lambda)$ are real for $\lambda \in (a, b)$, one has that

$$\rho_\pm(\lambda) = \frac{1}{2}\left(D(\lambda) \pm i\sqrt{4 - D(\lambda)^2}\right),$$

$|\rho_\pm(\lambda)| = 1$, and $\rho_-(\lambda) = \overline{\rho_+(\lambda)}$. Let

$$S := \{z \in \mathbb{C} : z = \lambda + i\eta \;\text{ where }\; a < \lambda < b \;\text{ and }\; \eta \in \mathbb{R}\}$$

be the vertical strip in the complex plane containing (a, b). For $z = \lambda + i\eta \in S$ one has that each of the following are equivalent: (i) $|\rho_\pm(z)|$ =1, (ii) $\eta = 0$, and (iii) $D(z) \in (-2, 2)$. That $(ii) \Rightarrow (iii)$ and $(iii) \Rightarrow (i)$ are clear. To see that $(i) \Rightarrow (ii)$ assume that (i) is true for some $z = \lambda + i\eta$ where $\eta \neq 0$. As both $\rho_\pm$ have modulus 1, then all solutions of (2.1) are bounded. By Weyl's alternative, however, if $\eta \neq 0$, then there exists a solution in L^2 near $+\infty$, the Weyl solution, while all other solutions are unbounded. This is a contradiction.

The arguments above imply that $\rho_\pm$ have analytic continuations to all of S, and that the only possible algebraic singularities occur at a and b. In addition, as $\rho_+(z)\rho_-(z) = \det[g_0(z)] = W[u_N, u_D] = 1$, we have by continuity of $|\cdot|$ that exactly one of ρ_+ and ρ_- satisfies $|\rho(z)| < 1$ in the upper half plane. Without loss of generality, let us denote by ρ_+ the eigenvalue for which $|\rho_+(\lambda + i\eta)| < 1$ for all $\eta > 0$ and $\lambda \in (a, b)$. This corresponds to choosing a branch of the square root in (2.3). Then, ρ_- satisfies $|\rho_-(\lambda + i\eta)| > 1$ for all $\eta > 0$ and $\lambda \in (a, b)$. Since we also have $|\rho_\pm(\lambda)| = 1$ for $\lambda \in (a, b)$, it follows from the Schwarz reflection principle (apply a linear fractional transformation) that $|\rho_+(\lambda + i\eta)| > 1$ and $|\rho_-(\lambda + i\eta)| < 1$ for all $\eta < 0$ and $\lambda \in (a, b)$.

For $z \in S$, let $v_\pm(z)$ be the eigenvectors of $g_0(z)$ corresponding to $\rho_\pm(z)$ with first component normalized to be one; i.e.

$$v_\pm(z) = \begin{pmatrix} 1 \\ c_\pm(z) \end{pmatrix}. \tag{2.4}$$

One may easily calculate that

$$c_\pm(z) = \frac{\rho_\pm(z) - u_N(1/2, z)}{u_D(1/2, z)}. \tag{2.5}$$

As $u_D(1/2, z)$ is never zero in S (Dirichlet eigenvalues of H_0 are either in the gaps of $\sigma(H_0)$ or at the band edges), $v_\pm$ are analytic in S. In particular, as u_D^{-1} has at most a pole at a and b, then $c_\pm$, and hence $v_\pm$, have at worst algebraic singularities at a and b. (More precisely, $c_\pm$ are branches of a multi-valued analytic function with algebraic singularities at the boundaries of stability intervals.) Let $\phi_\pm(\cdot, z)$ be the Floquet solutions of (2.1); i.e. the solutions satisfying

$$\begin{pmatrix} \phi_\pm(-1/2, z) \\ \phi'_\pm(-1/2, z) \end{pmatrix} = v_\pm(z). \tag{2.6}$$

We first note that $\phi_\pm(\cdot, \lambda + i\eta) \in L^2$ near $\pm\infty$ if $\eta > 0$, and $\phi_\pm(\cdot, \lambda + i\eta) \in L^2$ near $\mp\infty$ if $\eta < 0$. Thus in this setting the Floquet solutions are the Weyl solutions. Secondly, for fixed x, $\phi_\pm(x, \cdot)$ are analytic in S with at worst algebraic singularities at a and b arising from the singularities in the initial conditions $v_\pm$. Lastly, $\{\phi_+(\cdot, z), \phi_-(\cdot, z)\}$ are a fundamental system of (2.1) for every $z \in S$, as $\rho_+(z) \neq \rho_-(z)$ on S.

2.2. **Scattering with respect to a periodic background.** Let $0 \neq f \in L^1_{loc}(\mathbb{R})$ be real valued with $\mathrm{supp}(f) \subset [-1/2, 1/2]$. Consider the operator

$$H = -\frac{d^2}{dx^2} + V_{\mathrm{per}} + f$$

in $L^2(\mathbb{R})$. Take $z \in S$ and let u_+ be the solution of

$$-u'' + (V_{\mathrm{per}} + f)u = zu \tag{2.7}$$

satisfying

$$u_+(x, z) = \begin{cases} \phi_+(x, z) & \text{for } x \leq -1/2 \\ a(z)\phi_+(x, z) + b(z)\phi_-(x, z) & \text{for } x \geq 1/2. \end{cases} \tag{2.8}$$

Since $\phi_\pm$ are linearly independent for $z \in S$, this defines $a(z)$ and $b(z)$ uniquely. If $V_{\mathrm{per}} = 0$, then a and b are related to the classical transmission and reflection coefficients by $a = t^{-1}$ and $b = rt^{-1}$. In particular, vanishing of b is equivalent to vanishing reflection coefficient. Thus b and u_+ take the role of a (modified) reflection coefficient and Jost solution relative to the periodic background V_{per}.

Since for $\lambda \in (a, b)$ we know $\phi_-(x, \lambda) = \overline{\phi_+(x, \lambda)}$ from (2.4), (2.5), and (2.6), by taking u_- to be the solution of (2.7) with $u_-(x, \lambda) = \phi_-(x, \lambda)$ for $x \leq -1/2$ we get that for $x \geq 1/2$:

$$u_-(x, \lambda) = \overline{a(\lambda)}\phi_-(x, \lambda) + \overline{b(\lambda)}\phi_+(x, \lambda). \tag{2.9}$$

Using constancy of the non-zero Wronskian, we arrive at the familiar relation

$$|a(\lambda)|^2 - |b(\lambda)|^2 = 1, \tag{2.10}$$

for $\lambda \in (a, b)$, corresponding to $|r|^2 + |t|^2 = 1$.

Proposition 2.1. *$a(\cdot)$ and $b(\cdot)$, defined on S as above, are branches of multi-valued analytic functions with at most algebraic singularities at boundaries of stability intervals.*

Proof. Recall that u_+ is the solution of (2.7) with

$$\begin{pmatrix} u_+(-1/2, z) \\ u'_+(-1/2, z) \end{pmatrix} = \begin{pmatrix} \phi_+(-1/2, z) \\ \phi'_+(-1/2, z) \end{pmatrix} = v_+(z).$$

Thus $\left(u_+(1/2, z), u'_+(1/2, z)\right)^t$ is analytic in S with at worst algebraic singularities at a and b, i.e. the boundaries of the stability interval. But, as was determined before, both $\left(\phi_\pm(1/2,z),\phi'_\pm(1/2,z)\right)^t$ have identical analytic properties, and by the definition of $a(z)$ and $b(z)$,

$$\begin{pmatrix} a(z) \\ b(z) \end{pmatrix} = \begin{pmatrix} \phi_+(1/2,z) & \phi_-(1/2,z) \\ \phi'_+(1/2,z) & \phi'_-(1/2,z) \end{pmatrix}^{-1} \begin{pmatrix} u_+(1/2,z) \\ u'_+(1/2,z) \end{pmatrix}, \tag{2.11}$$

so we are done. □

Lemma 2.2. *If $b(\lambda)=0$ for all $\lambda\in(a,b)$, then f is identically zero.*

Proof. Suppose that $b(\lambda)=0$ for all $\lambda\in(a,b)$, and hence all $z\in S$ by analyticity. We know then that for every $\lambda\in(a,b)$ and $\eta>0$, the Jost solution $u_+(x,\lambda+i\eta)$ is $a(\lambda+i\eta)\phi_+(x,\lambda+i\eta)$ (for all $x\geq 1/2$); i.e. u_+ is exponentially decaying in this region of the upper half plane. Thus u_+ is the Weyl solution for the perturbed equation (2.7). We may therefore calculate the Weyl-Titchmarsh M-function (see e.g. [14]), $M_{V_{\rm per}+f}$, for (2.7) on the half line $(-1/2,\infty)$,

$$M_{V_{\rm per}+f}(\lambda+i\eta)=\frac{u'_+(-1/2,\lambda+i\eta)}{u_+(-1/2,\lambda+i\eta)}=\frac{\phi'_+(-1/2,\lambda+i\eta)}{\phi_+(-1/2,\lambda+i\eta)}=M_{V_{\rm per}}(\lambda+i\eta),$$

where the latter is the M-function of (2.1) on $(-1/2,\infty)$. As the M-functions are analytic in the entire upper half plane, we conclude that $M_{V_{\rm per}+f}(z)=M_{V_{\rm per}}(z)$ for all $z\in\mathbb{C}^+$. Thus, by the recent results of Simon [40] and Gesztesy and Simon [24], which provide generalizations (applicable in our setting) to the classical results of [7, 36, 37], this implies f is identically zero. □

As we assume throughout that f is not identically zero, then we know that $\{\lambda\in(a,b):b(\lambda)=0\}$ is in fact finite as accumulations at algebraic singularities are not possible.

Now we consider a gap. Take α such that $-\infty\leq\alpha<a<b$ and (α,a) is a maximal, non-trivial gap in the spectrum of H_0 (if $a=\inf\sigma(H_0)$, then $\alpha=-\infty$). Consider the following split strip:

$$S':=\{z=\lambda+i\eta:\alpha<\lambda<b,\eta\in\mathbb{R}\}\setminus[\mathrm{a},\mathrm{b}).$$

For $i=1,2$, let $\rho_i(z)$ be the branches of (2.2) with $|\rho_1(z)|<1$ and $|\rho_2(z)|>1$ for all $z\in S'$, which are well defined since $[a,b)$ is excluded. We first note that $\rho_1=\rho_+$ on the upper half of S, but $\rho_1=\rho_-$ on the lower half of S. Secondly, as before, it can be seen that ρ_i are analytic in S' and have at worst algebraic singularities at α, a, and b, and therefore they may be continued analytically across (a,b). In particular, ρ_i is the analytic continuation of ρ_j, where $i,j\in\{1,2\}$ with $i\neq j$.

For $z\in S'$, choose analytic eigenvectors $v_i(z)$ of $g_0(z)$ to corresponding $\rho_i(z)$, see Ch.II.4 [27]. Taking ϕ_i to be the solutions of (2.1) with $(\phi_i(-1/2,z),\phi'_i(-1/2,z))^t=v_i(z)$ for $z\in S'$, we see again that $\phi_1(\cdot,z)$ (resp., $\phi_2(\cdot,z)$) is in L^2 near $+\infty$ (resp., $-\infty$); i.e. they are the Weyl solutions. Set u_i to be the Jost solutions of (2.7) satisfying

$$u_i(\cdot,z)=\begin{cases} \phi_i(x,z) & x\leq -1/2 \\ a_i(z)\phi_i(x,z)+b_i(z)\phi_j(x,z) & x\geq 1/2 \end{cases} \tag{2.12}$$

for $z\in S'$ and the same i,j convention used above. As in (2.11) above, one sees that $a_1(z)$ and $b_1(z)$ are analytic in S'. In fact, in the upper half S_+ of S, they coincide with $a(z)$ and $b(z)$, since for $z\in S_+$, we have that $v_+(z)$ and $v_1(z)$ coincide up to a constant multiple, which cancels in (2.11) and the corresponding expression for

$a_1(z)$ and $b_1(z)$. Thus $a_1(z)$ and $b_1(z)$ are analytic continuations of the restrictions of $a(z)$ and $b(z)$ to S_+. Similarly, it is seen that $a_2(z)$ and $b_2(z)$ are analytic continuations of the restrictions of $a(z)$ and $b(z)$ to the lower half S_- of S. Thus, $a_1(z), b_1(z), a_2(z)$, and $b_2(z)$ have at most algebraic singularities at α, a, and b. In particular, since they cannot vanish identically by $f \neq 0$, the set

$$\{\lambda \in (\alpha, b) : a_1(\lambda)b_1(\lambda)a_2(\lambda)b_2(\lambda) = 0\} \tag{2.13}$$

is discrete and, if $\alpha \neq -\infty$, finite.

2.3. **Transfer matrices and positivity of γ.** To prove positivity of $\gamma(\lambda)$, we will investigate properties of the transfer matrices. As our minimal assumption is that the support of the distribution contains the points 0 and 1, we know that there are at least two non-trivial transfer matrices: the free matrix $g_0(\lambda)$ corresponding to (2.1) and the perturbed matrix $g_1(\lambda)$ corresponding to (2.7). Set $G(\lambda)$ to be the closed subgroup of $\mathrm{SL}(2,\mathbb{R})$ generated by $\{g_q(\lambda) : q \in \mathrm{supp}(\mu)\}$. Let $P(\mathbb{R}^2)$ be the projective space, i.e. the set of the directions in $\mathbb{R}^2$ and $\overline{v}$ be the direction of $v \in \mathbb{R}^2 \setminus \{0\}$. Note that $\mathrm{SL}(2,\mathbb{R})$ acts on $P(\mathbb{R}^2)$ by $g\overline{v} = \overline{gv}$. We say that $G \subset SL(2,\mathbb{R})$ is strongly irreducible if and only if there is no finite G-invariant set in $P(\mathbb{R}^2)$.

Theorem 2.3. *Given a family of operators $\{H_\omega\}$, as in* (1.2), (1.3), *with $f \neq 0$ and $0, 1 \in supp(\mu)$, then there exists a discrete set $M \subset \mathbb{R}$ such that $G(\lambda)$ is non-compact and strongly irreducible for all $\lambda \in \mathbb{R} \setminus M$. In particular, $\gamma(\lambda) > 0$ for all $\lambda \in \mathbb{R} \setminus M$.*

We will first prove that $G(\lambda)$ is not compact by showing that a sequence of elements has unbounded norm. This argument will be valid for all but a discrete set of λ's. Once non-compact, then the group G is known to be strongly irreducible if and only if for each $\overline{v} \in P(\mathbb{R}^2)$,

$$\#\{g\overline{v} : g \in G\} \geq 3, \tag{2.14}$$

see [8]. We shall prove that this condition is also satisfied for all but a slightly larger, yet still discrete set of λ's. The conclusion concerning positivity of γ follows from a general Theorem of Furstenberg which, in this context, states that if the group $G(\lambda)$ is non-compact and strongly irreducible, then $\gamma(\lambda) > 0$; see also [8]. Note that non-compactness of $G(\lambda)$ and (2.14) are both properties which are preserved if $\mathrm{supp}(\mu)$ is increased. Thus we will assume w.l.o.g. that $\mathrm{supp}(\mu) = \{0, 1\}$, i.e. $G(\lambda)$ is generated by $g_0(\lambda)$ and $g_1(\lambda)$.

In the construction of M we will treat stability intervals and spectral gaps of H_0 separately. We will first show that in any given stability interval (a, b), there is at most a finite number of λ's such that Furstenberg's Theorem does not apply. We then show the same for finite spectral gaps (α, a) of H_0. For the spectral gap $(-\infty, \inf \sigma(H_0))$, we allow for accumulation of exceptional energies at $-\infty$. Finally, we join all these exceptional sets and the endpoints of stability intervals to get the discrete set M.

Lemma 2.4. *Let λ be in a stability interval (a, b) of H_0. Then,*

$$g_1(\lambda) = C(\lambda)\tilde{g}_0(\lambda)s(\lambda)C(\lambda)^{-1}, \tag{2.15}$$

where

$$C(\lambda) = \begin{pmatrix} 1 & 0 \\ \mathrm{Re}[c(\lambda)] & \mathrm{Im}[c(\lambda)] \end{pmatrix}, \quad \tilde{g}_0(\lambda) = \begin{pmatrix} \mathrm{Re}[\rho(\lambda)] & \mathrm{Im}[\rho(\lambda)] \\ -\mathrm{Im}[\rho(\lambda)] & \mathrm{Re}[\rho(\lambda)] \end{pmatrix}, \tag{2.16}$$

$$s(\lambda) = \begin{pmatrix} \mathrm{Re}[a(\lambda)+b(\lambda)] & \mathrm{Im}[a(\lambda)+b(\lambda)] \\ -\mathrm{Im}[a(\lambda)-b(\lambda)] & \mathrm{Re}[a(\lambda)-b(\lambda)] \end{pmatrix}, \tag{2.17}$$

and $c(\lambda) = c_+(\lambda) = \overline{c_-(\lambda)}$, $\rho(\lambda) = \rho_+(\lambda) = \overline{\rho_-(\lambda)}$ *as given in* (2.3) *and* (2.5).

Proof. Since $\lambda \in (a,b)$, we have that $c_+(\lambda) = \overline{c_-(\lambda)}$, $\rho_+(\lambda) = \overline{\rho_-(\lambda)}$, and $\mathrm{Im}[c(\lambda)] \neq 0$ from (2.3) and (2.5). Thus $C(\lambda)$ is invertible. For the rest of the proof we will drop the fixed parameter λ. First note that expressing the solutions u_N and u_D in terms of the Jost solutions $u_\pm$ yields

$$g_1(\lambda) = \begin{pmatrix} u_+(1/2) & u_-(1/2) \\ u_+'(1/2) & u_-'(1/2) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ c & \overline{c} \end{pmatrix}^{-1}.$$

Clearly,

$$\begin{pmatrix} u_+(1/2) & u_-(1/2) \\ u_+'(1/2) & u_-'(1/2) \end{pmatrix} = \begin{pmatrix} \phi_+(1/2) & \phi_-(1/2) \\ \phi_+'(1/2) & \phi_-'(1/2) \end{pmatrix} \begin{pmatrix} a & \overline{b} \\ b & \overline{a} \end{pmatrix},$$

and using that $\phi_\pm$ are the Floquet solutions we have

$$\begin{pmatrix} \phi_+(1/2) & \phi_-(1/2) \\ \phi_+'(1/2) & \phi_-'(1/2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ c & \overline{c} \end{pmatrix} \begin{pmatrix} \rho & 0 \\ 0 & \overline{\rho} \end{pmatrix}.$$

Hence we see

$$g_1(\lambda) = \begin{pmatrix} 1 & 1 \\ c & \overline{c} \end{pmatrix} \begin{pmatrix} \rho & 0 \\ 0 & \overline{\rho} \end{pmatrix} \begin{pmatrix} a & \overline{b} \\ b & \overline{a} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ c & \overline{c} \end{pmatrix}^{-1}. \tag{2.18}$$

Let

$$Q := \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix},$$

then it is easily checked that

$$Q^{-1} \begin{pmatrix} a & \overline{b} \\ b & \overline{a} \end{pmatrix} Q = s, \quad Q^{-1} \begin{pmatrix} \rho & 0 \\ 0 & \overline{\rho} \end{pmatrix} Q = \tilde{g}_0, \text{ and } \begin{pmatrix} 1 & 1 \\ c & \overline{c} \end{pmatrix} Q = C,$$

with C, $\tilde{g}_0$ and s as defined in (2.16), (2.17). Thus (2.15) follows from (2.18). ☐

Since $g_0(\lambda)$ corresponds to trivial scattering, it follows from (2.15) that

$$g_0(\lambda) = C(\lambda)\tilde{g}_0(\lambda)C(\lambda)^{-1}. \tag{2.19}$$

Let $\tilde{G}(\lambda)$ be the subgroup of $\mathrm{SL}(2,\mathbb{R})$ generated by $\tilde{g}_0(\lambda)$ and $\tilde{g}_0(\lambda)s(\lambda)$ or equivalently, by $\tilde{g}_0(\lambda)$ and $s(\lambda)$. $\tilde{G}(\lambda)$ is conjugate to $G(\lambda)$, and thus $G(\lambda)$ is non-compact and strongly irreducible if and only if $\tilde{G}(\lambda)$ is non-compact and strongly irreducible.

Before proving Theorem 2.3, we note that for every λ in a stability interval (a,b), $\rho(\lambda) = e^{i\omega}$, where $\omega = \omega(\lambda) \in (0,\pi)$. Thus,

$$\tilde{g}_0(\lambda) = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix} \tag{2.20}$$

is merely a rotation. In particular, for any $\lambda \in (a,b)$ satisfying $b(\lambda) = 0$, one has that $a(\lambda) = e^{i\alpha}$, by (2.10), and therefore $s(\lambda)$ is reduced to a rotation matrix as well. For these finitely many λ's (see the paragraph immediately following the proof of Lemma 2.2), $\tilde{G}(\lambda)$ is a group of rotations and thereby compact. Moreover, by (2.15) and (2.19) we know that at such energies, the norm of products of the transfer matrices does not grow: hence $\gamma(\lambda) = 0$. Let $\tilde{M}^{(a,b)} := \{\lambda \in (a,b) : b(\lambda) = 0\}$. It is clear then that only away from this set, $\tilde{M}^{(a,b)}$, can we hope to apply Furstenberg's Theorem and prove positivity of γ.

Proof of Theorem 2.3. *The stability intervals:* We first consider energies in a stability interval (a,b). Let $\lambda \in (a,b) \setminus \tilde{M}^{(a,b)}$, i.e. such that $b(\lambda) \neq 0$. Set $a(\lambda) = Ae^{i\alpha}$, $b(\lambda) = Be^{i\beta}$ and see that (2.10) implies $A^2 - B^2 = 1$ with $A > B > 0$, $A + B > 1$, and $A - B < 1$. In this notation one may recalculate

$$s(\lambda) = A \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} + B \begin{pmatrix} \cos(\beta) & \sin(\beta) \\ \sin(\beta) & -\cos(\beta) \end{pmatrix}.$$

We wish to show that a sequence of elements in $\tilde{G}(\lambda)$ has unbounded norm. To do so, consider an arbitrary element of $P(\mathbb{R}^2)$, represented by

$$v(\theta) := \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

One sees that the relation

$$s(\lambda)v(\theta) = Av(\theta - \alpha) + Bv(\beta - \theta)$$

suggests the choice $\theta' := \frac{1}{2}(\alpha + \beta)$ yielding

$$s(\lambda)v(\theta') = (A + B)v([\beta - \alpha]/2).$$

Defining $R(\theta) := \|s(\lambda)v(\theta)\|^2$, a short calculation, using $A^2 - B^2 = 1$, shows that

$$R(\theta) - 1 = 2B\left[B + A\cos(\alpha + \beta - 2\theta)\right].$$

As a consequence, we see that $R(\theta) - 1$ is π-periodic and has exactly two roots in $[0, \pi)$. In particular, $R(\theta) = 1$ if and only if $\cos(\alpha + \beta - 2\theta) = -\frac{B}{A}$, which shows that the distance between the zeros of $R(\theta) - 1$ is not equal to $\frac{\pi}{2}$: recall $B \neq 0$ (and $B \neq A$). Similarly, $R(\theta) > 1$ if and only if $\cos(\alpha + \beta - 2\theta) > -\frac{B}{A}$ and hence $|\{\theta \in [0, \pi) : R(\theta) > 1\}| > \frac{\pi}{2}$. As a result, there exists a compact interval K (not necessarily in $[0, \pi)$) with $|K| > \frac{\pi}{2}$ and $\|s(\lambda)v(\theta)\| > c > 1$ for all $\theta \in K$.

Now, applying $s(\lambda)$ once to $v(\theta')$, θ' as above, produces a new vector with norm greater than one, but the direction, initially θ', is possibly altered. A vector with this new direction may not increase in norm by directly applying $s(\lambda)$ again. We may, however, apply the modified free matrix, (2.20), and rotate this new direction by ω. As $\omega \in (0, \pi)$, then finitely many applications of $\tilde{g}_0(\lambda)$ produces a vector v with direction in K. Once in K, a direct application of $s(\lambda)$ does increase the norm size uniformly by $c > 1$, as indicated above. In this manner, we can produce a sequence of elements in $\tilde{G}(\lambda)$ with unbounded norm.

To finish the proof of Theorem 2.3 in the stability interval (a,b), we choose

$$M^{(a,b)} := \{\lambda \in (a,b) : \lambda \in \tilde{M}^{(a,b)} \text{ or } D(\lambda) = 0\}.$$

Clearly, $M^{(a,b)}$ is finite, and by the above $\tilde{G}(\lambda)$ is not compact for $\lambda \in (a,b) \setminus M^{(a,b)}$. It remains the check (2.14) for these values of λ. But this is trivial as $D(\lambda) \neq 0$ implies $\omega \neq \frac{\pi}{2}$, and therefore the free transfer matrix produces three distinct elements in projective space.

The spectral gaps: For all energies $\lambda \in (\alpha, a)$, a maximal spectral gap of H_0, $G(\lambda)$ is non-compact. This is easily seen as $\lambda \in (\alpha, a)$ implies $|D(\lambda)| > 2$ and hence the free transfer matrix, g_0, has an eigenvalue $|\rho| = |\rho_2| > 1$. Thus repeated iteration of g_0 on $\overline{v}_2$, the direction of v_2, produces an unbounded sequence of elements. It remains to check (2.14), which we will do for $\lambda \in (\alpha, a) \setminus M^{(\alpha,a)}$, where

$$M^{(\alpha,a)} := \{\lambda \in (\alpha, a) : a_1(\lambda)b_1(\lambda)a_2(\lambda)b_2(\lambda) = 0\}.$$

As was clear in (2.13), $M^{(\alpha,a)}$ is discrete, and finite if $\alpha \neq -\infty$.

Let $\lambda \in (\alpha, a) \setminus M^{(\alpha,a)}$, i.e. $a_i(\lambda) \neq 0$ and $b_i(\lambda) \neq 0$ for $i = 1, 2$. By (2.12), this is equivalent to

$$s(\lambda)\overline{v}_i \notin \{\overline{v}_1, \overline{v}_2\}, \tag{2.21}$$

where $\overline{v}_i$ are the directions of the eigenvectors v_i of $g_0(\lambda)$, again for $i = 1, 2$. If $\overline{v} \notin \{\overline{v}_1, \overline{v}_2\}$, then $\#\{g_0(\lambda)^n\overline{v} : n \in \mathbb{Z}\} = \infty$. If, on the other hand, $\overline{v} \in \{\overline{v}_1, \overline{v}_2\}$, then we use (2.21) to conclude that an initial application of $s(\lambda)$ followed by iteration of $g_0(\lambda)$ gives an infinite orbit. This shows (2.14) and completes the proof for energies in a spectral gap.

We can now complete the proof by taking $M := M_1 \cup M_2 \cup M_3$, where

$$M_1 := \bigcup_{(a,b)} M^{(a,b)},$$

the union being taken over all stability intervals in $\sigma(H_0)$,

$$M_2 := \bigcup_{(\alpha,a)} M^{(\alpha,a)},$$

the union being taken over all maximal gaps of H_0, and M_3 defined to be the set containing all the endpoints of the stability intervals. □

3. Hölder Continuity of the Lyapunov Exponent

In this section, we prove that the Lyapunov exponent for the family of random operators defined by (1.2), (1.3) is Hölder continuous on every compact interval of $\mathbb{R}$ which doesn't contain energies in the exceptional set M from Theorem 2.3. A result of this type is proven in [13] for an analogue of (1.1) defined on a discrete strip. We give an outline of the proof indicating what changes are necessary in order to see that the argument from [13] can be adapted to our setting. The main changes are due to less explicit forms of the continuum transfer matrices.

Fix a compact interval $I \subset \mathbb{R} \setminus M$, where M is the discrete set from Theorem 2.3.

Theorem 3.1. *The Lyapunov exponent for* (1.2), (1.3) *is uniformly Hölder continuous on I, i.e. there exists a number $\alpha > 0$ and a constant C for which*

$$|\gamma(\lambda) - \gamma(\lambda')| \leq C|\lambda - \lambda'|^{\alpha},$$

for all $\lambda, \lambda' \in I$.

The rest of this section will be used to prove Theorem 3.1, where we rely on general facts from Appendix B.

As in the introduction let $g_\lambda(n, \omega)$ be the transfer matrix of (1.5) from $n - 1/2$ to $n + 1/2$. If $u_N(\cdot, \lambda, \omega)$ and $u_D(\cdot, \lambda, \omega)$ are the solutions of (1.5) with $u_N(n - 1/2) = u_D'(n - 1/2) = 1$, $u_N'(n - 1/2) = u_D(n - 1/2) = 0$, then

$$g_\lambda(n, \omega) = \begin{pmatrix} u_N(n + 1/2, \lambda, \omega) & u_D(n + 1/2, \lambda, \omega) \\ u_N'(n + 1/2, \lambda, \omega) & u_D'(n + 1/2, \lambda, \omega) \end{pmatrix}. \tag{3.1}$$

This shows that for fixed n and λ, $g_\lambda(n,\cdot) : \Omega \to \mathrm{SL}(2,\mathbb{R})$ is measurable, and for fixed λ, $g_\lambda(n,\omega)$ are real-valued, i.i.d. random matrices. Here we have used $\mathrm{supp}(f) \subset [-1/2, 1/2]$.

Let μ_λ denote the distribution of g_λ in $\mathrm{SL}(2,\mathbb{R})$. Using (3.1) and the boundedness of the distribution μ of q_n, we get from Lemma A.1 that

$$\|g_\lambda(n,\omega)\|^2 \le \exp(C_1 + |\lambda| + C_2|q_n(\omega)|) \le C_3, \tag{3.2}$$

and from Lemma A.2

$$\|g_\lambda(n,\omega) - g_{\lambda'}(n,\omega)\| \le C|\lambda - \lambda'|, \tag{3.3}$$

where all constants are uniform in ω,n, and $\lambda, \lambda' \in I$.

Take $G = \mathrm{SL}(2,\mathbb{R})$, $B = P(\mathbb{R}^2)$, and the action $g\overline{v} = \overline{gv}$ as described in Section 2. The projective distance

$$\delta(\overline{x},\overline{y}) = \frac{|x_1y_2 - x_2y_1|}{\|x\| \cdot \|y\|}, \quad \text{for } x, y \in \mathbb{R}^2$$

defines a metric on $P(\mathbb{R}^2)$. (Note that $\delta(\overline{x},\overline{y}) = |\sin(\psi)|$, where ψ is the angle between x and y.) As in Appendix B, Example B.1, set $\tilde{B} := B \times B \setminus \{(b,b) : b \in B\}$ and define a cocycle $\sigma_1 : G \times \tilde{B} \to (0,\infty)$ by

$$\sigma_1\,(g,(\overline{x},\overline{y})) := \frac{\delta(g\overline{x}, g\overline{y})}{\delta(\overline{x},\overline{y})}. \tag{3.4}$$

A short calculation shows that

$$\sigma_1\,(g,(\overline{x},\overline{y})) = \frac{\|x\| \cdot \|y\|}{\|gx\| \cdot \|gy\|},$$

and hence

$$\overline{\sigma}_1(g) := \sup_{b \in \tilde{B}} \sigma_1(g,b) \le \|g\|^2,$$

thus (3.2) implies that $\overline{\sigma}_1(g) \le e^{C+|\lambda|}$ for μ_λ-a.e. g. In particular, this shows that for every $t \in \mathbb{R}$,

$$\sup_{\lambda \in I} \int \overline{\sigma}_1(g)^t d\mu_\lambda(g) < \infty.$$

This proves that for (3.4), assumption (i) of Lemma B.2 is satisfied uniformly with respect to $\lambda \in I$. Proving that assumption (ii) of Lemma B.2 also holds for (3.4), again uniformly with respect to $\lambda \in I$, requires more work.

To this end, for any $\overline{v} \in P(\mathbb{R}^2)$, define

$$\Phi_\lambda(\overline{v}) := \mathbb{E}\left\{\log \frac{\|g_\lambda v\|}{\|v\|}\right\}.$$

The next lemma corresponds to Lemma V.4.7 of [13].

Lemma 3.2. (i) *The mapping $\Phi_\lambda : P(\mathbb{R}^2) \to \mathbb{R}$ is continuous.*
(ii) *The mapping $\Phi : I \times P(\mathbb{R}^2) \to \mathbb{R}$ with $\Phi(\lambda,\overline{v}) = \Phi_\lambda(\overline{v})$ satisfies*

$$\sup_{\overline{v} \in P(\mathbb{R}^2)} |\Phi_\lambda(\overline{v}) - \Phi_{\lambda'}(\overline{v})| \le C|\lambda - \lambda'|, \tag{3.5}$$

for some constant $C = C(I)$. In particular, Φ is continuous.

Proof. (i) By (3.2), $\log \|g_\lambda v\|/\|v\| \le \log \|g_\lambda\| \le C$ uniformly in ω and v. Thus by dominated convergence, $\lim_{n\to\infty} \Phi_\lambda(\overline{v}_n) = \Phi_\lambda(\overline{v})$, if $\overline{v}_n \to \overline{v}$.

(ii) We have that

$$|\Phi_\lambda(\overline{v}) - \Phi_{\lambda'}(\overline{v})| = \left| \mathbb{E}\left\{ \log \frac{\|g_\lambda v\|}{\|g_{\lambda'} v\|} \right\} \right| \le \mathbb{E}\left\{ \log \|g_\lambda g_{\lambda'}^{-1}\| \right\}$$

$$\le \mathbb{E}\left\{ \log\left(\|g_\lambda - g_{\lambda'}\| \cdot \|g_{\lambda'}^{-1}\| + 1 \right) \right\} \le C|\lambda - \lambda'|,$$

where the last inequality follows as $\|g_\lambda - g_{\lambda'}\| \cdot \|g_{\lambda'}^{-1}\| \le C|\lambda - \lambda'|$ for all $\lambda, \lambda' \in I$ by (3.2) and (3.3). □

As we know that for each fixed $\lambda \in I$, $G(\lambda)$ is non-compact and strongly irreducible, then combining Proposition IV.4.11 and Theorem IV.4.14 of [13] we have that there exists a unique μ_λ-invariant probability measure ν_λ on $P(\mathbb{R}^2)$ for which

$$\gamma(\lambda) = \int \log \frac{\|gv\|}{\|v\|} d\mu_\lambda(g) d\nu_\lambda(\overline{v}). \tag{3.6}$$

Corollary 3.3. (i) *γ is continuous on I, and*

(ii)

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left\{ \log \frac{\|U_\lambda(n) v\|}{\|v\|} \right\} = \gamma(\lambda),$$

uniformly with respect to $\lambda \in I$ and $\overline{v} \in P(\mathbb{R}^2)$.

Proof. (i) We first show that $\nu_{\lambda'} \to \nu_\lambda$ weakly as $\lambda' \to \lambda$. Note that in general if $\mu_{\lambda_n} \to \mu_{\lambda_0}$ weakly and $\nu_{\lambda_n} \to \nu_{\lambda_0}$ weakly, then $\mu_{\lambda_n} * \nu_{\lambda_n} \to \mu_{\lambda_0} * \nu_{\lambda_0}$ weakly (for the definition of the convolution, see Appendix B). By (3.3) we have that $\mu_{\lambda'} \to \mu_\lambda$ weakly and an application of Arzela-Ascoli implies that every subsequence of $\{\nu_{\lambda'}\}$ contains a weakly convergent subsequence. Using this, the convergence of the convolutions, and the uniqueness of the invariant measures, we see that $\nu_{\lambda'} \to \nu_\lambda$. Now, continuity follows from (3.6), i.e. noting that $\gamma(\lambda) = \nu_\lambda(\Phi_\lambda)$, and the estimate (3.5), see Corollary V.4.8 of [13].
(ii) This follows as in the proof of (3.5), with (3.2) and (3.3) replaced by corresponding estimates for $U_\lambda(n, 1)$, which again follow from Lemmas A.1 and A.2, see Proposition V.4.9 of [13]. □

Lemma 3.4. *There exists an integer N for which*

$$\sup_{\lambda \in I, b \in \tilde{B}} \int \log\left[\sigma_1(g, b)\right] d\mu_\lambda^N(g) < 0.$$

Proof. One calculates that

$$\frac{1}{n} \int \log \frac{\delta(g\overline{x}, g\overline{y})}{\delta(\overline{x}, \overline{y})} d\mu_\lambda^n(g) = -\frac{1}{n} \int \log \frac{\|gx\|}{\|x\|} d\mu_\lambda^n(g) - \frac{1}{n} \int \log \frac{\|gy\|}{\|y\|} d\mu_\lambda^n(g)$$

$$= -\frac{1}{n} \mathbb{E}\left\{ \log \frac{\|U_\lambda(n) x\|}{\|x\|} \right\} - \frac{1}{n} \mathbb{E}\left\{ \log \frac{\|U_\lambda(n) y\|}{\|y\|} \right\}.$$

The result follows by letting $n \to \infty$ and noting both (i) and (ii) of the above Corollary. □

We have thus shown that Lemma B.2 is applicable to the cocycle σ_1 defined in (3.4). Following the lines of [13], this leads to Hölder continuity properties with respect to λ of the invariant measures ν_λ and the operators T_λ defined by $(T_\lambda f)(\overline{x}) = \int f(g\overline{x})\,d\mu_\lambda(g)$ for f in the Hölder spaces $\mathcal{L}_\alpha(P(\mathbb{R}^2))$, see Appendix B.

Lemma 3.5. *There exists an $\alpha_0 > 0$ such that for $0 < \alpha \le \alpha_0$, there exists $\rho_\alpha < 1$ and $C_\alpha < \infty$ such that:*

$$\sup_{\lambda\in I, \overline{x}\neq\overline{y}\in P(\mathbb{R}^2)} \mathbb{E}\left\{ \frac{\delta(U_\lambda(n)\overline{x}, U_\lambda(n)\overline{y})^\alpha}{\delta(\overline{x},\overline{y})^\alpha} \right\} \le C_\alpha \rho_\alpha^n,$$

and

$$\sup_{\lambda\in I} \|T_\lambda^n f - \nu_\lambda(f)\| \le C_\alpha \|f\|_\alpha \rho_\alpha^n,$$

for $n = 1, 2, \ldots$

Proof. Lemma 3.4 confirms that both assumptions of Lemma B.2 hold uniformly with respect to $\lambda \in I$. That this implies the above estimates, again uniformly with respect to $\lambda \in I$, follows from observing that the proofs of Propositions IV.3.5 and IV.3.15 yield uniform results under uniform assumptions. □

Lemma 3.6. *There exists an $\alpha_0 > 0$ such that for $0 < \alpha \le \alpha_0$, there exists a $C_\alpha < \infty$ such that:*
(i)

$$\|T_\lambda f - T_{\lambda'} f\|_{\alpha/2} \le C_\alpha \|f\|_\alpha |\lambda - \lambda'|^{\alpha/2},$$

and
(ii)

$$|\nu_\lambda(f) - \nu_{\lambda'}(f)| \le C_\alpha \|f\|_\alpha |\lambda - \lambda'|^{\alpha/2},$$

for all $\lambda, \lambda' \in I$ and $f \in \mathcal{L}_\alpha$.

Proof. (i) follows as in V.4.13 of [13] using that

$$\delta(g_\lambda\overline{x}, g_{\lambda'}\overline{x}) = \frac{|((g_\lambda - g_{\lambda'})x)_1\,(g_{\lambda'}x)_2 + (g_{\lambda'}x)_1\,((g_{\lambda'} - g_\lambda)x)_2|}{\|g_\lambda x\|\|g_{\lambda'}x\|}$$

$$\le 2\cdot\frac{\|(g_\lambda - g_{\lambda'})x\|}{\|g_\lambda x\|} \le 2\cdot\|g_\lambda - g_{\lambda'}\|\cdot\|g_\lambda^{-1}\| \le C|\lambda - \lambda'|.$$

(ii) follows as in Proposition V.4.14 of [13] using Lemma B.2 (c) as done therein. □

Compiling these results, Theorem 3.1 follows as in Theorem V.4.15 of [13].

4. Hölder Continuity of the Integrated Density of States

In this section we prove an analogue of Theorem 3.1 for the integrated density of states. We begin with its definition. Let, for $L \in \mathbb{N}$ and $\Lambda_L(0) = [-L/2, L/2]$,

$$H_{\Lambda_L(0)}(\omega) := H_\omega \quad \text{on} \quad L^2(\Lambda_L(0)) \tag{4.1}$$

with H_ω as in (1.2), (1.3) and Dirichlet boundary conditions set at $\pm L/2$. Fix $\lambda \in \mathbb{R}$ and let $N_{L,\omega}(\lambda)$ denote the number of eigenvalues of $H_{\Lambda_L(0)}(\omega)$ less than or equal to λ. By Kingman's super-additive ergodic theorem, we have that the following limit

$$N(\lambda) := \lim_{L\to\infty} \frac{N_{L,\omega}(\lambda)}{L} \tag{4.2}$$

exists for almost every ω and in expectation. In particular,

$$N(\lambda) = \sup_{L\geq 1} \frac{1}{L}\mathbb{E}\left(N_{L,\omega}(\lambda)\right). \tag{4.3}$$

The limit $N(\lambda)$ in (4.2) does not depend on the particular choice of boundary conditions one sets on $\Lambda_L(0)$, however, the monotonicity of the limit in expectation (4.3) may be changed, e.g. with Neumann boundary conditions one gets an "inf" rather than a "sup." $N(\lambda)$ is called the integrated density of states.

As in Section 3, fix a compact interval $I \subset \mathbb{R}\setminus M$, M as in Theorem 2.3.

Theorem 4.1. *The integrated density of states for* (1.2), (1.3) *is uniformly Hölder continuous on I.*

The proof of this theorem follows easily from Theorem 3.1 once we have a means of relating N and γ. For our model, one has

Proposition 4.2. (The Thouless formula) *Let N and γ respectively be the integrated density of states and Lyapunov exponent for* (1.2), (1.3). *Then there exists $\alpha \in \mathbb{R}$ such that, for every $\lambda \in \mathbb{R}$,*

$$\gamma(\lambda) = -\alpha + \int_{\mathbb{R}} \log\left|\frac{\lambda - t}{t-i}\right| dN(t). \tag{4.4}$$

Remarks: (i) The above version of the Thouless formula arises in Kotani's work, e.g. [34]. It differs from the discrete version, see Proposition VI.4.3 of [13], essentially by the normalization term $t-i$ in (4.4), which compensates for the non-compact support of dN. Alternatively, one can work with the normalized Lyapunov exponent $\gamma-\gamma_0$ and integrated density of states $N-N_0$, where γ_0 and N_0 are the corresponding 'free' quantities, see [3]. We decided to work with (4.4) since it was shown in Appendix 2 of [29] that Kotani theory applies to singular potentials. While [29] works with L^2-assumptions on the potential, it is pointed out that all what is needed to make Kotani theory work are two basic estimates for the m-function $m(z,\omega)$ on $(0,\infty)$ of an ergodic random operator: If K is a compact subset of $\mathbb{C}^+$, the complex, upper-half plane, then there are constants $C(K)<\infty$ and $\delta(K)>0$ such that

$$|m(z,\omega)| \leq C(K), \tag{4.5}$$

and

$$|\mathrm{Im}[m(z,\omega)]| \geq \delta(K), \tag{4.6}$$

uniformly for $z \in K$ and $\omega \in \Omega$. In our model (1.2), (1.3) we have only used L^1_{loc}-assumptions on the potentials. That (4.5) is satisfied in this case follows from results on m-function asymptotics in [24]. To get (4.6) we use that

$$\mathrm{Im}[m(z)] = \mathrm{Im}[z] \cdot \int_0^\infty |u|^2 dx, \tag{4.7}$$

where u is the solution of $-u''+(V_{\rm per}+V_\omega)u = zu$ with $u(0)=1$ and $u'(0)=m(z)$, e.g. [14]. Combining Lemma A.1 and A.3, we get

$$\int_0^2 |u|^2 dx \geq C_1\left(|u(1)|^2+|u'(1)|^2\right) \geq C_1C_2\left(1+|m(z)|^2\right).$$

Thus (4.7) yields (4.6).

(ii) To apply Kotani's results from [34], one must have an $\mathbb{R}$-ergodic system, i.e. there must exist a group $\{\theta_t : t\in\mathbb{R}\}$ of measure preserving transformations for which the dynamical system $(\Omega,\mathcal{F},\theta_t,\mu)$ is ergodic and the random potential v_ω satisfies $v_{\theta_t\omega}(x) = v_\omega(x+t)$. Kirsch's result in [28] shows how a $\mathbb{Z}$-ergodic system, i.e. one for which the transformations θ_t are parametrized by $t\in\mathbb{Z}$, can be associated with an $\mathbb{R}$-ergodic system, embedded in a larger probability space. The associated system is constructed in such a way that the corresponding integrated density of states and Lyapunov exponent for both systems are equal. Our model is $\mathbb{Z}$-ergodic when equipped with translations. Thus we must apply Kotani's results to the corresponding $\mathbb{R}$-ergodic system after using Kirsch's suspension procedure.

(iii) In Kotani's work [34], N and γ arise as the real and imaginary parts of the non-tangential limit of a specific Herglotz function (the w-function). To see that the integrated density of states and the Lyapunov exponent actually coincide with these non-tangential limits, see Proposition V.12 and Proposition VI.1 of [11].

Based on (4.4), the proof of Theorem 4.1 is very similar to the proof of the corresponding result in the discrete case, see Proposition VI.3.9 of [13]. Some small changes arise from the use of the slightly different Thouless formula (4.4) and the necessity to exclude the set M. We use some basic facts about Hilbert transforms, which for a square integrable function ψ is defined by:

$$(T\psi)(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0^+}\int_{|x-t|>\varepsilon}\frac{\psi(t)}{x-t}dt.$$

One has two basic results, as stated in [13],
(i) $T\psi$ is a square integrable function and $T^2\psi = -\psi$ almost everywhere with respect to Lebesgue measure.
(ii) If ψ is Hölder continuous on some interval $[-a,a]$, then $T\psi$ is Hölder continuous on $[-\frac{a}{2},\frac{a}{2}]$.

Note that if $\psi_a(t) := \psi(t-a)$, then $T\psi_a = (T\psi)_a$. Thus if ψ is Hölder continuous on some interval $[x_0-a, x_0+a]$, then $T\psi$ is Hölder continuous on $[x_0-\frac{a}{2}, x_0+\frac{a}{2}]$.

Proof of Theorem 4.1. First, observe that dN-integrability of $\log|(\lambda-t)/(t-i)|$ yields

$$\lim_{\varepsilon\to 0^+}\int_{\lambda-\varepsilon}^{\lambda+\varepsilon}\left|\log|\frac{\lambda-t}{t-i}|\right| dN(t) = 0,$$

which easily implies that

$$\lim_{\varepsilon\to 0^+}|\log(\varepsilon)|\,(N(\lambda+\varepsilon)-N(\lambda-\varepsilon)) = 0, \tag{4.8}$$

i.e. log-Hölder continuity and, in particular, continuity of N.

Let $\lambda_0\in I$ and pick $a>0$ such that $[\lambda_0-4a,\lambda_0+4a]\subset\mathbb{R}\setminus M$ and thus γ is Hölder continuous in $[\lambda_0-4a,\lambda_0+4a]$. Take $\psi(t) := N(t)\chi_{\{t:|t-\lambda_0|\leq 4a\}}$, and note that $(T^2\psi)(t) = -N(t)$ for almost every t with $|t-\lambda_0| \le 4a$ by (i). For $|\lambda-\lambda_0| < 4a$ calculate,

$$\begin{aligned}
\gamma(\lambda) + \alpha - \int_{|t-\lambda_0|>4a} \log\left|\frac{\lambda-t}{t-i}\right| dN(t) &= \int_{\lambda_0-4a}^{\lambda_0+4a} \log\left|\frac{\lambda-t}{t-i}\right| dN(t) \\
&= \lim_{\varepsilon\to 0^+}\left(\int_{\lambda+\varepsilon}^{\lambda_0+4a} \log|\lambda-t| dN(t) + \int_{\lambda_0-4a}^{\lambda-\varepsilon} \log|\lambda-t| dN(t)\right) \\
&\quad - \frac{1}{2}\int_{\lambda_0-4a}^{\lambda_0+4a} \log(t^2+1) dN(t).
\end{aligned}$$

Integrating the first two integrals above by parts and rearranging yields

(4.9)
$$\pi(T\psi)(\lambda) = \gamma(\lambda) + \alpha - \int_{|t-\lambda_0|>4a} \log\left|\frac{\lambda-t}{t-i}\right| dN(t) + \frac{1}{2}\int_{\lambda_0-4a}^{\lambda_0+4a} \log(t^2+1) dN(t)$$
$$- \log(\lambda_0-\lambda+4a)N(\lambda_0+4a) + \log(\lambda-\lambda_0+4a)N(\lambda_0-4a),$$

where (4.8) has been used. From observing (4.9), we see that $T\psi$ is Hölder continuous in $[\lambda_0-2a, \lambda_0+2a]$, and thus $T^2\psi$ is Hölder continuous in $[\lambda_0-a, \lambda_0+a]$ by (ii). Since N is continuous we get $(T^2\psi)(t) = -N(t)$ for all $t \in [\lambda_0-a, \lambda_0+a]$. Thus N is Hölder continuous in $[\lambda_0-a, \lambda_0+a]$. A compactness argument yields uniform Hölder continuity over all of I. □

5. The Wegner Estimate

In this section we prove a Wegner estimate which constitutes one of the two ingredients that will enable us to start the multiscale induction. Given Hölder continuity of the integrated density of states, as proven in the preceding section, our proof of the Wegner estimate can be carried out in a way analogous to [12]. Since a few modifications of the arguments in [12] are required, we present sufficiently many details for the reader's convenience.

Fix throughout this section a compact interval $I = [a, b] \subset \mathbb{R}\setminus M$, where M is the discrete set found in Theorem 2.3. Our goal will be to prove estimates uniformly in I. It will therefore be convenient to have these two properties in a ball of fixed radius around each point in I. By our above results we know that there is some $\xi > 0$ such that for every $\lambda \in I_\xi = [a-\xi, b+\xi]$, $G(\lambda)$ is non-compact and strongly irreducible. In other words, for every $\lambda \in I$, we have that $G(\lambda')$ is non-compact and strongly irreducible for every $\lambda' \in [\lambda-\xi, \lambda+\xi]$.

Let, for odd $L \in \mathbb{N}$ and $\Lambda = \Lambda_L(0) = [-L/2, L/2]$, $H_\Lambda(\omega)$ be the restriction of H_ω to Λ with Dirichlet boundary conditions at $-L/2$ and $L/2$. We will prove the following theorem:

Theorem 5.1 (Wegner estimate)**.** *For every $\beta \in (0,1)$ and every $\sigma > 0$, there exist $L_0 \in \mathbb{N}$ and $\alpha > 0$ such that*

$$\mathbb{P}\{\mathrm{dist}(\lambda, \sigma(H_\Lambda(\omega)) \le e^{-\sigma L^\beta}\} \le e^{-\alpha L^\beta} \qquad (5.1)$$

for all $\lambda \in I$ and $L \ge L_0$.

To prove Theorem 5.1, we will need two lemmas, Lemma 5.2 and Lemma 5.3 below.

Lemma 5.2. *There exist $\alpha_1 > 0, \delta > 0, n_0 \in \mathbb{N}$ such that for all $\lambda \in I, n \geq n_0$, and x normalized, we have*

$$\mathbb{E}\{\|U_\lambda(n)x\|^{-\delta}\} \leq e^{-\alpha_1 n}.$$

Proof. This lemma can be proved in exactly the same way Lemma 5.1 is proved in [12]. For the reader's convenience, we sketch the argument briefly. By our choice of the interval I and the results from Sections 2 and 3, in particular Corollary 3.3 (i), we have $\gamma := \inf\{\gamma(\lambda) : \lambda \in I\} > 0$. Using the inequality $e^y \leq 1 + y + y^2 e^{|y|}$ and Hölder's inequality, one shows as in [12] that for every $\lambda \in I$ and every $\delta > 0$, we have

$$\begin{aligned}
\mathbb{E}\{\|U_\lambda(n,\omega)x\|^{-\delta}\} =& \mathbb{E}\{\|g_\lambda(n,\omega)\cdot\ldots\cdot g_\lambda(1,\omega)x\|^{-\delta}\} \\
\leq& 1 - \delta\mathbb{E}\{\log\|g_\lambda(n,\omega)\cdot\ldots\cdot g_\lambda(1,\omega)x\|\} \\
&+ \delta^2 n \left[\mathbb{E}\{(\log\|g_\lambda(1,\omega)\|)^4\}\right]^{1/2}\left[\mathbb{E}\{\|g_\lambda(1,\omega)\|^{2\delta}\}\right]^{n/2} \\
\leq& 1 - \delta\mathbb{E}\{\log\|g_\lambda(n,\omega)\cdot\ldots\cdot g_\lambda(1,\omega)x\|\} + \delta^2 n^2 C_1 C_2^n
\end{aligned}$$

for some finite constants $C_1 = C_1(I)$, $C_2 = C_2(I)$. Hence, by Corollary 3.3 (ii), we have for some $n_0 = n_0(I)$, uniformly in $\lambda \in I$ and x in the unit sphere,

$$\mathbb{E}\{\|U_\lambda(n_0,\omega)x\|^{-\delta}\} \leq 1 - \tfrac{1}{2}n_0\delta\gamma + \delta^2 n_0^2 C_1 C_2^{n_0} \leq 1 - \varepsilon \tag{5.2}$$

for some $\varepsilon > 0$, provided δ is small enough. Iterating (5.2) as in [12] yields

$$\mathbb{E}\{\|U_\lambda(n,\omega)x\|^{-\delta}\} \leq C(1-\varepsilon)^{\lfloor n/n_0\rfloor} \leq e^{-\alpha_1 n}$$

for all $n \geq n_1$ and $\lambda \in I$, for some $\alpha_1 = \alpha_1(I) > 0$ and $n_1 = n_1(I)$, where $\lfloor n/n_0\rfloor$ is the largest integer less or equal n/n_0. ☐

Lemma 5.3. *There exist $\rho > 0$ and $C < \infty$ such that for every $\lambda \in I$ and every $\varepsilon > 0$, we have for $L \geq L_0$,*

$$\begin{aligned}
\mathbb{P}\{&\textit{There exist } \lambda' \in (\lambda-\varepsilon,\lambda+\varepsilon) \textit{ and } \phi \in D(H_\Lambda),\ \|\phi\| = 1, \textit{ such that} \\
&(H_\Lambda(\omega)-\lambda')\phi = 0,\ |\phi'(-L/2)|^2 + |\phi'(L/2)|^2 \leq \varepsilon^2\} \leq CL\varepsilon^\rho.
\end{aligned} \tag{5.3}$$

Proof. We will follow the same strategy as Carmona et al. in their proof of [12, Lemma 5.2], that is, we will use Hölder continuity of the integrated density of states to derive the estimate (5.3). The only difficulty that arises is that the cutoff of eigenfunctions as performed by Carmona et al. in the discrete case may produce elements outside the domain of the local Hamiltonian. We therefore use a smooth cutoff procedure and show that the argument still goes through.

Note first that it suffices to prove (5.3) for small $\varepsilon > 0$. It follows from Theorem 4.1 that there are constants $\rho = \rho(I_\xi) > 0$ and $C_1 = C_1(I_\xi) < \infty$ such that for every $\lambda, \lambda' \in I_\xi$,

$$|N(\lambda) - N(\lambda')| \leq C_1|\lambda - \lambda'|^\rho. \tag{5.4}$$

Now fix $\lambda \in I$, $\varepsilon > 0$, and L. Let, for $k \in \mathbb{Z}$, Λ_k be the interval $[kL - L/2, kL + L/2]$ and denote by $H_{\Lambda_k}(\omega)$ the operator $H(\omega)$ restricted to Λ_k with Dirichlet boundary conditions.

Let A_k be the event $A_k = A_k(\lambda, \varepsilon, L) = \{\omega \in \Omega : H_{\Lambda_k}(\omega)$ has an eigenvalue $\lambda_k \in (\lambda - \varepsilon, \lambda + \varepsilon)$ such that the corresponding normalized eigenfunction ϕ_k satisfies $|\phi_k'(kL - L/2)|^2 + |\phi_k'(kL + L/2)|^2 \leq \varepsilon^2\}$. Let $p = p(\lambda, \varepsilon, L) = \mathbb{P}\{A_k\}$. Clearly, p is independent of k and equals the left hand side of (5.3).

Fix some $n \in \mathbb{N}$ and let $H_n(\omega)$ be the operator $H(\omega)$ restricted to $\Lambda(n) := \bigcup_{k=-n}^{n} \Lambda_k = [-nL - L/2, nL + L/2]$ with Dirichlet boundary conditions at $-nL - L/2$ and $nL + L/2$. Let $k_1, \ldots, k_j \in \{-n, \ldots, n\}$ be distinct and such that the event A_{k_l} occurs. For each such l, we shall construct a normalized function $\tilde{\phi}_l$ in the domain of $H_n(\omega)$ which is supported on Λ_{k_l} (hence $\{\tilde{\phi}_1, \ldots, \tilde{\phi}_j\}$ form an orthonormal set) such that for every l,

$$\|(H_n(\omega) - \lambda)\tilde{\phi}_l\| \leq C_2 \varepsilon, \tag{5.5}$$

where C is a constant which only depends on the single site potential f and the single site distribution μ. By disjointness of supports, we also have

$$\langle \tilde{\phi}_l, H_n(\omega)\tilde{\phi}_{l'} \rangle = 0 = \langle H_n(\omega)\tilde{\phi}_l, H_n(\omega)\tilde{\phi}_{l'} \rangle \text{ for } l \neq l'. \tag{5.6}$$

By [41, Lemma A.3.2], (5.5) and (5.6) imply that the number of eigenvalues of $H_n(\omega)$ (counted with multiplicity) in $[\lambda - C_2\varepsilon, \lambda + C_2\varepsilon]$ is bounded from below by j. In other words,

$$\#\{l \in \{-n, \ldots, n\} : A_l \text{ occurs}\} \leq \#\{\text{eigenvalues of } H_n(\omega) \text{ in } [\lambda - C_2\varepsilon, \lambda + C_2\varepsilon]\}.$$

Thus, if ε is small enough (namely, such that $[\lambda - C_2\varepsilon, \lambda + C_2\varepsilon] \subseteq I_\xi$ or, equivalently, $\varepsilon \leq \xi/C_2$), we have

$$\begin{aligned}
p &= \lim_{n\to\infty} \frac{1}{2n+1} \#\{l \in \{-n, \ldots, n\} : A_l \text{ occurs}\} \\
&\leq \lim_{n\to\infty} \frac{1}{2n+1} \#\{\text{eigenvalues of } H_n(\omega) \text{ in } [\lambda - C_2\varepsilon, \lambda + C_2\varepsilon]\} \\
&= L \lim_{n\to\infty} \frac{1}{(2n+1)L} \#\{\text{eigenvalues of } H_n(\omega) \text{ in } [\lambda - C_2\varepsilon, \lambda + C_2\varepsilon]\} \\
&= L(N(\lambda + C_2\varepsilon) - N(\lambda - C_2\varepsilon)) \\
&\leq LC_1(2C_2\varepsilon)^\rho \\
&=: CL\varepsilon^\rho,
\end{aligned}$$

where the intermediate steps hold true for almost every ω.

It remains to construct $\tilde{\phi}_l$ with the desired properties. Fix some l and consider the function ϕ_l which is defined on Λ_{k_l} and vanishes at the boundary points. In principle we would like to extend ϕ_l by zero on $\Lambda(n) \setminus \Lambda_{k_l}$. However, this function will in general not belong to the domain of H_n, so we could not even evaluate $(H_n - \lambda)$ applied to this function. Instead we will use a smooth extension of ϕ_l to $\Lambda(n)$.

Fix once and for all a smooth function χ which obeys $0 \leq \chi \leq 1$, $\chi(x) = 0$ for $x \leq 0$, and $\chi(x) = 1$ for $x \geq 1$. Let $x_l^L = k_l L - L/2$ and $x_l^R = k_l L + L/2$ and define

$\hat\phi_l$ for $x \in \Lambda(n)$ by

$$\hat\phi_l(x) = \begin{cases} 0 & \text{if} \quad x \notin \Lambda_{k_l} \\ \chi(x - x_l^L)\phi_l(x) & \text{if} \quad x_l^L \le x \le x_l^L + 1 \\ \phi_l(x) & \text{if} \quad x_l^L + 1 \le x \le x_l^R - 1 \\ \chi(-x + x_l^R)\phi_l(x) & \text{if} \quad x_l^R - 1 \le x \le x_l^R \end{cases}$$

Then $\hat\phi_l$ clearly belongs to the domain of H_n and it has norm bounded by $\|\phi_l\| = 1$. We want to estimate $\|(H_n - \lambda)\hat\phi_l\|$. Now ϕ_l is an eigenfunction corresponding to the eigenvalue $\lambda_{k_l} \in (\lambda - \varepsilon, \lambda + \varepsilon)$, so we write $\|(H_n - \lambda)\hat\phi_l\| \le \|(H_n - \lambda_{k_l})\hat\phi_l\| + \varepsilon$. To estimate $\|(H_n - \lambda_{k_l})\hat\phi_l\|$, we consider

$$(H - \lambda_{k_l})\hat\phi_l(x) = -\hat\phi_l''(x) + V_\omega(x)\hat\phi_l(x) - \lambda_{k_l}\hat\phi_l(x)$$

for $x \in \Lambda(n)$. Hence, we have

$$\begin{aligned} \|(H_n - \lambda_{k_l})\hat\phi_l\|^2 = &\int_{x_l^L}^{x_l^L+1} |\chi''(x - x_l^L)\phi_l(x) + 2\chi'(x - x_l^L)\phi_l'(x)|^2 dx \\ &+ \int_{x_l^R-1}^{x_l^R} |\chi''(-x + x_l^R)\phi_l(x) + 2\chi'(-x + x_l^R)\phi_l'(x)|^2 dx \\ \le &\left\| \begin{pmatrix} \phi_l \\ \phi_l' \end{pmatrix} \right\|^2_{L^\infty(x_l^L, x_l^L+1)} \cdot \int_{x_l^L}^{x_l^L+1} \left\| \begin{pmatrix} \chi''(x - x_l^L) \\ 2\chi'(x - x_l^L) \end{pmatrix} \right\|^2 dx \\ &+ \left\| \begin{pmatrix} \phi_l \\ \phi_l' \end{pmatrix} \right\|^2_{L^\infty(x_l^R-1, x_l^R)} \cdot \int_{x_l^R-1}^{x_l^R} \left\| \begin{pmatrix} \chi''(-x + x_l^R) \\ 2\chi'(-x + x_l^R) \end{pmatrix} \right\|^2 dx \\ \le &C_3 \left[\left\| \begin{pmatrix} \phi_l(x_l^L) \\ \phi_l'(x_l^L) \end{pmatrix} \right\|^2 + \left\| \begin{pmatrix} \phi_l(x_l^R) \\ \phi_l'(x_l^R) \end{pmatrix} \right\|^2 \right] \\ = &C_3 \left[|\phi_l'(x_l^L)|^2 + |\phi_l'(x_l^R)|^2 \right] \\ \le &C_3 \varepsilon^2, \end{aligned}$$

where the constant C_3 depends on the single site potential, the single site distribution, and the function χ. Here we have used Lemma A.1 in step 3. Let us define $\tilde\phi_l := \hat\phi_l / \|\hat\phi_l\|$. By construction and Lemma A.1, we have, for ε sufficiently small, $\|\hat\phi_l\| \ge 1/2$ and hence (5.5) holds true with a suitable C_2 which depends only on the single site potential, the single site distribution, and the function χ. Moreover, by construction $\{\tilde\phi_1, \dots, \tilde\phi_j\}$ form an orthonormal set and obey (5.6). This concludes the proof of the lemma. ☐

We are now in position to give the

Proof of Theorem 5.1. We closely follow the proof of Theorem 4.1 in [12] and make the necessary modifications. Let β, σ, I be as above and for each odd $L \in \mathbb{N}$, we set $n_L = \lfloor \tau(L/2)^\beta \rfloor + 1$ with some $\tau > 0$ to be chosen later. For every $\lambda \in I$ and $\theta > 0$, we define the events

$$\begin{aligned} A_\theta^{(\lambda,L)} &= \{\|g_\lambda(-(L+1)/2 + n_L) \cdot \ldots \cdot g_\lambda(-(L-1)/2)(0,1)^t\| > e^{\theta(L/2)^\beta}\}, \\ B_\theta^{(\lambda,L)} &= \{\|g_\lambda((L+1)/2 - n_L)^{-1} \cdot \ldots \cdot g_\lambda((L-1)/2)^{-1}(0,1)^t\| > e^{\theta(L/2)^\beta}\}. \end{aligned}$$

Let $\kappa = \tau\alpha_1/2\delta$ with α_1 and δ from Lemma 5.2. Then

$$\begin{aligned}
&\mathbb{P}\{\mathrm{dist}(\lambda, \sigma(H_\Lambda(\omega)) \leq e^{-\sigma L^\beta}\} \leq \\
&\quad \leq \mathbb{P}\left\{ \{\mathrm{dist}(\lambda, \sigma(H_\Lambda(\omega)) \leq e^{-\sigma L^\beta}\} \cap \bigcap_{|\lambda-\lambda'|\leq e^{-\sigma L^\beta}} \left(A_{\kappa/2}^{(\lambda',L)} \cap B_{\kappa/2}^{(\lambda',L)} \right) \right\} \\
&\quad + \mathbb{P}\left\{ A_\kappa^{(\lambda,L)} \cap B_\kappa^{(\lambda,L)} \cap \bigcup_{|\lambda-\lambda'|\leq e^{-\sigma L^\beta}} (A_{\kappa/2}^{(\lambda',L)})^c \right\} \\
&\quad + \mathbb{P}\left\{ A_\kappa^{(\lambda,L)} \cap B_\kappa^{(\lambda,L)} \cap \bigcup_{|\lambda-\lambda'|\leq e^{-\sigma L^\beta}} (B_{\kappa/2}^{(\lambda',L)})^c \right\} \\
&\quad + \mathbb{P}\{(A_\kappa^{(\lambda,L)})^c\} + \mathbb{P}\{(B_\kappa^{(\lambda,L)})^c\} \\
&\quad = \text{(i)} + \text{(ii)} + \text{(iii)} + \text{(iv)}
\end{aligned}$$

and Lemma 5.2 immediately implies

$$\text{(iv)} \leq 2e^{-\frac{1}{2}\tau\alpha_1(L/2)^\beta} \tag{5.7}$$

provided L is large enough. Together with Lemma A.3, Lemma 5.3 yields for L large enough,

$$\text{(i)} \leq \tilde{C}L \max\left\{ e^{-\sigma L^\beta \rho}, e^{-\frac{1}{2}\kappa(L/2)^\beta \rho} \right\} \tag{5.8}$$

with $\tilde{C}$ independent of λ and L.

Finally, using (3.3), for L large enough, one proves similarly to [12]

$$\text{(ii)} + \text{(iii)} \leq 2e^{-\alpha_2(L/2)^\beta} \tag{5.9}$$

for some suitable $\alpha_2 > 0$ if τ is chosen small enough. The assertion now follows from (5.7)—(5.9). ☐

6. The Initial Length Scale Estimate

Fix $\lambda \in \mathbb{R} \setminus M$ and let ν_λ be the unique μ_λ-invariant measure on $P(\mathbb{R}^2)$, see (3.6), and $\delta(\overline{x}, \overline{y})$ the projective distance of $\overline{x}, \overline{y} \in P(\mathbb{R}^2)$.

The measure ν_λ is Hölder continuous:

Lemma 6.1. *There exist $\rho > 0$ and $C > 0$ such that for all $\overline{x} \in P(\mathbb{R}^2)$ and $\varepsilon > 0$, one has*

$$\nu_\lambda(\{\overline{y} : \delta(\overline{x}, \overline{y}) \leq \varepsilon\}) \leq C\varepsilon^\rho.$$

Proof. This follows from Corollary VI.4.2 of [8]. Note that in the case $G(\lambda) \subset \mathrm{SL}(2, \mathbb{R})$ the assumptions required there are equivalent to $\int \|g\|^\tau d\mu_\lambda(g) < \infty$ for some $\tau > 0$ and that $G(\lambda)$ is non-compact and strongly irreducible. In particular, non-compactness is equivalent to the contractivity required in [8]; see [13, Prop. IV.4.11]. Integrability of $\|g\|^\tau$ with respect to μ_λ for all $\tau > 0$ follows from (3.2) and boundedness of the distribution of q_n. ☐

We know that $\frac{1}{n} \log \|U_\lambda(n)\|$ converges to $\gamma(\lambda)$ in expectation. We will need a large deviation result for this limit. In fact, the following result on the asymptotics of $\|U_\lambda(n)x\|$ for any initial vector $x \neq 0$ holds.

Lemma 6.2. *There exists $\alpha > 0$ such that for every $\varepsilon > 0$ and $x \neq 0$, one has*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(|\log \|U_\lambda(n)x\| - n\gamma(\lambda)| > n\varepsilon) < \alpha.$$

Proof. This follows from Theorem V.6.2 of [8], whose assumptions, that is, $G(\lambda)$ strongly irreducible and $\int \|g\|^\tau d\mu_\lambda(g) < \infty$ for some $\tau > 0$, are satisfied. ☐

An immediate consequence of Lemma 6.2 is that for every $\varepsilon > 0$ and $x \neq 0$, there exists $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}\left(e^{(\gamma(\lambda)-\varepsilon)n} \leq \|U_\lambda(n)x\| \leq e^{(\gamma(\lambda)+\varepsilon)n}\right) \geq 1 - e^{-\alpha n} \text{ for } n \geq n_0. \tag{6.1}$$

Next, we establish a large deviation result for $|\langle U_\lambda(n)x, y\rangle|$, that is, in particular for the matrix elements of the transfer matrices.

Lemma 6.3. *Fix y with $\|y\| = 1$. For all $\varepsilon > 0$, there are $n_0 \in \mathbb{N}$ and $\delta_0 > 0$ such that*

$$\sup_{x\neq 0} \mathbb{P}\left(\frac{|\langle U_\lambda(n)x, y\rangle|}{\|U_\lambda(n)x\|} < e^{-\varepsilon n}\right) < e^{-\delta_0 n} \text{ for } n \geq n_0.$$

Proof. We closely follow the proof of Prop. VI.2.2 in [8]. Define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq e^{-\varepsilon n}, \\ 2 - te^{\varepsilon n} & \text{if } e^{-\varepsilon n} \leq t \leq 2e^{-\varepsilon n}, \\ 0 & \text{if } 2e^{-\varepsilon n} \leq t \leq 1. \end{cases}$$

Then

$$|f_n(t) - f_n(t')| \leq |t - t'|e^{\varepsilon n} \text{ for all } t, t' \in [0,1]. \tag{6.2}$$

Define $\Phi_n : P(\mathbb{R}^2) \to \mathbb{R}$ by

$$\Phi_n(\overline{z}) = f_n\left(|\langle \tfrac{z}{\|z\|}, y\rangle|\right).$$

For $\overline{z}, \overline{z}' \in P(\mathbb{R}^2)$, choose representatives z and z' such that $\|z\| = \|z'\| = 1$ and the angle between z and z' is at most $\frac{\pi}{2}$. Then, by (6.2),

$$|\Phi_n(\overline{z}) - \Phi_n(\overline{z}')| \leq ||\langle z, y\rangle| - |\langle z', y\rangle||e^{\varepsilon n} \leq \|z - z'\|e^{\varepsilon n} \leq \sqrt{2}\delta(\overline{z}, \overline{z}')e^{\varepsilon n}, \tag{6.3}$$

recall that $\delta(\overline{z}, \overline{z}') = |\sin(\text{angle}(z, z'))|$. This implies, using notation from Section 3,

$$\|\Phi_n\|_\alpha = \|\Phi_n\|_\infty + m_\alpha(\Phi_n) \leq \sqrt{2}1 + e^{\varepsilon n} \text{ for all } 0 < \alpha < 1. \tag{6.4}$$

The definition of Φ_n shows

$$\begin{aligned} \mathbb{P}\left(\frac{|\langle U_\lambda(n)x, y\rangle|}{\|U_\lambda(n)x\|} < e^{-\varepsilon n}\right) &\leq \mathbb{E}(\Phi_n(U_\lambda(n)\overline{x})) \\ &\leq \left|\mathbb{E}(\Phi_n(U_\lambda(n)\overline{x})) - \int \Phi_n d\nu_\lambda\right| + \int \Phi_n d\nu_\lambda. \end{aligned} \tag{6.5}$$

Noting that $\mathbb{E}(\Phi_n(U_\lambda(n)\overline{x})) = (T_\lambda^n \Phi_n)(\overline{x})$, we may use Lemma 3.5 above to conclude that for $0 < \alpha \leq \alpha_0$, there exists $\rho < 1$ and $n_0 \in \mathbb{N}$ such that for $n \geq n_0$ and all x,

$$\left\| \mathbb{E}(\Phi_n(U_\lambda(n)\overline{x})) - \int \Phi_n d\nu_\lambda \right\| \leq C_\alpha \rho^n \|\Phi_n\|_\alpha \leq \rho^n(\sqrt{2}e^{\varepsilon n}+1) \leq \tfrac{1}{2}e^{-\delta_1 n}, \tag{6.6}$$

where in the last step, ε is assumed to be sufficiently small, that is, such that $\log\rho+\varepsilon<0$, and $\delta_1 := \frac{1}{2}|\log\rho+\varepsilon|$.

Next choose the unit vector $w = (w_1, w_2)^t = (y_2, -y_1)^t$. We therefore have $|\langle u/\|u\|, y\rangle| = \delta(\overline{u}, \overline{w})$ for all $u \neq 0$. It then follows from Lemma 6.1 that there exist $\beta > 0$ and $C > 0$ such that

$$\begin{aligned} \int \Phi_n d\nu_\lambda &\leq \nu_\lambda\{\overline{u} : |\langle u/\|u\|, y\rangle| \leq 2e^{-\varepsilon n}\} \\ &= \nu_\lambda\{\overline{u} : \delta(\overline{u}, \overline{w}) \leq 2e^{-\varepsilon n}\} \leq 2^\beta C e^{-\varepsilon\beta n}. \end{aligned} \tag{6.7}$$

Inserting (6.6) and (6.7) into (6.5) completes the proof of Lemma 6.3 with $0 < \delta_0 < \min\{\delta_1, \varepsilon\beta\}$. We have assumed that ε is sufficiently small, but the result extends to large ε with unchanged δ_0. ∎

Note that results corresponding to Lemma 6.2 and Lemma 6.3 also hold for $n \to -\infty$ since the result from [8] can be applied in the same way to the products of random matrices $g_\lambda^{-1}(n) \times \cdots \times g_\lambda^{-1}(-1)$ for $n < 0$ and the Lyapunov exponent can equivalently be defined as

$$\gamma(\lambda) = \lim_{n\to-\infty} \frac{1}{|n|} \mathbb{E}\left(\|g_\lambda^{-1}(n) \cdot \ldots \cdot g_\lambda^{-1}(-1)\|\right).$$

We can now combine Lemma 6.2 and Lemma 6.3 to show that with large probability, the matrix elements $|\langle U_\lambda(n)x, y\rangle|$ grow exponentially at the rate of almost the Lyapunov exponent.

Corollary 6.4. *Let $\|x\| = \|y\| = 1$. Then for every $\varepsilon > 0$, there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that*

$$\mathbb{P}\left(|\langle U_\lambda(n)x, y\rangle| \geq e^{(\gamma(\lambda)-\varepsilon)n}\right) \geq 1 - e^{-\delta n} \textit{ for } n \geq n_0.$$

Proof. From (6.1) and Lemma 6.3 we get for all $\varepsilon > 0$,

$$|\langle U_\lambda(n)x, y\rangle| \geq e^{-\varepsilon n}\|U_\lambda(n)x\| \geq e^{(\gamma(\lambda)-2\varepsilon)n} \text{ for } n \geq n_0$$

with probability at least $1 - e^{-\alpha n} - e^{-\delta_0 n}$. This yields the assertion. ∎

We are now ready to state and prove the main result of this section, an initial length scale estimate at energy λ. Let $L \in 3\mathbb{Z} \setminus 6\mathbb{Z}$ and $\Lambda = \Lambda_L(0) = [-L/2, L/2]$. Let $H_\Lambda(\omega)$ be the restriction of H_ω to Λ with Dirichlet boundary conditions at $-L/2$ and $L/2$. For $\lambda \notin \sigma(H_\Lambda(\omega))$, let $R_\Lambda(\lambda) = (H_\Lambda(\omega) - \lambda)^{-1}$. Define the characteristic functions $\chi^{\text{int}} = \chi_{\Lambda_{L/3}(0)} = \chi_{[-L/6, L/6]}$ and $\chi^{\text{out}} = \chi_{\Lambda_L(0)\setminus\Lambda_{L-2}(0)} = \chi_{[-L/2,-L/2+1]\cup[L/2-1,L/2]}$.

For $\gamma, \lambda \in \mathbb{R}$ and $\omega \in \Omega$, let us call the cube Λ (γ, λ)-good for ω if $\lambda \notin \sigma(H_\Lambda(\omega))$ and $\|\chi^{\text{out}} R_\Lambda(\lambda) \chi^{\text{int}}\| \leq e^{-\gamma L/3}$. We have the following theorem:

Theorem 6.5. *For every $\varepsilon > 0$, there exist $\delta > 0$ and $L_0 \in \mathbb{N}$ such that for $L \geq L_0$ with $L \in 3\mathbb{Z} \setminus 6\mathbb{Z}$, we have*

$$\mathbb{P}\left\{\Lambda \textit{ is } (\gamma(\lambda)-\varepsilon, \lambda)\textit{-good for } \omega\right\} \geq 1 - e^{-\delta L}. \tag{6.8}$$

Proof. Let $u_\pm$ be the solutions of $H_\omega u = \lambda u$ with Dirichlet boundary conditions at $\pm L/2$, that is, $u_+(L/2) = u_-(-L/2) = 0$, $u_+'(L/2) = u_-'(-L/2) = 1$. Then the Green's function $G_\Lambda(\lambda, x, y)$ (i.e., the kernel of $R_\Lambda(\lambda)$) is given by

$$G_\Lambda(\lambda, x, y) = \frac{1}{W(u_+, u_-)} \begin{cases} u_+(x)u_-(y) & \text{for } x \geq y \\ u_-(x)u_+(y) & \text{for } x < y \end{cases}, \tag{6.9}$$

where the Wronskian

$$W(u_+, u_-)(x) = u_+(x)u_-'(x) - u_+'(x)u_-(x)$$

is constant in x. Setting $x = L/2$, we get

$$W(u_+, u_-) = \langle (u_-(L/2), u_-'(L/2))^t, (-1, 0)^t \rangle = \langle A_\lambda(L/2, -L/2)(0,1)^t, (-1,0)^t \rangle,$$

where $A_\lambda(x, y)$ denotes the transfer matrix from y to x. By stationarity we can use Corollary 6.4 to conclude

$$\mathbb{P}\left(|W(u_+, u_-)| \geq e^{(\gamma(\lambda)-\varepsilon)L}\right) \geq 1 - e^{-\delta L} \text{ for } L \geq L_0. \tag{6.10}$$

Note that $\lambda \in \sigma(H_\Lambda(\omega))$ if and only if $W(u_+, u_-) = 0$. Thus the event in (6.10) implies $\lambda \notin \sigma(H_\Lambda(\omega))$.

Let $x \in [L/2 - 1, L/2]$ and $y \in [-L/6, L/6]$. Then

$$|u_+(x)| \leq C \text{ uniformly in } \omega \text{ and } L \tag{6.11}$$

by Lemma A.1. Also

$$\begin{aligned} |u_-(y)| &\leq \|(u_-(y), u_-'(y))^t\| = \|A_\lambda(y, -L/2)(0,1)^t\| \\ &\leq \|A_\lambda(y, \lfloor y + 1/2 \rfloor - 1/2)\| \cdot \|A_\lambda(\lfloor y + 1/2 \rfloor - 1/2, -L/2)(0,1)^t\|. \end{aligned}$$

Again by Lemma A.1,

$$\|A_\lambda(y, \lfloor y + 1/2 \rfloor - 1/2)\| \leq C, \tag{6.12}$$

and by stationarity and (6.1) (note $-L/6 \leq \lfloor y + 1/2 \rfloor - 1/2 \leq L/6$),

(6.13)
$$\mathbb{P}\left(\|A_\lambda(\lfloor y + 1/2 \rfloor - 1/2, -L/2)(0,1)^t\| \leq e^{(\gamma(\lambda)+\varepsilon)2L/3}\right) \geq 1 - e^{-\alpha L/3} \text{ if } L/3 \geq L_0.$$

Combining (6.10)—(6.13) and using $G_\Lambda(\lambda, x, y) = u_-(y)u_+(x)/W(u_+, u_-)$, we get

$$\mathbb{P}\left(|G_\lambda(\lambda, x, y)| \leq Ce^{-(\gamma(\lambda)-5\varepsilon)L/3}\right) \geq 1 - e^{-\delta L} - e^{-\alpha L/3}$$

for L sufficiently large. In a completely analogous way the same estimate is found if $x \in [-L/2, -L/2 + 1]$, $y \in [-L/6, L/6]$. From this it can now be seen easily that for every $\varepsilon > 0$, there exist $\delta > 0$ and $L_0 \in \mathbb{N}$ such that for $L \geq L_0$ and $L \in 3\mathbb{Z} \setminus 6\mathbb{Z}$, we have

$$\sup_x \int |\chi^{\text{out}}(x) G_\Lambda(\lambda, x, y) \chi^{\text{int}}(y)| dy \leq e^{-(\gamma(\lambda)-\varepsilon)L/3}$$

and

$$\sup_y \int |\chi^{\text{out}}(x) G_\Lambda(\lambda, x, y) \chi^{\text{int}}(y)| dx \leq e^{-(\gamma(\lambda)-\varepsilon)L/3}$$

with probability at least $1 - e^{-\delta L}$. The theorem now follows by Schur's test. □

Theorem 6.5 can be regarded as a fixed-energy initial length scale estimate. Our goal is to start a variable-energy multiscale induction which will ultimately yield both Theorem 1.1 and Theorem 1.2. We therefore need an estimate of the form (6.8) where the energy is not fixed, but rather varying over an interval. A result of this kind will be established, using the Wegner estimate and an argument from [19], in the following corollary.

Corollary 6.6. *For every $\lambda \in I$ and every $\beta \in (0,1)$, $\sigma, \varepsilon > 0$, let $\alpha > 0, \delta > 0$ and $L_0 \in \mathbb{N}$ be as in Theorem 5.1 and Theorem 6.5, respectively. For every $0 < \varepsilon < \varepsilon'$ and every $L \geq L_0$, let*

$$\kappa_L = \frac{1}{2} e^{-2\sigma L^\beta} \left(e^{-(\gamma(\lambda)-\varepsilon')L/3} - e^{-(\gamma(\lambda)-\varepsilon)L/3} \right).$$

Then we have for every $L \geq L_0$

$$\mathbb{P}\left\{\forall \lambda' \in (\lambda - \kappa_L, \lambda + \kappa_L) : \Lambda \text{ is } (\gamma(\lambda) - \varepsilon', \lambda')\text{-good}\right\} \geq 1 - e^{-\delta L} - e^{-\alpha L^\beta}. \tag{6.14}$$

Proof. With probability $1 - e^{-\delta L} - e^{-\alpha L^\beta}$ we have that both the event in (6.8) and the complementary event in (5.1) hold. Thus, by assumption we have that for every $\lambda' \in (\lambda - \kappa_L, \lambda + \kappa_L)$, we have $\lambda' \notin \sigma(H_\Lambda(\omega))$ and moreover, by the resolvent equation,

$$\begin{aligned}
\|\chi^{\text{out}} R_\Lambda(\lambda')\chi^{\text{int}}\| &= \|\chi^{\text{out}} \left(R_\Lambda(\lambda) + (\lambda - \lambda')R_\Lambda(\lambda')R_\Lambda(\lambda)\right) \chi^{\text{int}}\| \\
&\leq \|\chi^{\text{out}} R_\Lambda(\lambda)\chi^{\text{int}}\| + |\lambda - \lambda'| \cdot \|\chi^{\text{out}} R_\Lambda(\lambda')R_\Lambda(\lambda)\chi^{\text{int}}\| \\
&\leq \|\chi^{\text{out}} R_\Lambda(\lambda)\chi^{\text{int}}\| + |\lambda - \lambda'| \cdot \|R_\Lambda(\lambda')\| \cdot \|R_\Lambda(\lambda)\| \\
&\leq e^{-(\gamma(\lambda)-\varepsilon)L/3} + 2 \cdot |\lambda - \lambda'| \cdot e^{2\sigma L^\beta} \\
&\leq e^{-(\gamma(\lambda)-\varepsilon')L/3}.
\end{aligned}$$

Thus for these ω's, the cube Λ is $(\gamma(\lambda) - \varepsilon', \lambda')$-good. □

7. Proof of the Main Theorems

In the preceding two sections we have established the two ingredients, namely, a Wegner estimate and an initial length scale estimate, that are necessary to start the multiscale induction which, by known results, implies both Theorem 1.1 and Theorem 1.2. In this section we briefly show how to reduce these two theorems to known results, given Theorem 5.1 and Corollary 6.6 above.

Let M be the discrete set found in Theorem 2.3.

Proof of Theorem 1.1. Fix an arbitrary compact interval $I \subset \mathbb{R} \setminus M$. It follows from Theorem 5.1 and Theorem 6.5 that we have both a Wegner estimate and a fixed energy initial length scale estimate for every $\lambda \in I$. Corollary 6.6 shows that these two results imply a variable-energy initial length scale estimate for a ball $B(\lambda)$ of explicit radius around λ . The variable-energy multiscale analysis as presented, for example, in [44] then establishes variable-energy resolvent decay estimates on a sequence $(L_k)_{k\in\mathbb{N}}$ of length scales for energies in $B(\lambda)$. These estimates, together with the existence of polynomially bounded eigenfunctions for spectrally almost every energy, yield pure point spectrum in $B(\lambda)$ with exponentially decaying eigenfunctions for almost every $\omega \in \Omega$; see, for example, [44] for details. Thus, we have exponential localization in $\mathbb{R} \setminus M$ for almost every $\omega \in \Omega$. Finally, since, by general principles [13], the set M carries almost surely no spectral measure, we have exponential localization in $\mathbb{R}$ for almost every $\omega \in \Omega$ and hence Theorem 1.1. $\square$

Proof of Theorem 1.2. It essentially follows from [16] that the variable-energy resolvent decay estimates, as given by the output of the the variable-energy multiscale analysis, imply strong dynamical localization in the sense of Theorem 1.2. For the curious reader we briefly sketch the argument, referring him to [16] for necessary notation. Given a compact interval $I \subset \mathbb{R} \setminus M$, a compact set $K \subset \mathbb{R}$, and $p > 0$, we first let $\gamma = \min\{\gamma(\lambda) : \lambda \in I\} > 0$. Next we choose L_1 large enough so that, for every $\lambda \in I$, Theorem 5.1 and Corollary 6.6 imply both $W(I, L, \Theta, q)$, $L \geq L_1$, and $G(B(\lambda), L_1, \gamma - \varepsilon', \xi)$ of [16] with parameters sufficient to cover the desired p. Having this length scale fixed, we decompose the interval I into a finite disjoint union of intervals $I_1, \dots, I_m$, each of them having length bounded by κ_{L_1}. We split the projection $P_I(H_\omega)$ in (1.4) into the finite sum $\sum_{i=1}^m P_{I_i}(H_\omega)$ and treat each term separately. For every i, we can apply Theorem 3.1 of [16], with initial length scale L_1, since all the other conditions (e.g., (INDY), (GRI), (WEYL), (EDI), (i)) are known to hold for the concrete operators H_ω under consideration [44]. After establishing (1.4), with I replaced by I_i, for every i, we get (1.4) and hence Theorem 1.2 $\square$

Appendix A. A Priori Solution Estimates

Here we provide several a priori estimates for solutions of the Schrödinger equation which are used repeatedly in the main text. For convenience we include proofs of these standard facts.

Lemma A.1. *Let $q \in L^1_{\mathrm{loc}}(\mathbb{R})$, u a solution of $-u'' + qu = 0$ and $x, y \in \mathbb{R}$. Then*

$$|u(x)|^2 + |u'(x)|^2 \leq \left(|u(y)|^2 + |u'(y)|^2\right) \exp\left\{\int_{\min(x,y)}^{\max(x,y)} |q(t) + 1|\, dt\right\}.$$

Proof. For $R(t) := |u(t)|^2 + |u'(t)|^2$ one has

$$|R'(t)| = |2(q(t) + 1)\mathrm{Re} u'(t)\overline{u(t)}| \leq |q(t) + 1| R(t),$$

i.e. $|(\ln R(t))'| \leq |q(t) + 1|$, which implies the lemma. $\square$

Lemma A.2. *For $i = 1, 2$, let $q_i \in L^1_{\mathrm{loc}}(\mathbb{R})$, u_i solutions of $-u_i'' + q_i u_i = 0$ with $u_1(y) = u_2(y)$ and $u_1'(y) = u_2'(y)$ for some $y \in \mathbb{R}$. Then for any $x \in \mathbb{R}$,*

$$\begin{aligned}
&\left(|u_1(x) - u_2(x)|^2 + |u_1'(x) - u_2'(x)|^2\right)^{1/2} \leq \\
&\quad \left(|u_1(y)|^2 + |u_1'(y)|^2\right)^{1/2} \exp\left\{\int_{\min(x,y)}^{\max(x,y)} (|q_1(t)| + |q_2(t)| + 2)\, dt\right\} \times \\
&\quad \times \int_{\min(x,y)}^{\max(x,y)} |q_1(t) - q_2(t)|\, dt.
\end{aligned}$$

Proof. Without restriction let $y \le x$. The solutions u_1 and u_2 satisfy

$$\begin{pmatrix} u_1(x)-u_2(x) \\ u_1'(x)-u_2'(x) \end{pmatrix} = \int_y^x \begin{pmatrix} 0 \\ (q_1(t)-q_2(t))u_1(t) \end{pmatrix} dt + \\ + \int_y^x \begin{pmatrix} 0 & 1 \\ q_2(t) & 0 \end{pmatrix} \begin{pmatrix} u_1(t)-u_2(t) \\ u_1'(t)-u_2'(t) \end{pmatrix} dt,$$

and thus

$$\left\| \begin{pmatrix} u_1(x)-u_2(x) \\ u_1'(x)-u_2'(x) \end{pmatrix} \right\| \le \int_y^x |q_1(t)-q_2(t)||u_1(t)|\,dt + \\ + \int_y^x (|q_2(t)|+1) \left\| \begin{pmatrix} u_1(t)-u_2(t) \\ u_1'(t)-u_2'(t) \end{pmatrix} \right\| dt. \tag{1.1}$$

Gronwall's lemma, e.g. [47], yields

$$\left\| \begin{pmatrix} u_1(x)-u_2(x) \\ u_1'(x)-u_2'(x) \end{pmatrix} \right\| \le \int_y^x |q_1(t)-q_2(t)||u_1(t)|\,dt \exp\left\{ \int_y^x (|q_2(t)|+1)\,dt \right\}.$$

By Lemma A.1 we have for all $t \in [y,x]$ that

$$|u_1(t)| \le \exp\left\{ \frac{1}{2}\int_y^x |q_1(s)+1|\,ds \right\} \left(|u_1(y)|^2 + |u_1'(y)|^2 \right)^{1/2}.$$

Inserting this into (1.1) yields the result. ☐

Lemma A.3. *Let $q \in L^1_{\mathrm{loc,unif}}(\mathbb{R})$, i.e. $\|q\|_{1,\mathrm{unif}} := \sup_x \int_x^{x+1} |q(t)|\,dt < \infty$. Then there is a $C > 0$, only depending on $\|q\|_{1,\mathrm{unif}}$ such that for all solutions u of $-u'' + qu = 0$ and all $x \in \mathbb{R}$*

$$\int_{x-1}^{x+1} |u(t)|^2\,dt \ge C\left(|u(x)|^2 + |u'(x)|^2 \right). \tag{1.2}$$

Proof. By Lemma A.1 there are constants $0 < C_1, C_2 < \infty$ only depending on $\|q\|_{1,\mathrm{unif}}$ such that for all $t \in [x-1, x+1]$

$$C_1(|u(x)|^2 + |u'(x)|^2) \le |u(t)|^2 + |u'(t)|^2 \le C_2(|u(x)|^2 + |u'(x)|^2).$$

With $C_3 := (C_1/2)^{1/2}$ and $C_4 := (2C_2)^{1/2}$ we get

$$C_3(|u(x)| + |u'(x)|) \le |u(t)| + |u'(t)| \le C_4(|u(x)| + |u'(x)|).$$

It now follows from elementary geometric considerations, e.g. [45, p.218], that $[x-1, x+1]$ contains an interval of length $\min(2, C_3/4C_4)$ on which $|u| \ge C_3(|u(x)| + |u'(x)|)/4$. This yields (1.2). ☐

Appendix B. Cocycles and invariant measures

In this appendix we collect some basic facts about cocycles and the existence and uniqueness of invariant measures for group actions, which are used in Section 3 to prove Hölder continuity of the Lyapunov exponent. All this can be found in Chapter IV of [13].

Let G be a metric group, with unit e, that is both locally compact and σ-compact. Let B be a metrizable topological space such that G acts on B, i.e. to each $(g,b) \in G \times B$ one can continuously associate an element $gb \in B$ for which $(g_1g_2)\cdot b = g_1\cdot(g_2\cdot b)$ for g_1, $g_2 \in G$ and $b \in B$, and $e \cdot b = b$ for $b \in B$. A continuous map $\sigma : G \times B \to (0,\infty)$ is called a cocycle if for all $g_1, g_2 \in G$ and $b \in B$ one has $\sigma(g_1g_2, b) = \sigma(g_1, g_2b)\sigma(g_2, b)$.

Note that if σ is a cocycle, then clearly σ^t is also a cocycle for all $t \in \mathbb{R}$. There is one particular example of primary importance:

Example B.1. *Let B be a compact metric space, and G be a group acting on B. Take $\tilde{B} := B \times B \setminus \{(b,b) : b \in B\}$, and consider the induced action of G on $\tilde{B}$, i.e. $g \cdot (a,b) := (ga, gb)$ for all $g \in G$ and $(a,b) \in \tilde{B}$. If δ is the metric on B, then*

$$\sigma_1(g,(a,b)) := \frac{\delta(ga,gb)}{\delta(a,b)},$$

defines a cocycle on $G \times \tilde{B}$.

A crucial property of cocycles is that they satisfy certain integral estimates. For this reason, the "pseudo-convolution" of a probability measure μ on G and a measure ν on B is introduced by:

$$(\mu * \nu)(f) := \int f(gb) d\mu(g) d\nu(b),$$

for all $f \in \mathcal{B}(B)$, the bounded measurable functions on B. Here if B=G, then μ^n is $\mu * \mu * ... * \mu$, n-times, and the above definition coincides with that of the ordinary convolution on G. We note that a cocycle is said to be μ-integrable if $\overline{\sigma}(g) := \sup_{b \in B} \sigma(g,b)$ is μ-integrable.

The concept of invariance will also be important. A measure ν on B is said to be μ-invariant if $\mu * \nu = \nu$. For B compact, the existence of an invariant measure is trivial. In particular, any weak limit of the sequence $\frac{1}{n}\sum_{j=1}^{n} \mu^n * m$, where m is an arbitrary probability measure on B, is a μ-invariant probability measure. In addition an operator $T : \mathcal{B}(B) \to \mathcal{B}(B)$ is defined by:

$$(Tf)(b) := \int f(gb) d\mu(g).$$

Note that the operator T^n is given by the formula above with μ^n replacing μ. The relationship between these operators and μ-invariant probability measures is illustrated in Lemma $B.2$ below.

Lastly, if B is a compact space metric space, with metric δ, then for any real number α the space of α-Hölder continuous functions, $\mathcal{L}_\alpha(B)$, is defined to be

$$\mathcal{L}_\alpha(B) := \{f \in C(B) : m_\alpha(f) < \infty\},$$

where

$$m_\alpha(f) := \sup_{(a,b) \in \tilde{B}} \frac{|f(a) - f(b)|}{\delta^\alpha(a,b)}.$$

Equipped with the norm

$$\|f\|_\alpha = \|f\|_\infty + m_\alpha(f),$$

$\mathcal{L}_\alpha(B)$ is a Banach space.

Lemma B.2. *Consider the cocycle σ_1, as defined in Example $B.1$, on $G \times \tilde{B}$. Suppose*
(i) *there exists a positive number τ such that σ_1^t is μ- integrable for $|t| \leq \tau$, and*
(ii) *there exists an integer N such that*

$$\sup_{b \in \tilde{B}} \int \log[\sigma_1(g,b)] d\mu^N(g) < 0.$$

Then, there exists a real number α_0 such that for any α with $0 < \alpha \leq \alpha_0$ there exists constants $C_\alpha < \infty$ and $\rho_\alpha < 1$ for which:
(a)

$$\sup_{b \in B} \int \sigma_1(g,b)^\alpha d\mu^n(g) \leq C_\alpha \rho_\alpha^n$$

for $n = 1, 2, \ldots.$
(b) *T is a bounded operator on $\mathcal{L}_\alpha$ satisfying*

$$\|T^n f - \nu(f)\|_\alpha \leq \|f\|_\alpha C_\alpha \rho_\alpha^n$$

for $n = 1, 2, \ldots$ and $f \in \mathcal{L}_\alpha$, where ν is any μ-invariant probability measure. In particular, this proves uniqueness of the invariant measure.
(c) *The operator T on $\mathcal{L}_\alpha$ has eigenvalue 1 and the rest of the spectrum is contained in a disk of radius strictly less than 1. Moreover, T admits the following decomposition:*

$$T^n f = \nu(f) + Q^n f \quad \text{for } f \in \mathcal{L}_\alpha,$$

where ν is the invariant probability measure and Q is an operator on $\mathcal{L}_\alpha$ of spectral radius strictly less than 1.

Proof. See Proposition IV.3.5, Proposition IV.3.15, and Corollary IV.3.16 of [13]. □

References

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

[2] M. Aizenman, J. Schenker, R. Friedrich, and D. Hundertmark, Finite-volume Criteria for Anderson localization. Preprint 1999, math-ph/9910022v2, to appear in *Commun. Math. Phys.*

[3] J.Avron and B. Simon, Almost periodic Schrödinger operators, II. The integrated density of states. *Duke Math. J.* **50**, (1983), 369–391

[4] J. Barbaroux, W. Fischer, and P. Müller, Dynamical properties of random Schrödinger operators. Preprint 1999, math-ph/9907002

[5] M. Benderskii and L. Pastur, On the asymptotics of the solutions of second-order equations with random coefficients (in Russian). *Teoria Funkcii, Func. Anal. i Priloz.* (Kharkov University) **N 22** (1975), 3–14

[6] S. De Bièvre and F. Germinet, Dynamical Localization for the Random Dimer Schrödinger Operator. *J. Stat. Phys.* **98** (2000), 1135–1148

[7] G. Borg, Uniqueness theorems in the spectral theory of $y'' + (\lambda - q(x))y = 0$. *Proc. 11th Scandinavian Congress of Mathematicians*, Johan Grundt Tanums Forlag, Oslo (1952) 276–287

[8] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators.* Birkhäuser, Boston–Stuttgart (1985)

[9] M. Campanino and A. Klein, A supersymmetric transfer matrix and differentiability of the density of states in the one-dimensional Anderson model. *Commun. Math. Phys.* **104** (1986), 227–241

[10] R. Carmona, One-dimensional Schrödinger operators with random potential. *Physica A* **124** (1984), 181–188

[11] R. Carmona, Random Schrödinger Operators. *École d'été de probabilités de Saint-Flour*, XIV—1984, 1–124, Lecture Notes in Math. **1180**, Springer, Berlin-New York, 1986

[12] R. Carmona, A. Klein, and F. Martinelli, Anderson localization for Bernoulli and other singular potentials. *Commun. Math. Phys.* **108** (1987), 41–66

[13] R. Carmona and J. Lacroix, *Spectral theory of random Schrödinger operators.* Birkhäuser, Basel–Berlin (1990)

[14] E.A. Coddington and N. Levinson: Theory of ordinary differential equations. McGraw-Hill, New York 1955

[15] J. M. Combes and P. D. Hislop: Localization for some continuous random Hamiltonians in d-dimensions. *J. Funct. Anal.* **124** (1994), 149–180

[16] D. Damanik and P. Stollmann, Multi-scale analysis implies strong dynamical localization. Preprint 1999, mp-arc/99-461, to appear in *Geom. Funct. Anal.*

[17] P. K. Datta and K. Kundu, The absence of localization in one-dimensional disordered harmonic chains. *J. Phys. Condens. Matter* **6** (1994), 4465–4478.

[18] R. del Rio, S. Jitomirskaya, Y. Last, and B. Simon, Operators with singular continuous spectrum, IV. Hausdorf dimension, rank-one perturbations, and localization. *J. Anal. Math.* **69** (1996), 153–200

[19] H. von Dreifus and A. Klein, A new proof of localization in the Anderson tight binding Model. *Commun. Math. Phys.* **124** (1989), 285–299

[20] M.S.P. Eastham: The Spectral Theory of Periodic Differential Equations. Scottish Academic Press, Edinburgh–London 1973

[21] A. Figotin, On the exponential growth of solutions of finite difference equations with random coefficients (in Russian). *Dokl. Akad. Nauk Uzb. SSR* 1980(2), 9–11

[22] A. Figotin and A. Klein, Localization of classical waves. I. Acoustic waves. *Comm. Math. Phys.* **180** (1996), 439–482

[23] F. Germinet and S. De Bièvre, Dynamical localization for discrete and continuous random Schrödinger operators. *Commun. Math. Phys.* **194** (1998), 323–341

[24] F. Gesztesy and B. Simon, On local Borg-Marchenko uniqueness results. *Commun. Math. Phys.* **211** (2000) 273–287

[25] K. Ishii, Localization of eigenstates and transport phenomena in one-dimensional disordered systems. *Progress Theor. Phys. Suppl.* **53** (1973), 77–118

[26] K. Ishii and H. Matsuda, Localization of normal modes and energy transport in the disordered harmonic chain. *Progress Theor. Phys. Suppl.* **45** (1970), 56–86

[27] T. Kato, *Perturbation theory for linear operators.* Springer-Verlag, Berlin–Heidelberg (1966)

[28] W. Kirsch, On a class of random Schrödinger operators. *Adv. in Appl. Math.* **6** (1985), 177–187

[29] W. Kirsch, S. Kotani and B. Simon: Absence of absolutely continuous spectrum for some one dimensional random but deterministic potentials. *Ann. Inst. Henri Poincaré* **42**, 383 – 406 (1985)

[30] W. Kirsch, P. Stollmann, and G. Stolz, Anderson localization for random Schrödinger operators with long range interactions. *Commun. Math. Phys.* **195** (1998), 495–507

[31] F. Klopp, Localization for some continuous random Schrödinger operators. *Commun. Math. Phys.* **167** (1995), 553–569

[32] V. Kostrykin and R. Schrader, Scattering theory approach to random Schrödinger operators in one-dimension. *Rev. Math. Phys.* **11** (1999), 187–242

[33] V. Kostrykin and R. Schrader, Global bounds for the Lyapunov exponent and the integrated density of states or random Schrödinger operators in one dimension. Preprint 2000, mp-arc/00-226

[34] S. Kotani, One-dimensional Schrödinger operators and Herglotz functions. In: *Probabilistic Methods in Mathematical Physics*, ed. K. Ito and N. Ikeda, 219–250, Academic Press, Boston 1987

[35] I. Lifshits, S. Gredeskul, and L. Pastur, Introduction to the theory of disordered systems (Translated from the Russian by Eugene Yankovsky). *A Wiley-Interscience Publication.* John Wiley & Sons, Inc., New York, 1988

[36] V. Marchenko, Certain problems in the theory of second-order differential operators. *Doklady Ahad. Nauk SSSR* **72**, 457 –460 (1950) (Russian)

[37] V. Marchenko, Some questions in the theory of one-dimensional linear differential operators of the second order. I, *Trudy Moskov. Mat. Obsc.* **1**, 327–420 (1952) (Russian); English transl. in *Amer. Math. Soc. Transl. (2)* **101**, 1–104 (1973)

[38] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators.* Springer Verlag, Berlin-Heidelberg-New York (1992)

[39] C. Shubin, R. Vakilian, and T. Wolff, Some harmonic analysis questions suggested by Anderson-Bernoulli models. *Geom. Funct. Anal.* **8** (1998), 932–964

[40] B. Simon, A new approach to inverse spectral theory, I. Fundamental formalism. *Ann. of Math. (2)* **150** (1999) 1029–1057

[41] B. Simon and M. Taylor, Harmonic analysis on $SL(2, \mathbb{R})$ and smoothness of the density of states in the one-dimensional Anderson model. *Commun. Math. Phys.* **101** (1985), 1–19

[42] R. Sims and G. Stolz, Localization in one dimensional random media: a scattering theoretic approach. Preprint 2000, mp-arc/00-64, to appear in *Commun. Math. Phys.*

[43] P. Stollmann, Wegner estimates and localization for continuum Anderson models with some singular distributions. Preprint 1998, to appear in *Arch. Math.*

[44] P. Stollmann, *Caught by Disorder: Lectures on Bound States in Random Media*. Book, in preparation.

[45] G. Stolz, Bounded solutions and absolute continuity of Sturm-Liouville operators. *J. Math. Anal. Appl.* **169** (1992), 210–228

[46] G. Stolz, Non-monotonic Random Schrödinger Operators: The Anderson Model. *J. Math. Anal. Appl.* **248** (2000), 173–183

[47] W. Walter, *Ordinary Differential Equations*. Graduate Texts in Mathematics, Vol. **182**, Springer, New York (1998)

Department of Mathematics, University of California at Irvine, CA 92697-3875, USA, partially supported by the German Academic Exchange Service through HSP III (Postdoktoranden)

Department of Mathematics, University of Alabama at Birmingham, AL 35394-1170, USA, partially supported by NSF grant DMS-9706076

Department of Mathematics, University of Alabama at Birmingham, AL 35394-1170, USA, partially supported by NSF grants DMS-9706076 and DMS-0070343