ESI The Erwin Schrödinger International Institute for Mathematical Physics
Boltzmanngasse 9 A-1090 Wien, Austria

Hölder Continuity of the Integrated Density of States for Quasiperiodic Schrödinger Equations and Averages of Shifts of Subharmonic Functions

Michael Goldstein
Wilhelm Schlag

Vienna, Preprint ESI 845 (2000) February 23, 2000

Supported by Federal Ministry of Science and Transport, Austria
Available via http://www.esi.ac.at

HÖLDER CONTINUITY OF THE INTEGRATED DENSITY OF STATES FOR QUASIPERIODIC SCHRÖDINGER EQUATIONS AND AVERAGES OF SHIFTS OF SUBHARMONIC FUNCTIONS

MICHAEL GOLDSTEIN AND WILHELM SCHLAG

Abstract. In this paper we consider various regularity results for discrete quasiperiodic Schrödinger equations

$$-\psi_{n+1} - \psi_{n-1} + V(\theta + n\omega)\psi_n = E\psi_n$$

with analytic potential V. We prove that on intervals of positivity for the Lyapunov exponent the integrated density of states is Hölder continuous in the energy provided ω has a typical continued fraction expansion. The proof is based on certain sharp large deviation theorems for the norms of the monodromy matrices and the "avalanche–principle". The latter refers to a mechanism that allows us to write the norm of a monodromy matrix as the product of the norms of many short blocks. In the multifrequency case the integrated density of states is shown to have a modulus of continuity of the form $\exp(-|\log t|^\sigma)$ for some $0 < \sigma < 1$, but currently we do not obtain Hölder continuity in the case of more than one frequency. We also present a mechanism for proving the positivity of the Lyapunov exponent for large disorders for a general class of equations. The only requirement for this approach is some weak form of a large deviation theorem for the Lyapunov exponents. In particular, we obtain an independent proof of the Herman–Sorets–Spencer theorem in the multifrequency case. The approach in this paper is related to the recent nonperturbative proof of Anderson localization in the quasiperiodic case by J. Bourgain and M. Goldstein.

1. Introduction

Given a real–valued function $V : \mathbb{T}^d \to \mathbb{T}^d$, an ergodic shift $\boldsymbol{\theta} \mapsto \boldsymbol{\theta} + \boldsymbol{\omega}$ on $\mathbb{T}^d$, and a real number μ consider the following family of discrete Schrödinger operators

$$(H_{\boldsymbol{\omega},\mu,\boldsymbol{\theta}}\, \psi)(n) = -\psi(n+1) - \psi(n-1) + \mu V(\boldsymbol{\theta} + n\boldsymbol{\omega})\psi(n) = E\psi(n), \quad n \in \mathbb{Z} \tag{1.1}$$

on $\ell^2(\mathbb{Z})$. It is a well–known consequence of ergodicity that the spectra of this family of self–adjoint operators are deterministic, i.e., there exists a fixed compact set $K \subset \mathbb{R}$ so that $\mathrm{spec}(H_{\boldsymbol{\omega},\mu,\boldsymbol{\theta}}) = K$ for a.e. $\boldsymbol{\theta} \in \mathbb{T}^d$. Moreover, the spectral parts are also deterministic, seeFigotin, Pastur [12]. It was shown by Shnol and Simon that a.e. energy E with respect to the spectral measure has polynomially bounded solutions of equation (1.1), see [30] and [31]. These generalized eigenfunctions exhibit different behavior in different domains of the (μ, E)–plane. This phenomenon was studied by physicists starting with the famous works by Anderson [1] and Harper [20]. In physical terminology the quantum system ruled by (1.1) experiences phase transitions in the plane of the main parameters μ (coupling constant) and E (energy). The rigorous analysis of this phenomenon was initiated by Sinai's Moscow seminar about thirty years ago, see Oseledec [28], Dinaburg, Sinai [9] and Goldsheid, Molchanov, Pastur [18]. Equations with random potentials have a particularly rich history with important contributions by many researchers. A list of basic references up to roughly 1991 can be found in the monographs Cycon, Fröse, Kirsch, Simon[8], Carmona, Lacroix [6], [12].

The first author was supported in part by Ben–Gurion University and a grant of the Israel Academy of Sciences. This paper was completed while he was a member of the Institute for Advanced Study, Princeton, where he was supported by a grant of the NEC Research Institute, Inc. The second author was supported in part by the NSF, grant number DMS-9706889 and by the Erwin Schrödinger Institute in Vienna, Austria. The authors wish to thank Alexander Brudnyi and Thomas Spencer for useful discussions.

The following notions are essential in the study of equation (1.1). For further details we refer the reader to [12].

(I) The Lyapunov Exponent. Rewriting (1.1) as a system of first order difference equations yields

$$\begin{pmatrix} \psi(n+1) \\ \psi(n) \end{pmatrix} = A(\boldsymbol{\theta}+n\boldsymbol{\omega},E)\begin{pmatrix} \psi(n) \\ \psi(n-1) \end{pmatrix}, \tag{1.2}$$

where

$$A(\boldsymbol{\theta},E) = \begin{pmatrix} \mu V(\boldsymbol{\theta})-E & -1 \\ 1 & 0 \end{pmatrix}. \tag{1.3}$$

By Kingman's subadditive ergodic theorem the limit

$$L(\mu,E) = \lim_{n\to\infty} n^{-1}\log\|A(\boldsymbol{\theta}+n\boldsymbol{\omega},E)\dots A(\boldsymbol{\theta}+\boldsymbol{\omega},E)\| = \lim_{n\to\infty}\frac{1}{n}\int_{\mathbb{T}^d}\log\|A(\boldsymbol{\eta}+n\boldsymbol{\omega},E)\dots A(\boldsymbol{\eta}+\boldsymbol{\omega},E)\|\,d\boldsymbol{\eta}$$

exists for a.e. $\boldsymbol{\theta}$. $L(\mu,E)$ is called the Lyapunov exponent. Since the matrix A in (1.3) is unimodular, one clearly has $L(\mu,E)\ge 0$. $M_n(\boldsymbol{\theta},E) = A(\boldsymbol{\theta}+n\boldsymbol{\omega},E)\dots A(\boldsymbol{\theta}+\boldsymbol{\omega},E)$ is referred to as the monodromy matrix associated with (1.1). As the propagator of that equation on the interval $[0,n]$ it is of fundamental importance in its study.

(II) The Integrated Density of States. Let $E_{\Lambda,j}(\mu,\theta)$, $j=1,\dots,b-a+1=|\Lambda|$ be the eigenvalues of the restriction of (1.1) to the interval $\Lambda=[a,b]$ with zero boundary conditions, $\varphi(a-1)=\varphi(b+1)=0$. Consider

$$N_\Lambda(\mu,E,\theta) = \frac{1}{|\Lambda|}\sum_j \chi_{(-\infty,E)}(E_{\Lambda,j}).$$

It is well–know that the weak limit (in the sense of measures)

$$\lim_{a\to-\infty,b\to+\infty} dN_\Lambda(\mu,\cdot,\theta) = dN(\mu,\cdot)$$

exists and does not depend on θ (up to a set of measure zero). The distribution function $N(\mu,\cdot)$ is called the integrated density of states. It is connected with the Lyapunov exponent via the Thouless formula

$$L(\mu,E) = \int \log|E-E'|\,dN(\mu,E'). \tag{1.4}$$

Assuming that $V(\theta)$ possesses a certain degree of regularity and that ω is a generic irrational number, the main conjecture about equation (1.1) is as follows:

(A) If $L(\mu_0,E_0)>0$, then there is some $\delta>0$ such that almost every θ satisfies the following property: For every $E\in(E_0-\delta,E_0+\delta)$ any generalized eigenfunction of (1.1) with that choice of E decays exponentially. This is equivalent to the following property:

(AL) The spectrum of (1.1) in $(E_0-\delta,E_0+\delta)$ is pure point and the corresponding eigenfunctions decay exponentially.

Property (AL) is called *Anderson localization.* Deciding in which cases Anderson localization holds has been at the center of research in this area. Consider the equation

$$-\Delta\psi(n)+\mu V_n\psi_n \quad = \quad E\psi_n \tag{1.5}$$

where $n\in\mathbb{Z}^d$ and $\{V_n\}$ is a random field on $\mathbb{Z}^d$. Equation (1.5) includes all relevant models, for example the quasiperiodic case (1.1) and the case of independent identically distributed random potentials (the latter is the classical "Anderson model"). The basic ideas in the analysis of AL were introduced in the following works:

- Goldsheid, Molchanov, Pastur [18]: Reduction of AL to Fürstenberg's theorem [17] on the positivity of the Lyapunov exponent.
- Fröhlich, Spencer [14]: A probabilistic KAM scheme for multi–dimensional Anderson model with large μ.

- Dinaburg, Sinai [9], Sinai [33], Fröhlich, Spencer, Wittwer [16]: KAM approach for quasiperiodic equations.

These techniques have been developed in a number of important publications, see the references in [6], and [12] for the literature up to roughly 1991. In the years since then a simple proof of AL for the Anderson model with large μ was given by Aizenman and Molchanov [2], the complete analysis of Floquet–Bloch solutions in the quasiperiodic case was obtained by Eliasson [11], the purely singular continuous nature of the spectrum for the almost Mathieu equation with $\mu = 2$ was established in the works by Avron, Gordon, Jitomirskaya, Last, van Mouche, Simon, Thouless [36], [3], [25], [26], [19], a nonperturbative proof of AL for the almost Mathieu equation with $\mu > 2$ was given by Jitmoriskaya [22].

In the recent preprint by Bourgain and Goldstein [4] Anderson localization for equation (1.1) was established in a nonperturbative regime for $d = 1, 2$ provided V is analytic. It seems reasonable to believe that the methods from [4] will lead to the solution of problem (A) for equation (1.1) with analytic potentials V.
To provide a more complete picture of the phase transitions in this models one needs to answer the following questions:

(i) How regular are the main thermodynamical functions $L(\mu, E)$ and $dN(\mu, E)$ with respect to μ and E? Is $dN(\mu, E)$ analytic in μ in some regions?

(ii) What are the typical spacings between the eigenvalues of (1.1) in a large interval $\Lambda = [a, b]$? What are the typical localization lengths of the corresponding eigenfunctions? What is the connection between these quantities? How are these quantities related to the Lyapunov exponent?
In this paper we study the regularity properties of $L(\mu, E)$ and $dN(\mu, E)$. This problem is considered difficult for any type of sequence of potentials, see [8]. Positive results are known only for independent random potentials $V(n)$ under certain assumptions, cf. Constantinescu, Fröhlich, Spencer [7], Wegner [37], Simon, Taylor [32], and Campanino, Klein [5]. We would like to emphasize, however, that our approach is completely different from these works. Although our methods also allow us to establish Hölder regularity of L in μ without significant changes, we have restricted ourselves to E. We plan to return to the issue of (possibly much greater) regularity in μ elsewhere.
Our method hinges on two basic tools. The first of these tools is the so called *avalanche principle*. This principle basically allows one to write the norm of the monodromy matrix on $[0, n]$ as the product of the norms of shifts of the monodromy matrix on $[0, \ell]$ provided the norms of all the monodromy matrices of size ℓ are large compared with n, see Section 2. It applies to any number of frequencies, i.e., $d = 1, 2, \ldots$, and provides the rescaling procedure in non-perturbative regimes.
The second basic tool are certain *sharp large deviation theorems* for the Lyapunov exponents. More precisely, we prove estimates of the form (setting $\mu = 1$ and $L(E) = L(1, E)$ for simplicity)

$$\operatorname{mes}\Big(\Big\{\theta \in \mathbb{T} : \Big|\frac{1}{n}\log\|M_n(\theta, E)\| - L(E)\Big| > \delta\Big\}\Big) \le \exp(-c\delta n). \tag{1.6}$$

These estimates are of crucial importance in our approach since they provide the aforementioned largeness hypothesis in the avalanche principle. More precisely, applying (1.6) to M_ℓ with $\ell = C\log n$ shows that the avalanche principle can be applied with this choice of ℓ up to an exceptional set of θ's of measure no larger than n^{-10}, say. This is essential, since the derivative of M_ℓ in the energy is only polynomially large in n rather than exponentially large, as for M_n itself. This is the key observation that allows us to prove Hölder continuity of $L(E)$ in E. The corresponding result for N then follows easily from (1.4) by well–known arguments. See Section 6 for details. We would like to emphasize that (1.6) has so far been established only for the case of one frequency. For the case of several frequencies (1.6) is known to hold with $\exp(-Cn^\sigma)$ on the right–hand side for some $\sigma \in (0, 1)$. This fact accounts for the weaker regularity results for $d = 2, 3 \ldots$ given below, see Section 10. In connection with (1.6) we would like to mention that the appearance of $\theta \in \mathbb{T}$ for which the deviations $\frac{1}{n}\log\|M_n(\theta, E)\| - L(E)$ are large is intimately connected with the essential support of eigenfunctions that was discovered in [33] and [16] for perturbative regimes. Moreover, the exponential

decay of the measure of this set reflects the exponential growth of the gaps between the points in the essential support. However, we do not exploit these facts here but plan to elaborate them elsewhere.
All proof methods of (1.6) known to the authors rely on the fact that for analytic potentials (and fixed E)

$$u_n(z_1, \ldots, z_d) \quad = \quad \frac{1}{n} \log \|M_n(z_1, \ldots, z_d, E)\| \tag{1.7}$$

is a plurisubharmonic bounded function on a neighborhood of the origin. The importance of subharmonicity or plurisubharmonicity was already recognized by Herman in his seminal paper [21]. As in [4] subharmonicity is exploited be means of Riesz's representation. More precisely, any subharmonic function u on the unit disk D can be written in the form

$$u(z) = \int \log|z - \zeta|\, d\mu(\zeta) + h(z) \quad \text{for all } z \in D. \tag{1.8}$$

Here h is harmonic and μ is a nonnegative measure that is finite on compact subsets of D. One easily checks from the definition (1.7) that

$$\sup_{\boldsymbol{\theta} \in \mathbb{T}^d} |u_n(\boldsymbol{\theta} + \boldsymbol{\omega}) - u_n(\boldsymbol{\theta})| < \frac{C}{n}.$$

This allows one to reduce the proof of (1.6) to similar estimates for averages of the form

$$K^{-1} \sum_{1 \le m \le K} u(\theta + m\omega), \tag{1.9}$$

where u is a bounded subharmonic function. Furthermore, in the case of one frequency, i.e., $d = 1$, property (1.8) allows one to reduce the large deviation estimates for these averages to the case $u(z) = \log|z|$. This is precisely the approach developed in this paper, see Section 3. Averages as in (1.9) appear already in [4]. However, the methods from [4], which are based on Fourier expansions, seem to be insufficient for the somewhat delicate regularity questions that we address in this paper.

In the case of several frequencies this straightforward reduction to averages of shifts of $\log|z|$ is not available. We therefore develop a different approach based on Cartan type estimates for plurisubharmonic functions, see Section 8. Given any bounded plurisubharmonic function $u(z_1, \ldots, z_d)$ and any $r > 0$, there exists a polydisk Π of size r and a set $\mathcal{B} \subset \Pi$ such that on $\Pi \setminus \mathcal{B}$ the deviation of u from its average is smaller than r^β for some $\beta > 0$. Moreover, $\mathrm{mes}(\mathcal{B}) < \exp(-r^{-\beta})$. Here $\beta > 0$ is some small absolute constant. Compare Theorems 8.3 and 8.5 below for more precise statements. This result was motivated by Cartan's theorem for subharmonic functions of one variable, see Levin [27]. Combining this statement with the dynamics then allows us to control the deviation of u_n as in (1.7) on the entire torus, cf. Section 9. We would like to emphasize that the approach based on Cartan type estimates is not limited to the dynamics of the shift, but applies to a much larger class of transformations.
We conclude this paper with Section 11 that is devoted to a proof of the positivity of the Lyapunov exponent for large disorders for equations of the form

$$-\psi_{n+1} - \psi_{n-1} + \lambda V(T^n \boldsymbol{\theta})\psi_n = E\psi_n.$$

Here $\boldsymbol{\theta} \in \mathbb{T}^d$, $T : \mathbb{T}^d \to \mathbb{T}^d$ is an ergodic transformation, and V a nonconstant real–analytic function on $\mathbb{T}^d$. We show that the Lyapunov exponents are positive for large λ *provided the following large deviation theorem holds*: For some $\sigma > 0$ and all n

$$\int_{\mathbb{T}^d} \left| \frac{1}{n} \log \|M_n(\boldsymbol{\theta}, \lambda, E)\| - L_n(\lambda, E) \right| d\boldsymbol{\theta} \le C\, S(\lambda, E) n^{-\sigma}, \tag{1.10}$$

where $S(\lambda, E)$ is some scaling factor. In particular, we obtain an independent proof of the Herman, Sorets–Spencer Theorem [21], [34] in the multifrequency case. We hope that this method will lead to positivity of the Lyapunov exponent for many interesting examples, although (1.10) is unknown in most cases and also seems rather difficult to establish.

2. Avalanche principle

Definition 2.1. Fix some unimodular 2×2 matrix K. We denote the normalized eigenvectors of $\sqrt{K^*K}$ by $\mathbf{u}_K^+$ and $\mathbf{u}_K^-$, respectively. One has $K\mathbf{u}_K^+ = \|K\|\mathbf{v}_K^+$ and $K\mathbf{u}_K^- = \|K\|^{-1}\mathbf{v}_K^-$ where $\mathbf{v}_K^+$ and $\mathbf{v}_K^-$ are unit vectors. Given two unimodular 2×2 matrices K and M, we let $b^{(+,+)}(K,M) = \mathbf{v}_K^+ \cdot \mathbf{u}_M^+$ and similarly for $b^{(+,-)}, b^{(-,+)}$, and $b^{(-,-)}$. Strictly speaking, these quantities are defined up to a sign, but we are only interested in the absolute value.

The letters C and c will denote absolute constants. Any dependence on parameters will usually be stated. As usual, $a \asymp b$ will mean $C^{-1}a \le b \le Ca$ for some C.

Proposition 2.2. *Let $A_1, \ldots, A_n$ be a sequence of unimodular 2×2–matrices. Suppose that*

$$\min_{1 \le j \le n} \|A_j\| \ge \mu > n \quad \text{and} \tag{2.1}$$

$$\max_{1 \le j < n} [\log \|A_{j+1}\| + \log \|A_j\| - \log \|A_{j+1}A_j\|] < \frac{1}{2}\log \mu. \tag{2.2}$$

Then

$$\left|\log \|A_n \cdot \ldots \cdot A_1\| - \sum_{j=1}^{n} \log \|A_j\| - \sum_{j=1}^{n-1} \log |b^{(+,+)}(A_j, A_{j+1})|\right| \ < \ C\frac{n}{\mu} \tag{2.3}$$

$$\left|\log \|A_n \cdot \ldots \cdot A_1\| + \sum_{j=2}^{n-1} \log \|A_j\| - \sum_{j=1}^{n-1} \log \|A_{j+1}A_j\|\right| \ < \ C\frac{n}{\mu}. \tag{2.4}$$

Proof. One checks from the definition that

$$\|K\|\|M\||b^{(+,+)}(K,M)| - \|K\|\|M\|^{-1} \le \|MK\| \le \|K\|\|M\||b^{(+,+)}(K,M)| + \|K\|^{-1}\|M\| + \|K\|\|M\|^{-1}.$$

In particular,

$$\frac{\|A_{j+1}A_j\|}{\|A_{j+1}\|\|A_j\|} - \frac{1}{\|A_j\|^2} \le |b^{(+,+)}(A_j, A_{j+1})| \le \frac{\|A_{j+1}A_j\|}{\|A_{j+1}\|\|A_j\|} + \frac{1}{\|A_j\|^2} + \frac{1}{\|A_{j+1}\|^2}.$$

In view of our assumptions therefore

$$1 - \frac{\sqrt{\mu}}{\mu^2} \quad \le \quad |b^{(+,+)}(A_j, A_{j+1})|\frac{\|A_{j+1}\|\|A_j\|}{\|A_{j+1}A_j\|} \le 1 + \frac{2\sqrt{\mu}}{\mu^2} \tag{2.5}$$

which implies $|b^{(+,+)}(A_j, A_{j+1})| \ge \frac{1}{\sqrt{\mu}}(1 - \mu^{-\frac{3}{2}}) \ge \frac{1}{2}\mu^{-\frac{1}{2}}$ if $n \ge 2$, say. One checks easily by induction that for any vector $\boldsymbol{u}$

$$A_n \cdot \ldots \cdot A_1 \boldsymbol{u} = \sum_{\epsilon_1, \ldots, \epsilon_n = \pm 1} \|A_n\|^{\epsilon_n} \prod_{j=1}^{n-1} \|A_j\|^{\epsilon_j} b^{(\epsilon_j, \epsilon_{j+1})}(A_j, A_{j+1}) (\boldsymbol{u}_{A_1}^{\epsilon_1} \cdot \boldsymbol{u})\, \boldsymbol{v}_{A_n}^{\epsilon_n}.$$

Hence

$$\|A_n \cdot \ldots \cdot A_1 \boldsymbol{u}\| = \|A_n\| \prod_{j=1}^{n-1} \|A_j\||b^{(+,+)}(A_j, A_{j+1})||\boldsymbol{u}_{A_1}^+ \cdot \boldsymbol{u}|[1 + R_n(\boldsymbol{u})]$$

where

$$
\begin{aligned}
|R_n(\boldsymbol{u})| \quad &\le \quad \sum_{\substack{\epsilon_1,\ldots,\epsilon_n=\pm1\\ \min_j \epsilon_j=-1}} \prod_{j=1}^{n} \|A_j\|^{\epsilon_j-1} \prod_{k=1}^{n-1} \left|\frac{b^{(\epsilon_k,\epsilon_{k+1})}(A_k,A_{k+1})}{b^{(+,+)}(A_k,A_{k+1})}\right| \\
&\le \quad \sum_{\ell=1}^{n} \binom{n}{\ell} \mu^{-2\ell}(2\sqrt{\mu})^{2\ell} = \sum_{\ell=1}^{n}\binom{n}{\ell}(4/\mu)^{\ell} = \Big(1+\frac{4}{\mu}\Big)^n - 1 < 4e^4\,\frac{n}{\mu}
\end{aligned}
$$

and (2.3) follows. In view of (2.5)

$$
\left|\sum_{j=1}^{n-1}\Big[\log|b^{(+,+)}(A_j,A_{j+1})| - \log\|A_{j+1}A_j\| + \log\|A_j\| + \log\|A_{j+1}\|\Big]\right| \le C\mu^{-\frac{3}{2}}n \le C\frac{n}{\mu}.
$$

Combining this with (2.3) yields (2.4). □

3. Large deviation theorem for sums of shifts of normalized 1-periodic subharmonic functions

In this section we consider subharmonic functions $u(z)$ defined on some neighborhood of the real axis satisfying $u(z) = u(z+1)$. Furthermore, we require $|u(z)| \le 1$ on that neighborhood. Recall Riesz's theorem, see Levin [27], Lecture 7: Given any subharmonic function on a domain G, there are a unique positive measure μ and a unique harmonic function h defined on G such that

$$
u(z) = \int_G \log|z-\zeta|\,d\mu(\zeta) + h(z) \quad \text{for all } z \in G.
$$

Furthermore, for any compact $K \subset G$ there is a constant $C(K,G)$ so that

$$
\mu(K) + \sup_{z\in K}|h(z)| \quad \le \quad C(K,G)\|u\|_\infty. \tag{3.1}
$$

This follows easily from Jensen's formula, see [27], and an explicit representation of h as boundary integral. See also Koosis [24] or the proof of Lemma 8.2 below.

In particular, with a periodic subharmonic function u as above, there is some positive measure μ and a harmonic function h both of which are defined on a neighborhood of the interval $[0,1]$ such that for all $0 \le x \le 1$

$$
u(x) = \int \log|x-\zeta|\,d\mu(\zeta) + h(x). \tag{3.2}
$$

Moreover, $\|\mu\| + \|h\|_\infty \le C$. The appearance of the logarithm in (3.2) should explain the following lemmas. For the relevance of subharmonicity for the Schrödinger equation see the beginning of the following section.

Fix some $a > 1$. Throughout this paper we assume that $\omega \in (0,1)$ satisfies the Diophantine condition

$$
\|n\omega\| \ge \frac{C_\omega}{n(\log n)^a} \quad \text{for all } n. \tag{3.3}
$$

It is well–known that for a fixed $a > 1$ a.e. ω satisfies (3.3). Consider the continued fraction expansion $\omega = [a_1, a_2, \ldots]$ with convergents $\frac{p_s}{q_s}$ for $s = 1, 2, \ldots$. One has $\|q_s\omega\| \le q_{s+1}^{-1}$ and in view of condition (3.3) therefore

$$
q_{s+1} \le C q_s (\log q_s)^a. \tag{3.4}
$$

One checks by induction that this implies

$$
q_s \le \exp(2as\log s) \tag{3.5}
$$

for sufficiently large $s > s_0(a)$.

Let $\{x\} = x - [x]$. For any positive integer q, complex number $\zeta = \xi + i\eta$, and $0 < x < 1$ define

$$f_{q,\zeta}(x) = \sum_{0\le k<q} \log|\{x - k/q\} - \zeta|,\quad F_{q,\zeta}(x) = \sum_{0\le k<q} \log|\{x - k\omega\} - \zeta|,\quad I(\zeta) = \int_0^1 \log|y-\zeta|\,dy. \tag{3.6}$$

We will always assume that $-1 < \xi < 2$ and $|\eta| \le 1$. In what follows dist will denote the distance mod 1, i.e.,

$$\mathrm{dist}(x, y) = \min_{n\in\mathbb{Z}} |x - y + n|.$$

This is the same as $\mathrm{dist}(x, y) = \|x - y\|$, where $\|\cdot\|$ denotes the distance to the nearest integer.

Lemma 3.1. *Let $d(x, q) = \mathrm{dist}(x, \{k/q : 0 \le k < q\})$ and $D(x, \omega, q) = \mathrm{dist}(x, \{m\omega, k/q : 0 \le k, m < q\})$. Then for all $0 \le x < 1$*

$$\begin{aligned} |f_{q,\zeta}(x) - qI(\zeta)| &\le C(|\log d(x-\xi, q)| + \log q)\\ |F_{q_s,\zeta}(x) - f_{q_s,\zeta}(x)| &\le C(|\log D(x-\xi, \omega, q_s)| + \log q_s). \end{aligned}$$

Proof. Let $g_\zeta(y) = \log|\{y\} - \zeta|$. Clearly, g_ζ has at most two monotonicity intervals on $[0, 1]$. Arranged in increasing order, the points $\{x - \frac{k}{q}\}$ for $k = 0, 1, \dots, q-1$ form an arithmetic progression with increment $\frac{1}{q}$. An elementary consideration involving Riemann integrals therefore implies that

$$\Big|\frac{1}{q}\sum_{k=0}^{q-1} g_\zeta\Big(x - \frac{k}{q}\Big) - I(\zeta)\Big| \le \frac{C}{q}\max_{0\le k<q}|g_\zeta\Big(x-\frac{k}{q}\Big)| - \int_{|y|<\frac{C}{q}} \log|y|\,dy \le \frac{C}{q}(|\log d(x-\xi, q)| + \log q)$$

and the first assertions holds. To obtain the second assertion, arrange the points $\{x - \frac{k}{q_s}\} - \zeta$ in increasing order on the line $\Im = -\eta$. The distance between any two adjacent points is exactly $1/q_s$ and for each of them there is a point of the form $\{x - k\omega\} - \zeta$ at a distance less than $1/q_{s+1}$. Fix any $x \in [0, 1]$ and let k_0 be that value of k for which $|\{x - kp_s/q_s\} - \zeta| + |\{x - k\omega\} - \zeta|$ is minimal. Then

$$\begin{aligned} |F_{q_s,\zeta}(x) - f_{q_s,\zeta}(x)| &\le \sum_{|k-k_0|<3}\Big[\big|\log|\{x - kp_s/q_s\} - \zeta|\big| + \big|\log|\{x - k\omega\} - \zeta|\big|\Big] + \\ &\quad + \sum_{|k-k_0|\ge 3}\Big|\log\Big[1 + \frac{|\{x - kp_s/q_s\} - \{x - k\omega\}|}{|\{x - kp_s/q_s\} - \zeta|}\Big]\Big| \\ &\le C|\log D(x-\xi, \omega, q_s)| + C\sum_{\ell=1}^{q_s}\frac{q_{s+1}^{-1}}{\ell q_s^{-1}} \\ &\le C|\log D(x-\xi, \omega, q_s)| + C\frac{q_s\log q_s}{q_{s+1}}, \end{aligned}$$

as claimed. □

Lemma 3.2. *Let $\Omega \subset \mathbb{T}$ be an arbitrary finite set. Then*

$$\int_{\mathbb{T}} \exp(\lambda|\log\mathrm{dist}(x, \Omega)|)\,dx \le \frac{2^\lambda}{1-\lambda}(\#\Omega)^\lambda$$

for any $0 < \lambda < 1$.

Proof. Let $\Omega = \{y_1, \dots, y_m\}$ and set $\{x : \mathrm{dist}(x, \Omega) = \|x - y_j\|\} = I_j$ for each $j = 1, \dots, m$. The intervals I_j intersect at most at the endpoints. Thus

$$\begin{aligned} &\int_{\mathbb{T}} \exp(\lambda|\log\mathrm{dist}(x, \Omega)|)\,dx = \sum_{j=1}^m \int_{I_j} \exp(-\lambda\log\|x - y_j\|)\,dx \le \\ &\le \frac{2^\lambda}{1-\lambda}\sum_{j=1}^m |I_j|^{1-\lambda} \le \frac{2^\lambda}{1-\lambda}\Big(\sum_j |I_j|\Big)^{1-\lambda} m^\lambda \le \frac{2^\lambda}{1-\lambda} m^\lambda \end{aligned}$$

as claimed. ☐

Proposition 3.3. *Let q_s and ζ be as above. Then for sufficiently small $\lambda > 0$*

$$\int_0^1 \exp(\lambda |F_{q_s,\zeta}(x) - q_s I(\zeta)|)\, dx \le C q_s^{C\lambda}.$$

Proof. This is an immediate consequence of the previous two lemmas and Cauchy–Schwarz. ☐

We shall now obtain a version of this proposition for arbitrary n instead of q_s. Let $q_s \le n < q_{s+1}$ and write $n = \ell q_s + r$ where $0 \le r < q_s$ and $\ell < q_{s+1}/q_s$. Then

$$F_{n,\zeta}(x) = \sum_{h=0}^{\ell-1} F_{q_s,\zeta}(x_h) + F_{r,\zeta}(x_\ell)$$

where $x_h = x - hq_s\omega \mod 1$. By Lemma 3.1

$$\begin{aligned} |F_{n,\zeta}(x) - nI(\zeta)| &\le \sum_{h=0}^{\ell-1} |F_{q_s,\zeta}(x_h) - q_s I(\zeta)| + |F_{r,\zeta}(x_\ell) - rI(\zeta)| \\ &\le C\sum_{h=0}^{\ell-1} |\log D(x_h - \xi, q_s, \omega)| + |F_{r,\zeta}(x_\ell) - rI(\zeta)| + C\ell \log q_s. \end{aligned} \tag{3.7}$$

Let

$$\Omega_s = \bigcup_{h=0}^{\ell-1} \Big(\{k/q_s, m\omega : 0 \le k, m < q_s\} + hq_s\omega \Big) \mod 1.$$

Lemma 3.4. *With q_s and n as above,*

$$\exp\Big(\lambda \sum_{h=0}^{\ell-1} |\log D(x_h - \xi, q_s, \omega)|\Big) \le \exp(C\lambda\ell \log n) \cdot \exp(\lambda |\log \operatorname{dist}(x - \xi, \Omega_s)|)$$

for any $\lambda > 0$.

Proof. Let $D(x_{h_0} - \xi, q_s, \omega) = \min_{0 \le h < \ell} D(x_h - \xi, q_s, \omega)$. Suppose that

$$D(x_{h_0} - \xi, q_s, \omega) < q_s^{-6}$$

and moreover, that there is some $h_1 \ne h_0$ so that

$$D(x_{h_1} - \xi, q_s, \omega) < q_s^{-6}.$$

By definition, there are $y, z \in \{k/q_s, \ell\omega : k, \ell = 0, 1, \dots, q_s - 1\}$ such that

$$\|x_{h_0} - \xi - y\| < q_s^{-6}, \ \ \|x_{h_1} - \xi - z\| < q_s^{-6}.$$

This implies that

$$\Big\| (h_1 - h_0) q_s \omega + \frac{k}{q_s} + m\omega \Big\| < \frac{2}{q_s^6}$$

for some $-q_s < k, m < q_s$. In other words,

$$\Big\| u\omega + \frac{k}{q_s} \Big\| < \frac{2}{q_s^6} \tag{3.8}$$

where $1 \le |u| < q_s|h_1 - h_0| + m < q_{s+1}$. One has $q_s u\omega = t + \rho$ where $-\frac{1}{2} < \rho < \frac{1}{2}$ and $t \in \mathbb{Z}$. By (3.3), $|\rho| > (q_s u)^{-2}$, and therefore

$$\begin{aligned} \left\| u\omega + \frac{k}{q_s} \right\| &= \left\| \frac{t+k+\rho}{q_s} \right\| \ge \min\Big(\frac{1}{2q_s}, \frac{1}{q_s(q_s u)^2} \Big) \\ &\ge \frac{1}{q_s(q_s q_{s+1})^2} \ge \frac{1}{q_s^5 (\log q_s)^{2a}}. \end{aligned}$$

This contradicts (3.8). The conclusion is that there is always some h_0 such that

$$\min_{h \ne h_0} D(x_h - \xi, \omega, q_s) > q_s^{-6}.$$

Since $D(x_{h_0} - \xi, \omega, q_s) \le \operatorname{dist}(x - \xi, \Omega_s)$ and $|\log D(x_h - \xi, \omega, q_s)| \le 6 \log q_s$ if $h \ne h_0$, the lemma follows. □

Lemma 3.5. *With n and q_r as above,*

$$\int_0^1 \exp(\lambda |F_{n,\zeta}(x) - nI(\zeta)|)\, dx \le \exp\Big(C\lambda (\log n)^{a+1} \Big) \left(\int_0^1 \exp(2\lambda |F_{r,\zeta}(x) - rI(\zeta)|)\, dx \right)^{\frac{1}{2}}$$

for small $\lambda > 0$.

Proof. In view of Lemmas 3.4 and 3.2 and (3.7)

$$\begin{aligned} \int_0^1 \exp(\lambda |F_{n,\zeta}(x) - nI(\zeta)|)\, dx &\le \exp(C\lambda \ell \log n) \left(\int_0^1 \exp(2\lambda |\log \operatorname{dist}(x - \xi, \Omega_s)|)\, dx \right)^{\frac{1}{2}} \cdot \\ &\quad \cdot \left(\int_0^1 \exp(2\lambda |F_{r,\zeta}(x) - rI(\zeta)|)\, dx \right)^{\frac{1}{2}} \\ &\le C \exp\Big(C\lambda (\log n)^{a+1} \Big) (\#\Omega_s)^\lambda \left(\int_0^1 \exp(2\lambda |F_{r,\zeta}(x) - rI(\zeta)|)\, dx \right)^{\frac{1}{2}}. \end{aligned}$$

Here we have used that $\ell < \frac{q_{s+1}}{q_s} \le (\log q_s)^a \le (\log n)^a$. Since $\#\Omega_s \le C q_s \ell \le C q_{s+1} < n^2$, the lemma follows. □

Proposition 3.6. *With n and s as above,*

$$\int_0^1 \exp(\lambda |F_{n,\zeta}(x) - nI(\zeta)|)\, dx \le \exp\Big(C\lambda \sum_{j=0}^{s+1} (\log q_j)^{a+1} \Big) \le \exp\Big(C\lambda (\log n)^A \Big) \tag{3.9}$$

for small $\lambda > 0$ and any $A > a + 2$.

Proof. Recall that $n = \ell q_s + r$ where $0 \le r < q_s$. Hence $q_t \le r < q_{t+1}$ with $t + 1 \le s$. The first inequality now follows by induction in s using Lemma 3.5. To pass to the second inequality one invokes (3.5) and the general fact $q_s \ge 2^{(s-1)/2}$, see Theorem 12 in [23]. □

We shall now combine Propositions 3.3 and 3.6 with Riesz's representation (3.2) to obtain analogous results for subharmonic functions. To begin with, we need a deviation theorem for harmonic functions.

Lemma 3.7. *Let h be a 1–periodic harmonic function defined on a neighborhood of the real axis. Suppose further that $\|h\|_\infty \le 1$. Then*

$$\sup_x \Big| \sum_{k=1}^{q} h(x - k\omega) - q \int_0^1 h(y)\, dy \Big| \le C,$$

where the constant depends only on the width of the neighborhood.

Proof. Clearly,

$$\Big|\sum_{k=1}^{q} h(x-k\omega) - q\int_0^1 h(y)\,dy\Big| \le \sum_{n\neq 0} |\hat{h}(n)| \min(q, |\sum_{k=1}^{q} e^{-2\pi i k n\omega}|).$$

One has $|\hat{h}(n)| \le \frac{C}{n^4}$ with some constant depending on the width of the neighborhood. Combining this with the bound (see (3.3))

$$\Big|\sum_{k=1}^{q} e^{-2\pi i k n\omega}\Big| \le 2\|n\omega\|^{-1} \le Cn^2,$$

yields the desired result. □

The following theorem is the main result of this section.

Theorem 3.8. *Let u be a 1–periodic subharmonic function defined on a neighborhood of the real axis. Suppose furthermore that $|u(z)| \le 1$. Then for sufficiently small $\lambda > 0$*

$$\int_0^1 \exp\Big(\lambda\Big|\sum_{k=1}^{n} u(x-k\omega) - n\langle u\rangle\Big|\Big)\,dx \le \exp\Big(C\lambda(\log n)^A\Big).$$

Here A is as in Proposition 3.6 and $\langle u\rangle = \int_0^1 u(y)\,dy$. If $n = q_s$ where $\frac{p_s}{q_s}$ is a convergent of ω, then

$$\int_0^1 \exp\Big(\lambda\Big|\sum_{k=1}^{n} u(x-k\omega) - n\langle u\rangle\Big|\Big)\,dx \le Cn^{C\lambda}.$$

In particular,

$$\text{mes}\Big(\{x : |\sum_{k=1}^{n} u(x-k\omega) - n\langle u\rangle| > \delta n\}\Big) < \exp(-c\delta n + r_n) \tag{3.10}$$

where $r_n \le C(\log n)^A$ for general n and $r_n \le C\log n$ if $n = q_s$ for any s.

Proof. In view of (3.2)

$$\sum_{k=1}^{n} u(x-k\omega) - n\langle u\rangle = \sum_{k=1}^{n}\int \log|\{x-k\omega\}-\zeta|\,d\mu(\zeta) - n\int I(\zeta)\,d\mu(\zeta) + \sum_{k=1}^{n} h(\{x-k\omega\}) - n\int_0^1 h(y)\,dy.$$

By the previous lemma, it suffices to consider the contribution from the logarithmic integral. In view of (3.6)

$$\sum_{k=1}^{n}\int \log|\{x-k\omega\}-\zeta|\,d\mu(\zeta) = \int F_{n,\zeta}(x)\,d\mu(\zeta).$$

Since $\exp(\lambda\cdot)$ is a convex function Jensen's inequality implies

$$\int_0^1 \exp\Big(\lambda\Big|\int \big(F_{n,\zeta}(x) - nI(\zeta)\big)\,d\mu(\zeta)\Big|\Big)\,dx \le \int_0^1\int \exp\Big(\lambda\|\mu\|\big|F_{n,\zeta}(x) - nI(\zeta)\big|\Big)\,\frac{d\mu(\zeta)}{\|\mu\|}\,dx.$$

The theorem therefore follows from Propositions 3.3 and 3.6. Finally, (3.10) follows immediately from the integral estimates via Markov's inequality. □

4. A large deviation theorem for monodromy matrices

We now turn to the equation

$$-\psi_{n+1} - \psi_{n-1} + V(\theta + n\omega)\psi_n = E\psi_n.$$

Let

$$A(\theta, E) = \begin{pmatrix} V(\theta) - E & -1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$M_n(\theta, E) = \prod_{j=1}^{n} A(\theta + j\omega, E)$$

is the monodromy matrix of the Schrödinger equation at energy E. Assuming that the potential V is analytic,

$$u_n(x) = \frac{1}{n} \log \|M_n(x, E)\|$$

can be extended to a neighborhood of the real axis as a subharmonic, 1–periodic function. This observation goes back to Herman [21]. Moreover, its size is bounded by a constant depending only on V, $|E|$, and the width of that neighborhood. As usual, $L_n(E) = \int_{\mathbb{T}} u_n(y)\, dy$. We start with the following lemma. Notice that it sharpens the large deviation theorem in [4], cf. Lemma 1.1.

Lemma 4.1. *For any $\delta > 0$ and any positive integer n,*

$$\mathrm{mes}\big(\{x \in \mathbb{T} : |u_n(x) - L_n(E)| > \delta\}\big) \leq \exp\Big(-c\delta^2 n + C(\log n)^A\Big).$$

The constants c, C only depend on the size of E, the potential, and ω. A is determined by the constant a, see (3.3) *and Proposition 3.6.*

Proof. We may assume that $\delta > n^{-\frac{1}{2}}$. Since $\sup_x |u_n(x - \omega) - u_n(x)| \leq \frac{C}{n}$, one can choose $K = [\delta n / C]$ with some large constant C so that

$$\Big| u_n(x) - \frac{1}{K} \sum_{k=1}^{K} u_n(x - k\omega) \Big| \leq \delta/2.$$

Thus

$$\begin{aligned} \mathrm{mes}\big(\{x : |u_n(x) - L_n(E)| > \delta\}\big) &\leq \mathrm{mes}\Big(\Big\{x : \Big|\frac{1}{K}\sum_{k=1}^{K} u_n(x - k\omega) - L_n(E)\Big| > \delta/2\Big\}\Big) \\ &\leq \exp\Big(-c\delta K + C(\log K)^A\Big), \end{aligned}$$

where the second inequality follows from Theorem 3.8 with some fixed choice of $\lambda > 0$. □

We shall sharpen this estimate in section 7 by replacing δ^2 with δ. This will be accomplished by means of the avalanche principle Proposition 2.2. In fact, our main application of Lemma 4.1 will be to prove an important form of the avalanche principle, see Lemma 4.3 below. First we need to derive an auxiliary fact concerning the speed of convergence of $L_n(E)$ to $L(E)$. The essential statement in the following lemma is that $L_n(E) \to L(E)$ *uniformly* on any compact interval on which $L(E)$ is positive.

Lemma 4.2. *Suppose $L(E) > \gamma > 0$. Then*

$$0 \leq L_n(E) - L(E) < C\frac{\log n}{n},$$

where $C = C(\gamma, |E|, V, \omega)$. In particular, $L_n(E) \to L(E)$ uniformly on any compact interval on which $L(E)$ is positive.

Proof. Clearly, $0 \le L_n(E) \le C_0 = C_0(V, |E|)$ for all n. Let t be a positive integer such that $t\gamma > 16C_0$. Given $n > 10$, let $\ell_0 = [C_1 \log n]$ with C_1 to be specified below. Consider the integers $\ell_0, 2\ell_0, \dots, 2^t\ell_0$. There is some $0 \le j < t$ such that (with $\ell_j = 2^j \ell_0$)

$$L_{\ell_j}(E) - L_{\ell_{j+1}}(E) < \frac{\gamma}{16}. \tag{4.1}$$

For if not, then $C_0 > L_{\ell_0}(E) - L_{2^t\ell_0}(E) \ge t\gamma/16 > C_0$, which is a contradiction. Let $\ell = 2^j\ell_0$ with the choice of j satisfying (4.1).

We shall now apply Proposition 2.2 to the matrices $A_j = A_j(x) = M_\ell(x + (j-1)\ell\omega)$ for $1 \le j \le m = [n/\ell]$ and with $\mu = \exp(\ell\gamma/2)$. Notice that $\mu > n^2$ if $C_1 > 4/\gamma$. By Lemma 4.1,

$$\min_{1\le j\le m} \|A_j(x)\| \ge \exp(\ell(L_\ell(E) - \gamma/2)) \ge \mu > n^2$$

up to a set of $x \in \mathbb{T}$ of measure not exceeding

$$m \exp(-c_0(\gamma/2)^2\ell) < n \exp(-c_0 C_1 \gamma^2 \log n/4) < n^{-2}$$

provided $C_1\gamma^2 > 12/c_0$. Furthermore, in view of (4.1) and Lemma 4.1,

$$\begin{aligned} &\max_{0\le j<m} \Big[\log\|A_{j+1}(x)\| + \log\|A_j(x)\| - \log\|A_{j+1}A_j(x)\|\Big] \\ \le\;& 2\ell(L_\ell(E) + \frac{\gamma}{32}) - 2\ell(L_{2\ell}(E) - \frac{\gamma}{32}) \le 2\ell(L_\ell(E) - L_{2\ell}(E) + \frac{\gamma}{16}) \\ \le\;& 2\ell(\frac{\gamma}{16} + \frac{\gamma}{16}) \le \frac{1}{2}\log\mu = \frac{\gamma}{4}\ell, \end{aligned}$$

up to a set of x of measure not exceeding

$$2m \exp(-c_0(\gamma/32)^2\ell) \le 2n \exp(-c_0(\gamma/32)^2 C_1 \log n) < n^{-2}$$

if $C_1\gamma^2$ is large. We have shown that (2.1) and (2.2) hold for all $x \in \mathcal{G}$, where $|\mathbb{T} \setminus \mathcal{G}| < 2n^{-2}$. We conclude from (2.4) that

$$\Big|\log\|M_{\ell m}(x,E)\| + \sum_{j=1}^{m-2} \log\|M_\ell(x + j\ell\omega, E)\| - \sum_{j=0}^{m-2} \log\|M_\ell(x + (j+1)\ell\omega, E) M_\ell(x + j\ell\omega, E)\|\Big| \le Cn^{-1}$$

for all $x \in \mathcal{G}$. Arguing similarly for $\log\|M_{\ell m}(x + \ell m\omega)\|$ and $\log\|M_{2\ell m}(x)\|$ and comparing the respective estimates yields

$$\begin{aligned} &\Big|\log\|M_{2\ell m}(x,E)\| - \log\|M_{\ell m}(x + \ell m\omega, E)\| - \log\|M_{\ell m}(x,E)\| + \log\|M_\ell(x + \ell m\omega, E)\| + \\ &+ \log\|M_\ell(x + (m-1)\ell\omega, E)\| - \log\|M_\ell(x + \ell m\omega, E) M_\ell(x + (m-1)\ell\omega, E)\|\Big| \le \frac{C}{n} \end{aligned} \tag{4.2}$$

up to a set of x not exceeding Cn^{-2} in measure. Since (recall $m = [n/\ell]$)

$$\Big|\log\|M_n(x)\| - \log\|M_{\ell m}(x)\|\Big| \le C\ell \quad \text{and} \quad \Big|\log\|M_\ell(x)\|\Big| \le C\ell, \tag{4.3}$$

we conclude from (4.2) that

$$\Big|\log\|M_{2n}(x,E)\| - \log\|M_n(x + n\omega, E)\| - \log\|M_n(x,E)\|\Big| \le C \log n$$

up to a set of x not exceeding Cn^{-2} in measure. Integrating over x therefore implies that

$$|L_{2n}(E) - L_n(E)| \le C\frac{\log n}{n},$$

where $C = C(\gamma, |E|, V, \omega)$. Summing over $2^k n$ finally proves the lemma. ∎

The following lemma represents a form of the avalanche principle which will turn out to be relevant for certain applications below. Its proof is similar to the proof of the previous lemma, but it will be important for us to know that the speed of convergence of $L_n(E) \to L(E)$ is controlled by γ.

Lemma 4.3. *Suppose $L(E) > \gamma > 0$. Let $n = \sum_{j=1}^m \ell_j$ with $[C_1 \log n] \le \ell_j \le 2[C_1 \log n]$ and set $s_j = \sum_{k=1}^j \ell_k$. Then one can choose $C_1 = C_1(\gamma)$ sufficiently large so that there exists a set $\mathcal{G} = \mathcal{G}(n, E) \subset \mathbb{T}$ satisfying $|\mathbb{T} \setminus \mathcal{G}| < n^{-2}$ with the property that*

$$\Big|\log \|M_n(x,E)\| + \sum_{j=2}^{m-1} \log \|M_{\ell_j}(x+s_{j-1}\omega, E)\| - \sum_{j=1}^{m-1} \log \|M_{\ell_{j+1}}(x+s_j\omega, E) M_{\ell_j}(x+s_{j-1}\omega, E)\|\Big| \le \frac{C_2}{n} \tag{4.4}$$

for all $x \in \mathcal{G}$. Both C_1 and C_2 depend only on γ, $|E|$, V, and ω.

Proof. Set $\ell = [C_1 \log n]$ where C_1 will be specified below. We shall apply Proposition 2.2 with $A_j = A_j(x) = M_{\ell_j}(x + s_{j-1}\omega, E)$ and $\mu = \exp(\ell\gamma/2)$. In view of Lemma 4.1 with $\delta = \delta_0 = \gamma/100$ and $C_1\gamma^2$ sufficiently large

$$\min_{1\le j\le m} \|A_j(x)\| \ge \exp(\ell_j L_{\ell_j}(E)/2) \ge \exp(\gamma[C_1 \log n]/2) = \mu > n^2 \tag{4.5}$$

up to a set of x of measure less than

$$m \exp(-c\delta_0^2 \ell) \le n \exp(-c\delta_0^2 C_1 \log n) \le n^{-2}. \tag{4.6}$$

As for the second condition (2.2), let $n_0(E)$ be sufficiently large such that

$$L_k(E) - L(E) < \delta_0 \tag{4.7}$$

for all $k > [C_1 \log n_0]$. Lemma 4.2 implies that n_0 depends only on γ and the size of E. Applying Lemma 4.1 again with the same choice of δ_0 yields (suppressing E for simplicity)

$$\begin{aligned} &\log \|A_{j+1}(x)\| + \log \|A_j(x)\| - \log \|A_{j+1}A_j(x)\| \le \ell_{j+1}(L_{\ell_{j+1}} + \delta_0) + \ell_j(L_{\ell_j} + \delta_0) \\ &\quad -(\ell_j + \ell_{j+1})(L_{\ell_j+\ell_{j+1}} - \delta_0) \\ &\le \ \ell_{j+1}(2\delta_0 + L_{\ell_{j+1}} - L) + \ell_j(2\delta_0 + L_{\ell_j} - L) \le 12\ell\delta_0 < \frac{1}{2}\log\mu = \frac{1}{4}\ell\gamma \end{aligned} \tag{4.8}$$

up to a set of x of measure at most n^{-2}, see (4.6). Let $\mathcal{G}$ be the set of x satisfying both (4.5) and (4.8). We have shown that $|\mathbb{T} \setminus \mathcal{G}| \le n^{-2}$ for an appropriate choice of C_1 and provided $n > n_0(\gamma)$. In view of Proposition 2.2 therefore

$$\Big|\log \|M_n(x)\| + \sum_{j=2}^{m-1} \log \|M_{\ell_j}(x+s_{j-1}\omega)\| - \sum_{j=1}^{m-1} \log \|M_{\ell_{j+1}}(x+s_j\omega) M_{\ell_j}(x+s_{j-1}\omega)\|\Big| \le \frac{C_2}{n}$$

for all $x \in \mathcal{G}$, as claimed. □

5. An estimate for the speed of convergence of $L_n(E)$ to $L(E)$

The purpose of the following theorem is to remove the $\log n$ factor in Lemma 4.2.

Theorem 5.1. *Suppose that for some fixed compact interval I one has $L(E) > \gamma > 0$ for all $E \in I$. Then the estimate*

$$0 \le L_n(E) - L(E) < \frac{C_0}{n}$$

holds for all $n = 1, 2, \ldots$ and $E \in I$ with some $C_0 = C_0(\gamma, I, V, \omega)$.

Proof. Fix some $E \in I$. It suffices to show that $L_n(E) - L_{2n}(E) < \frac{C}{n}$. Fix some large n and let $n = \sum_{j=1}^m \ell_j$ where $[C_1 \log n] \le \ell_j \le 2[C_1 \log n]$ with C_1 being the constant from Lemma 4.3. It is clear that this can be done with all but two of the ℓ_j being equal. In fact, we can assume that $\ell_1 = \ell_m$. Let $\mathcal{G} = \mathcal{G}(n, E)$ be as in Lemma 4.3. Then (4.4) holds simultaneously for all x and $x + n\omega$ provided $x \in \mathcal{G} \cap (\mathcal{G} - n\omega)$. Partitioning the interval $[1, 2n]$ into intervals of length $\ell_1, \dots, \ell_m, \ell_1, \dots, \ell_m$ in this order and applying Lemma 4.3 once more, one obtains a similar estimate for $\log \|M_{2n}(x, E)\|$ off a set of measure at most n^{-2}. Comparing the three estimates (4.4) for $\log \|M_n(x, E)\|$, $\log \|M_n(x + n\omega, E)\|$, and $\log \|M_{2n}(x, E)\|$, respectively, one concludes that

$$\begin{aligned} &\Big| \log \|M_{2n}(x, E)\| - \log \|M_n(x + n\omega, E)\| - \log \|M_n(x, E)\| + \log \|M_{\ell_1}(x + n\omega, E)\| + \\ &+ \log \|M_{\ell_m}(x + s_{m-1}\omega, E)\| - \log \|M_{\ell_1}(x + n\omega, E) M_{\ell_m}(x + s_{m-1}\omega, E)\| \Big| \le \frac{C}{n} \end{aligned} \tag{5.1}$$

up to set of x of size at most n^{-2}. Since the terms on the left–hand side of (5.1) are no bigger than Cn in absolute value, integrating (5.1) over $\mathbb{T}$ yields (recall $\ell_1 = \ell_m$)

$$|2n(L_{2n}(E) - L_n(E)) - 2\ell_1(L_{2\ell_1}(E) - L_{\ell_1}(E))| \le \frac{C}{n}.$$

This implies that the function $R(n) = 2n(L_{2n}(E) - L_n(E))$ satisfies

$$R(n) \le R(\ell_1) + \frac{C}{n}.$$

Since $[C_1 \log n] \le \ell_1 \le 2[C_1 \log n]$, iteration leads to

$$R(n) \le \frac{C}{n} + \frac{C}{C_1 \log n} + \frac{C}{C_1 \log(C_1 \log n)} + \dots + R(k_0) \tag{5.2}$$

where $k_0 < C_1^2$, say. As the sum in (5.2) clearly gives a bounded contribution, the theorem follows. ☐

The following proposition shows that the positive quantities $L_{2\ell}(E) - L(E)$ and $L_\ell(E) - L_{2\ell}(E)$ differ only by an amount that is *exponentially small in ℓ*.

Proposition 5.2. *Suppose $L(E) > \gamma > 0$. Then there exists a constant $c_1 = c_1(\gamma, |E|, V, \omega) > 0$ such that*

$$|L(E) - 2L_{2\ell}(E) + L_\ell(E)| \le \exp(-c_1 \ell) \quad \text{for all } \ell = 1, 2, \dots. \tag{5.3}$$

Moreover, there is $\ell_0 = \ell_0(c_1)$ so that if $L_{\ell_1}(E) - L(E) > 4\exp(-c_1 \ell_1)$ for some $\ell_1 \ge \ell_0$, then

$$L_{2^k \ell_1}(E) - L(E) > \frac{1}{2^{k+1}}(L_{\ell_1}(E) - L(E))$$

for all $k \ge 0$. In other words, on intervals of positivity of L either $L_\ell(E) \to L(E)$ exponentially fast, or $L_n(E) - L(E) > \frac{C(E)}{n}$ for infinitely many n.

Proof. Let C_1 be as in Lemma 4.3. Set $\ell = [C_1 \log n]$ and write $n = m\ell + r$ where $0 \le r < \ell$. In view of (4.4) and (4.3),

$$\Big| \log \|M_n(x)\| + \sum_{j=0}^{m-1} \log \|M_\ell(x + j\ell\omega)\| - \sum_{j=0}^{m-1} \log \|M_\ell(x + (j+1)\ell\omega) M_\ell(x + j\ell\omega)\| \Big| \le C\ell$$

for all x up to a set of measure at most n^{-2}. Since the left–hand side is no bigger than Cn for any x, integrating over $\mathbb{T}$ yields

$$|nL_n(E) + m\ell L_\ell(E) - 2m\ell L_{2\ell}(E)| \le C\ell$$

or

$$|L_n(E) + L_\ell(E) - 2L_{2\ell}(E)| \le \frac{C\ell}{n}.$$

Replacing $L_n(E)$ with $L(E)$ by means of Theorem 5.1 establishes (5.3).

Now assume that $L_{\ell_1}(E) - L(E) > 4\exp(-c_1\ell_1)$. In view of (5.3),

$$L_{2\ell_1}(E) - L(E) > \frac{1}{2}(L_{\ell_1}(E) - L(E)) - \frac{1}{2}\exp(-c_1\ell_1).$$

Continuing inductively one obtains that

$$\begin{aligned} L_{2^k\ell_1}(E) - L(E) \quad > \quad & \frac{1}{2^k}(L_{\ell_1}(E) - L(E)) - \frac{1}{2^k}\exp(-c_1\ell_1)\Big[1 + 2\exp\big(-(2-1)c_1\ell_1\big) + \\ & + \ldots + 2^{k-1}\exp\big(-(2^{k-1}-1)c_1\ell_1\big)\Big]. \end{aligned} \tag{5.4}$$

Now choose ℓ_0 so large that

$$\sum_{j=0}^{\infty} 2^j \exp\big(-(2^j - 1)c_1\ell_0\big) \le 2.$$

By (5.4) and our assumption,

$$L_{2^k\ell_1}(E) - L(E) > \frac{1}{2^k}(L_{\ell_1}(E) - L(E)) - \frac{2}{2^k}\exp(-c_1\ell_1) > \frac{1}{2^{k+1}}(L_{\ell_1}(E) - L(E))$$

as claimed. ∎

6. Hölder continuity of the Lyapunov exponent and the integrated density of states

Theorem 6.1. *Let $N(E)$ be the integrated density of states. Assume that $L(E) > \gamma > 0$ for all $E \in I$ where I is some compact interval. Then there exists $\beta = \beta(\gamma, I, V, \omega) > 0$ such that*

$$|L(E) - L(E')| + |N(E) - N(E')| \le C|E - E'|^{\beta}$$

for all $E, E' \in I$ where $C = C(\gamma, I, V, \omega)$.

Proof. Fix some $E, E' \in I$ and let C_1 be as in Lemma 4.3. Let n be a large integer to be specified below. Write $n = m\ell + r$ with $\ell = [C_1 \log n]$ and $0 \le r < \ell$. By Lemma 4.3 and (4.3)

$$\Big|\log\|M_n(x,E)\| + \sum_{j=0}^{m-1}\log\|M_\ell(x + j\ell\omega, E)\| - \sum_{j=0}^{m-1}\log\|M_\ell(x + (j+1)\ell\omega, E)M_\ell(x + j\ell\omega, E)\|\Big| \le C\ell \tag{6.1}$$

$$\Big|\log\|M_n(x,E')\| + \sum_{j=0}^{m-1}\log\|M_\ell(x + j\ell\omega, E')\| - \sum_{j=0}^{m-1}\log\|M_\ell(x + (j+1)\ell\omega, E')M_\ell(x + j\ell\omega, E')\|\Big| \le C\ell \tag{6.2}$$

provided $x \in \mathcal{G}(n, E) \cap \mathcal{G}(n, E')$. It is clear that

$$\sup_{x\in\mathbb{T}}\Big\|\frac{d}{dE}M_\ell(x, E)\Big\| \le \exp(C_3\ell) \tag{6.3}$$

with a constant C_3 depending only on the potential and the size of E. Since $\|M_\ell\| \ge 1$ one therefore has

$$\Big|\log\|M_\ell(y,E)\| - \log\|M_\ell(y,E')\|\Big| \le \Big|\log\Big[1 + \frac{\|M_\ell(y,E) - M_\ell(y,E')\|}{\|M_\ell(y,E)\|}\Big]\Big| \le \exp(C_3\ell)|E - E'| \tag{6.4}$$

for all $y \in \mathbb{T}$ and similarly for $M_{2\ell}$. Subtracting (6.1) from (6.2) yields by means of (6.3) that for all $x \in \mathcal{G}(n, E) \cap \mathcal{G}(n, E')$

$$\Big|\frac{1}{n}\log\|M_n(x,E)\| - \frac{1}{n}\log\|M_n(x,E')\|\Big| \le \exp(2C_3\ell)|E - E'| + \frac{C\ell}{n} \le \frac{C\log n}{n}$$

provided $|E - E'| < \frac{1}{n}\exp(-2C_3\ell)$. Integrating over $\mathbb{T}$ and invoking Theorem 5.1 finally implies that

$$|L(E) - L(E')| \le \frac{C\log n}{n} \quad \text{if } |E - E'| < n^{-4C_1C_3}.$$

This proves the stated bound for the Lyapunov exponent. The bound on the integrated density of states follows from the Hölder continuity of the Lyapunov exponent via the Thouless formula and standard properties of the Hilbert transform. This is well–known, see e.g. Figotin, Pastur [12], chapters 11.B and 11.C, and also Section 10 below. □

7. A sharp large deviation theorem for monodromy matrices

In this section we replace δ^2 with δ in Lemma 4.1. Notice that this increases the range of deviations we can control from roughly $[n^{-\frac{1}{2}}, 1]$ to $[(\log n)^A/n, 1]$. As the former region is precisely the one in the random case, one therefore sees that the quasi–periodic case behaves differently in this respect. In fact, the main point is that $\|M_n(x)\|$ can basically be written as a product of shifts of some function, cf. Proposition 2.2 and Lemma 4.3.

Theorem 7.1. *Let $u_n(x) = \frac{1}{n}\log\|M_n(x,E)\|$ and assume that $L(E) > \gamma > 0$. Then with some $c = c(\gamma)$*

$$\mathrm{mes}\Big(\{x\in\mathbb{T} : |u_n(x) - L(E)| > \delta\}\Big) \le \exp(-c\delta n + r_n) \tag{7.1}$$

where $r_n \le C(\log n)^A$ for general n and $r_n \le C\log n$ if $n = q_s$ for any s. Moreover, the set on the left–hand side of (7.1) is contained in no more than Cn many intervals. Also,

$$\int_0^1 \exp(\lambda|\log\|M_n(x,E)\| - nL(E)|)\,dx \le \exp(r_n) \tag{7.2}$$

provided $0 < \lambda < \lambda_0(\gamma, |E|, V, \omega)$.

Proof. For the sake of simplicity, we fix E and we shall not indicate the dependence on E. Take n large and let $\ell = [\delta n]$ where $\frac{C_4\log n}{n} < \delta < \frac{1}{10}$. We claim that for all x up to a set of measure not exceeding $\exp(-c\gamma^2\delta n)$

$$\Big|\log\|M_n(x)\| + \sum_{j=0}^{m-1}\log\|M_\ell(x+j\ell\omega)\| - \sum_{j=0}^{m-1}\log\|M_\ell(x+(j+1)\ell\omega)M_\ell(x+j\ell\omega)\|\Big| \le C\ell \tag{7.3}$$

where $m = [n/\ell]$. This follows from Proposition 2.2 and Lemma 4.1. More precisely, let $\delta_0 = \gamma/100$ and $\mu = \exp(\ell\gamma/2)$. By Lemma 4.1 and Lemma 4.2 with $C_4\gamma$ large,

$$\min_{0\le j<n}\|M_\ell(x+j\omega)\| \ge \exp(\ell L_\ell - \delta_0\ell) > \mu > n$$

$$\max_{0\le j<n}\Big[\log\|M_\ell(x+(j+1)\omega)\| + \log\|M_\ell(x+j\omega)\| - \log\|M_\ell(x+(j+\ell)\omega)M_\ell(x+j\omega)\|\Big] \le \frac{1}{2}\log\mu$$

up to a set of x of measure not exceeding

$$Cn\exp(-c\delta_0^2\ell) < \exp(-\frac{1}{2}c\delta_0^2\ell) \tag{7.4}$$

if $C_4\gamma^2$ is large. This guarantees the conditions (2.1) and (2.2) not just for x but also for $x + k\omega$ with $0 \le k < \ell$. We conclude that there is $\mathcal{G}\subset\mathbb{T}$ with $|\mathbb{T}\setminus\mathcal{G}|$ bounded by (7.4) such that (7.3) holds for all $x + k\omega$, $k = 0, 1, \ldots, \ell-1$ provided $x\in\mathcal{G}$. Consequently,

$$\Big|\frac{1}{\ell}\sum_{k=0}^{\ell-1}\log\|M_n(x+k\omega)\| + \sum_{j=0}^{n-1}\frac{1}{\ell}\log\|M_\ell(x+j\omega)\| - \sum_{j=0}^{n-1}\frac{1}{\ell}\log\|M_\ell(x+(j+\ell)\omega)M_\ell(x+j\omega)\|\Big| \le C\ell \tag{7.5}$$

for all $x\in\mathcal{G}$. Since

$$\Big|\log\|M_n(x)\| - \frac{1}{\ell}\sum_{k=0}^{\ell-1}\log\|M_n(x+k\omega)\|\Big| < C\ell = C\delta n$$

for all x, one can rewrite (7.5) in the form

$$\Big|\log\|M_n(x)\| + \sum_{j=0}^{n-1}\frac{1}{\ell}\log\|M_\ell(x+j\omega)\| - \sum_{j=0}^{n-1}\frac{1}{\ell}\log\|M_\ell(x+(j+\ell)\omega)M_\ell(x+j\omega)\|\Big| \le C\delta n.$$

In view of Theorem 3.8 the sums in this expression differ from a constant by more than δn on a set of measure at most $\exp(-c\delta n + r_n)$. Therefore, $\log\|M_n(x)\|$ differs from its mean by more than δn on a set of measure not exceeding (cf. (7.4))

$$|\mathbb{T}\setminus\mathcal{G}| + \exp(-c\delta n + r_n) < Cn\exp(-c\gamma^2\delta n) + \exp(-c\delta n + r_n),$$

as claimed. The boundedness of the integral (7.2) follows from (7.1) by integrating over level sets.

To obtain the statement about intervals we will basically show that the function $u_n(x)$ does not have more than Cn many intervals of monotonicity. Let

$$M_n(z) = \begin{bmatrix} f_n(z) & g_n(z) \\ r_n(z) & s_n(z) \end{bmatrix}$$

with analytic functions f_n, g_n, r_n, s_n on $D(0,2)$ (it is possible to identify these entries as certain determinants, see (11.13)). For any $x \in \mathbb{R}$,

$$\|M_n(x)\| \asymp f_n^2(x) + g_n^2(x) + r_n^2(x) + s_n^2(x). \tag{7.6}$$

Denote the right–hand side of (7.6) by v_n. Then v_n is analytic on $D(0,2)$, and $|v_n| \le \exp(Cn)$ on that disk. Therefore also $|v_n'(z)| \le \exp(Cn)$ for all $|z| \le \frac{7}{8}$. Since $\|M_n(x)\| \ge 1$ for all x, one has $|v_n(x)| \ge c$ for all x with some small absolute constant c. We claim that $|v_n'(x_0)| \ge 1$ for some $x_0 \in [-1/8, 1/8]$. Suppose not. Then $|v_n(x) - v_n(y)| \le 1$ for all $x, y \in [-1/8, 1/8]$ and thus $|\log v_n(x) - \log v_n(y)| \le C$ on that interval. Therefore,

$$|u_n(x) - u_n(y)| \le \frac{C}{n} + \frac{1}{n}|\log v_n(x) - \log v_n(y)| \le \frac{C}{n} \tag{7.7}$$

for any $x, y \in [-1/8, 1/8]$. Since $|u_n(x) - u_n(x+\ell\omega)| \le \frac{C\ell}{n}$, inequality (7.7) holds for all $x, y \in \mathbb{T}$. But this implies that

$$\sup_{x\in\mathbb{T}} |u_n(x) - L_n(E)| \le \frac{C}{n},$$

so that in view of Theorem 5.1 the left–hand side of (7.1) is empty provided $\delta > \frac{C}{n}$. But for smaller values of δ (7.1) is trivial. Hence the claim. By Jensen's formula

$$\int_0^1 \log|v_n'(x_0 + \tfrac{3}{2}e^{2\pi i\theta})|\,d\theta - \log|v_n'(x_0)| = \sum_j \log\frac{3}{2|z_j|},$$

where the sum runs over all the zeros z_j of $v_n'(\cdot + x_0)$. Since the left–hand side is no bigger than Cn and $|x_0| \le \frac{1}{8}$, we conclude that

$$\mathrm{card}\{j : |z_j| \le 1\} \le Cn.$$

Consequently, v_n has at most Cn monotonicity intervals on $\mathbb{T}$. Therefore,

$$\{x \in \mathbb{T} : |\log v_n - nL(E)| > \delta n\}$$

is contained in no more than Cn intervals for any n. Since $\Big|u_n - \frac{1}{n}\log v_n\Big| \le \frac{C}{n}$, the same statement holds for u_n and any $\delta > \frac{C}{n}$. Since these are the only relevant values of δ, the theorem follows. ∎

Choosing $\delta = n^{-\tau}$ for some $0 < \tau < 1$ in (7.1), one obtains

$$\text{mes}\Big(\{x \in \mathbb{T} : |u_n(x) - L(E)| > n^{-\tau}\}\Big) \quad \leq \quad \exp(-cn^{1-\tau}). \tag{7.8}$$

We shall now indicate that these estimates are sharp, at least if $\tau \leq \frac{1}{2}$. More precisely, pick an x_0 such that

$$\log \|M_n(x_0, E)\| > nL(E) - n^{1-\tau}.$$

Also let $\gamma = \max(1 - \tau, \frac{1}{2})$. We require the following bound:

$$\sup_{x \in \mathbb{T}} \frac{1}{n} \log \|M_n(x, E)\| \quad \leq \quad L_n(E) + C_\epsilon n^{-\frac{1}{2}+\epsilon} \tag{7.9}$$

for all n and $\epsilon > 0$. This is proved in [4], Lemma 2.1 provided on chooses the parameters there appropriately. (7.9) can also be proved by the methods of the previous sections, even without the ϵ. Furthermore, we shall use the following algebraic fact (Trotter's identity)

$$A_n A_{n-1} \ldots A_1 - B_n B_{n-1} \ldots B_1 \quad = \quad \sum_{j=1}^{n} \prod_{i=1}^{n-j} A_{j+i}(A_j - B_j) \prod_{k=1}^{j-1} B_k. \tag{7.10}$$

In view of (7.9), (7.10), and Theorem 5.1

$$\begin{aligned} &\|M_n(x_0, E) - M_n(x, E)\| \leq C|x_0 - x| \sum_{j=1}^{n} \|M_{n-j}(x + j\omega, E)\| \|M_{j-1}(x_0, E)\| \\ &\leq C|x - x_0| \sum_{j=1}^{n} \exp\Big((n-j)L(E) + Cn^{\frac{1}{2}+\epsilon}\Big) \exp\Big((j-1)L(E) + Cn^{\frac{1}{2}+\epsilon}\Big) \\ &\leq C|x - x_0| \, n \exp(nL(E) + Cn^{\frac{1}{2}+\epsilon}). \end{aligned}$$

Therefore, by our choice of x_0,

$$\begin{aligned} \Big|\frac{1}{n} \log \|M_n(x, E)\| - \frac{1}{n} \log \|M_n(x_0, E)\|\Big| \quad &\leq \quad \frac{1}{n} \log\Big[1 + \frac{\|M_n(x, E) - M_n(x_0, E)\|}{\|M_n(x_0, E)\|}\Big] \\ &\leq \quad C|x - x_0| \frac{\exp\big(nL(E) + Cn^{\frac{1}{2}+\epsilon}\big)}{\exp\big(nL(E) - n^{1-\tau}\big)} \leq C|x - x_0| \exp\big(Cn^{\gamma+\epsilon}\big). \end{aligned} \tag{7.11}$$

Hence if $|x - x_0| < \exp(-2n^{\gamma+2\epsilon})$ and n is large, then

$$\Big|\frac{1}{n} \log \|M_n(x, E)\| - \frac{1}{n} \log \|M_n(x_0, E)\|\Big| < \exp\big(-n^{\gamma+2\epsilon}\big).$$

Consequently, if the set on the left–hand side of (7.8) is nonempty, then it has to contain an interval of size at least $\exp(-2n^{\gamma+2\epsilon})$. This proves that for large n,

$$\text{mes}\Big(\{x \in \mathbb{T} : |u_n(x) - L(E)| > n^{-\tau}\}\Big) \geq \exp(-cn^{\gamma+\epsilon})$$

unless the set on the left–hand side is empty. Hence (7.8) cannot be improved if $\tau \leq \frac{1}{2}$.

8. Cartan's Theorem in higher dimensions

The purpose of this section is to develop analytical tools to prove large deviation theorems in the case of several frequencies. The approach chosen here is not the only available one. In fact, [4] contains a direct proof of a large deviation theorem for monodromy matrices in the case of several frequencies by means of Fourier series. The approach chosen here, however, is more flexible in terms of the dynamics and it also leads to better exponents. See the following section for further discussion.

Definition 8.1. Let $0 < H < 1$. For any subset $\mathcal{B} \subset \mathbb{C}$ we say that $\mathcal{B} \in \mathrm{Car}_1(H)$ if $\mathcal{B} \subset \bigcup_j D(z_j, r_j)$ with

$$\sum_j r_j \leq C_0 H. \tag{8.1}$$

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

$$\mathcal{B} = \{(z_1, z_2, \ldots, z_d) : (z_2, \ldots, z_d) \in \mathcal{B}_0 \text{ or } z_1 \in \mathcal{B}(z_2, \ldots, z_d) \text{ for some } \mathcal{B}(z_2, \ldots, z_d) \in \mathrm{Car}_1(H)\}.$$

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

Notice that the absolute constant C_0 is not specified in this definition. This allows one to say that the union of two Cartan sets (with the same parameters d and H) is again a Cartan set but with $2C_0$ instead of C_0. It is important, however, that C_0 will always be an absolute constant which is implicitly defined by the context in which it arises. The following lemma collects some well–known facts, see [27] and [24]. For the definition of Riesz measures see the beginning of Section 3.

Lemma 8.2. *Suppose $u : D(0, 2) \mapsto [-1, 1]$ is a subharmonic function. Let μ be the Riesz measure of u. For any $z_0 \in D(0, \frac{1}{2})$, $0 < r < \frac{1}{2}$, and $H \in (0, 1)$ there exists $\mathcal{B} \in \mathrm{Car}_1(H)$ so that*

$$|u(z) - u(z')| \quad < \quad C\Big[\mu(D(z_0, r)) \log \frac{1}{H} + |z - z'|\Big(1 + \int_{D(0,1)\setminus D(z_0,r)} \frac{d\mu(\zeta)}{|z_0 - \zeta|}\Big)\Big] \tag{8.2}$$

for all $z, z' \in D(z_0, r/2) \setminus \mathcal{B}$. In particular, if for some $A \geq 1$

$$M_1\mu(z_0) \quad = \quad \sup_{0<t<\frac{1}{2}} \frac{\mu(D(z_0, t))}{t} \leq A, \tag{8.3}$$

then

$$|u(z) - u(z')| < C\,A\Big[r \log \frac{1}{H} + |z - z'| \log \frac{1}{r}\Big] \tag{8.4}$$

for all $z, z' \in D(z_0, r/2) \setminus \mathcal{B}$.

Proof. Let $D_0 = D(z_0, r)$. It is well–known, see Koosis [24], that for any $z = re^{i\theta}$ with $r < 1$

$$\begin{aligned} u(z) \quad &= \quad \int_{D(0,1)} \log \frac{|z - \zeta|}{|1 - z\bar{\zeta}|}\, d\mu(\zeta) + \int_0^{2\pi} \frac{1 - r^2}{1 - 2r\cos(\phi - \theta) + r^2} u(e^{i\phi}) \frac{d\phi}{2\pi} &(8.5)\\ &= \quad \int_{D_0} \log|z - \zeta|\, d\mu(\zeta) + \int_{D(0,1)\setminus D_0} \log|z - \zeta|\, d\mu(\zeta) - \int_{D(0,1)} \log|1 - z\bar{\zeta}|\, d\mu(\zeta) + h(z) &(8.6)\\ &= \quad v(z) + w(z) + g(z). &(8.7)\end{aligned}$$

We denoted the Poisson integral in (8.5) by h, the functions v and w stand for the first and second integrals in line (8.6), respectively, and g is the sum of the other two. If $z, z' \in D(z_0, r/2)$, then $|z|, |z'| < \frac{3}{4}$ and thus

$$|g(z) - g(z')| < C \int_{D(0,1)\setminus D_0} \frac{|z - z'|}{1 - |z|}\, d\mu(\zeta) + \sup_{|\zeta| \leq \frac{3}{4}} |\nabla h(\zeta)||z - z'| \leq C|z - z'|\Big[1 + \mu(D(0, 1))\Big]. \tag{8.8}$$

According to Jensen's formula [27] section 7.2

$$\int_0^{2\pi} u\Big(\frac{3}{2} e^{i\theta}\Big) \frac{d\theta}{2\pi} - u(0) = \int_{|z| \leq \frac{3}{2}} \log\Big(\frac{3}{2|z|}\Big)\, d\mu(z)$$

and therefore

$$\mu(D(0, 1)) \leq \frac{2}{3 \log \frac{3}{2}}. \tag{8.9}$$

With z, z' as above,

$$|w(z) - w(z')| \leq C \int_{D(0,1)\setminus D_0} \frac{|z-z'|}{|z-\zeta|} \, d\mu(\zeta) \leq C \int_{D(0,1)\setminus D_0} \frac{|z-z'|}{|z_0-\zeta|} \, d\mu(\zeta). \tag{8.10}$$

Since $r < \frac{1}{2}$ one has $v \leq 0$ on D_0. By Cartan's theorem, see [27] Section 11.2, there is $\mathcal{B} \in \mathrm{Car}_1(H)$ with $C_0 = 5$ in (8.1) so that

$$v(z) > \mu(D_0) \log(H/e) \quad \text{if } z \in D_0 \setminus \mathcal{B}. \tag{8.11}$$

Estimate (8.2) follows from (8.7), (8.8), (8.9), (8.10), and (8.11). Finally, if (8.3) holds, then (8.4) follows from (8.2). □

The following theorem presents a version of the previous lemma that applies to functions of two variables which are subharmonic in each variable. Although there is a corresponding result for functions on $\mathbb{C}^d$ for any $d > 2$ we first present the proof on $\mathbb{C}^2$, as the argument turns out to be more efficient in that case.

Theorem 8.3. *Let u be a continuous function on $D(0,2) \times D(0,2) \subset \mathbb{C}^2$ so that $|u| \leq 1$. Suppose further that*

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

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

$$|u(z_1, z_2) - u(z_1', z_2')| < C_\gamma \, r^{1-2\gamma} \log \frac{1}{r} \quad \text{for all} \quad (z_1, z_2), (z_1', z_2') \in \Pi \setminus \mathcal{B} \tag{8.12}$$

$$H = \exp\Big(-r^{-\gamma}\Big). \tag{8.13}$$

Proof. For any $z_1 \in D(0,2)$ define

$$v(z_1) = \int_{-1}^{1} u(z_1, x_2) \, dx_2. \tag{8.14}$$

$v : D(0,2) \to \mathbb{R}$ is a subharmonic function such that $|v| \leq 2$ with Riesz measure μ_v. Let $M_1\mu_v$ be the maximal function given by (8.3). Clearly, M_1 satisfies the usual weak–type L^1 inequality

$$\mathrm{mes}(\{x_1 \in [-1,1] : M_1\mu_v(x_1) > \lambda\}) \leq \frac{C}{\lambda} \mu_v(D(0, \frac{3}{2})).$$

In particular, there is some $x_1^{(0)} \in [-1,1]$ so that $M_1\mu_v(x_1^{(0)}) \leq C$. For any $z_2 \in D(0,2)$ let

$$g_t(z_2) = \int_0^1 u(x_1^{(0)} + te^{2\pi i\theta}, z_2) \, d\theta - u(x_1^{(0)}, z_2). \tag{8.15}$$

By Jensen's formula, see Theorem 2 in Section 7.2 of [27],

$$g_t(z_2) = \int_{|z_1 - x_1^{(0)}| < t} \log \frac{t}{|z_1 - x_1^{(0)}|} \, \mu(dz_1, z_2) = \int_0^t \frac{n(s, z_2)}{s} \, ds \tag{8.16}$$

where $n(s, z_2) = \mu(D(x_1^{(0)}, s), z_2)$ with the Riesz measure $\mu(\cdot, z_2)$ of $u(\cdot, z_2)$. Clearly,

$$\mu_v(D(x_1^{(0)}, s)) = \int_{-1}^{1} n(s, x_2) \, dx_2.$$

Therefore, in view of (8.15), (8.16), and our choice of $x_1^{(0)}$,

$$\int_{-1}^{1} g_t(x_2) \, dx_2 = \int_0^t \frac{\mu_v(D(x_1^{(0)}, s))}{s} \, ds \leq Ct. \tag{8.17}$$

Now fix some $r \in (0, 1/2)$ and define

$$G = \sum_{0 \le j < C \log \frac{1}{r}} 2^{-j} g_{2^j r}.$$

The subharmonicity of $z_1 \mapsto u(z_1, z_2)$ implies that $g_t \ge 0$ so that G is the sum of nonnegative terms. By (8.17)

$$\int_{-1}^{1} G(x_2)\, dx_2 \le C r \log \frac{1}{r}$$

and thus

$$\operatorname{mes}\big(\{x_2 \in [-1,1] : G(x_2) > C r \log \frac{1}{r}\}\big) < \frac{1}{2} \tag{8.18}$$

provided C is a sufficiently large absolute constant. For technical reasons we introduce the auxiliary subharmonic function

$$h(z_2) \;=\; \int_0^1 u(x_1^{(0)} + r^2 e^{2\pi i \theta}, z_2)\, d\theta \quad \text{for any } z_2 \in D(0,2). \tag{8.19}$$

Clearly, $|h| \le 1$ and we denote the Riesz measure of h by μ_h. As before, $\mu_h(D(0,3/2)) \le C$. The function g_t introduced in (8.15) is the difference of two subharmonic functions on $D(0,2)$. Let μ_t and μ_0 be their respective Riesz measures. As before, most points $x_2 \in [-1,1]$ satisfy

$$M_1\Big(\sum_{0 \le j < C \log \frac{1}{r}} \mu_{2^j r} + \mu_0 + \mu_h\Big)(x_2) \le C \log \frac{1}{r}. \tag{8.20}$$

In view of Lemma 8.2, for any such x_2 there exists $\mathcal{B}_0(x_2) \in \mathrm{Car}_1(\exp(-r^{-\gamma}))$ so that

$$\sup_{0 \le j < C \log \frac{1}{r}} |g_{2^j r}(z_2) - g_{2^j r}(z_2')| < C r^{1-\gamma} \log \frac{1}{r} \quad \text{for all } \ z_2, z_2' \in D(x_2, r) \setminus \mathcal{B}_0(x_2). \tag{8.21}$$

Combining (8.18) and (8.21) yields a point $x_2^{(0)} \in [-1,1]$ with the property that

$$g_{2^j r}(z_2) \le C[2^j r + r^{1-\gamma}] \log \frac{1}{r} \quad \text{for all } \ z_2 \in D(x_2^{(0)}, r) \setminus \mathcal{B}_0 \ \text{ and all } 0 \le j < C \log \frac{1}{r}.$$

Here we have set $\mathcal{B}_0 = \mathcal{B}_0(x_2^{(0)})$. Using (8.16) this immediately leads to

$$\mu(D(x_1^{(0)}, 2^j r), z_2) \le C[2^j r + r^{1-\gamma}] \log \frac{1}{r}$$

for all z_2 and j as before. Inserting this bound into (8.2) with $H = \exp(-r^{-\gamma})$ and $r^{1-\gamma}$ instead of r one obtains for any such z_2 a Cartan set $\mathcal{B}(z_2) \in \mathrm{Car}_1(H)$ so that

$$\begin{aligned} |u(z_1, z_2) - u(z_1', z_2)| \;&\le\; C\Big[r^{1-\gamma} \log \frac{1}{r} \log \frac{1}{H} + |z_1 - z_1'| \log^2 \frac{1}{r}\Big] \\ &\le\; r^{1-2\gamma} \log \frac{1}{r} \quad \text{for any } \ z_1, z_1' \in D(x_1^{(0)}, r^{1-\gamma}) \setminus \mathcal{B}(z_2). \end{aligned} \tag{8.22}$$

To control the deviation in z_2 we invoke the auxiliary subharmonic function h from above. Because of (8.20) Lemma 8.2 implies that

$$|h(z_2) - h(z_2')| \;\le\; C r^{1-\gamma} \log \frac{1}{r} \quad \text{for all } \ z_2, z_2' \in D(x_2^{(0)}, r) \setminus \mathcal{B}_1 \tag{8.23}$$

where $\mathcal{B}_1 \in \mathrm{Car}_1(H)$, $H = \exp(-r^{-\gamma})$. By the definition of a Cartan set and (8.22),

$$|h(z_2) - u(z_1, z_2)| \le C[r^{1-2\gamma} \log \frac{1}{r} + r^{-2} H] \ \text{ for all } \ z_2 \in D(x_2^{(0)}, r) \setminus \mathcal{B}_0, z_1 \in D(x_1^{(0)}, r) \setminus \mathcal{B}(z_2). \tag{8.24}$$

Let $\Pi = D(x_1^{(0)}, r^{1-\gamma}) \times D(x_2^{(0)}, r)$ and

$$\mathcal{B} = \{(z_1, z_2) : z_2 \in \mathcal{B}_0 \cup \mathcal{B}_1 \ \text{ or } z_2 \in D(x_2^{(0)}, r) \setminus \mathcal{B}_0 \cup \mathcal{B}_1 \ \text{ and } z_1 \in \mathcal{B}(z_2)\}.$$

In view of Definition 8.1, $\mathcal{B} \in \mathrm{Car}_2(H)$ with $H = \exp(-r^{-\gamma})$. Combining (8.24) with (8.23) implies that

$$|u(z_1, z_2) - u(z_1', z_2')| \le C r^{1-2\gamma} \log \frac{1}{r} \quad \text{for all } \ (z_1, z_2), (z_1', z_2') \in \Pi \setminus \mathcal{B},$$

as claimed. □

Remark 8.4. Under the same assumptions as in Theorem 8.3 the previous proof implies the following statement: Given $r \in (0,1)$ there exists a polydisk $\Pi = D(z_1^{(0)}, r) \times D(z_2^{(0)}, r^2) \subset D(0,1) \times D(0,1)$ with $z_1^{(0)}, z_2^{(0)} \in D(0,1)$ and a set $\mathcal{B} \in \mathrm{Car}_2(H)$ so that

$$|u(z_1, z_2) - u(z_1', z_2')| < C\, r \ \text{ for all } \ (z_1, z_2), (z_1', z_2') \in \Pi \setminus \mathcal{B}$$
$$H = \exp\Big(-r^{-1}\Big).$$

The point here is that the center of the polydisk is no longer restricted to the real plane. Since this fact is not useful to us, we do not supply a detailed proof (which is, however, very similar to the previous one).

We now turn to the case of higher dimensions. The following theorem is formulated in all dimensions for technical reasons, but Theorem 8.3 is superior to it if $d = 2$.

Theorem 8.5. *Let d be a positive integer. Suppose $u : D(0,2)^d \to [-1,1]$ is subharmonic in each variable, i.e., $z_1 \mapsto u(z_1, z_2, \ldots, z_d)$ is subharmonic for any choice of $(z_2, \ldots, z_d) \in D(0,2)^{d-1}$ and similarly for each of the other variables. Given $r \in (0,1)$ there exists a polydisk $\Pi = D(x_1^{(0)}, r) \times \ldots \times D(x_d^{(0)}, r) \subset \mathbb{C}^d$ with $x_1^{(0)}, \ldots, x_d^{(0)} \in [-1,1]$ and a Cartan set $\mathcal{B} \in \mathrm{Car}_d(H)$ so that*

$$|u(z_1, \ldots, z_d) - u(z_1', \ldots, z_d')| < C\, r^{\beta} \ \ \textit{for all } (z_1, \ldots, z_d), (z_1', \ldots, z_d') \in \Pi \setminus \mathcal{B} \tag{8.25}$$

$$H = \exp\Big(-r^{-\beta}\Big). \tag{8.26}$$

The constant $\beta > 0$ depends only on the dimension d. Furthermore, given $u_1, \ldots, u_k$ each of which satisfies the hypotheses of the theorem, there are Π and $\mathcal{B}$ as above so that (8.25) *holds simultaneously for each of the $u_1, \ldots, u_k$ with a constant Ck instead of C.*

Proof. We start with the case $d = 1$. Given subharmonic functions $u_1, \ldots, u_k$ each of which is bounded by one on $D(0,2)$ we let $\mu_1, \ldots, \mu_k$ be their respective Riesz measures. There exists a point $x^{(0)} \in [-1,1]$ such that

$$M_1[\mu_1 + \ldots + \mu_k](x^{(0)}) < Ck.$$

The theorem now follows from Lemma 8.2 with $\Pi = D(x_0, r)$ and $\beta = \frac{1}{2}$.
Now let $d \ge 2$ and suppose the theorem is true for $d-1$ and we will prove it for d. The proof is similar to that of the previous theorem and we shall only sketch the argument. Fix some $r \in (0,1)$ and let v be the bounded subharmonic function on $D(0,2)$ given by

$$v(z_1) = \int_{-1}^{1} \ldots \int_{-1}^{1} u(z_1, x_2, \ldots, x_d)\, dx_2 \ldots dx_d.$$

We denote the Riesz measure of v by μ_v. Pick some $x_1^{(0)}$ so that $M_1 \mu_v(x_1^{(0)}) \le C$. For any $(z_2, \ldots, z_d) \in D(0,2)^{d-1}$ define

$$\begin{aligned} g_t(z_2, \ldots, z_d) &= \int_0^1 u(x_1^{(0)} + te^{2\pi i\theta}, z_2, \ldots, z_d)\, d\theta - u(x_1^{(0)}, z_2, \ldots, z_d) \\ h(z_2, \ldots, z_d) &= \int_0^1 u(x_1^{(0)} + r^{2d} e^{2\pi i\theta}, z_2, \ldots, z_d)\, d\theta. \end{aligned} \tag{8.27}$$

Applying the induction hypothesis (with $d-1$ and $k \asymp [\log \frac{1}{r}]$) to the functions given by the right–hand sides of (8.27) for all $t = 2^j r^d$ with $j = 0, \dots, C[\log \frac{1}{r}]$ and h, one obtains a polydisk $\Pi = D(x_2^{(0)}, r) \times \dots \times D(x_d^{(0)}, r) \subset D(0,1)^{d-1}$ with real $x_2^{(0)}, \dots, x_d^{(0)}$, and a Cartan set $\mathcal{B}_1 \in \mathrm{Car}_{d-1}(H_1)$ with $H_1 = \exp(-r^{-\beta})$ such that

$$\sup_{0 \le j < C \log \frac{1}{r}} |g_{2^j r^d}(q) - g_{2^j r^d}(q')| + |h(q) - h(q')| < C r^\beta \log \frac{1}{r} \tag{8.28}$$

for any $q = (z_2, \dots, z_d)$, $q' = (z_2', \dots, z_d') \in \Pi \setminus \mathcal{B}_1$. As above we let

$$G = \sum_{0 \le j < C \log \frac{1}{r}} 2^{-j} g_{2^j r^d}.$$

The same calculation as in (8.17) yields

$$\int_{-1}^{1} \dots \int_{-1}^{1} G(x_2, \dots, x_d)\, dx_2 \dots dx_d \le C r^d \log \frac{1}{r}.$$

Therefore,

$$\mathrm{mes}\Big(\{ \Pi \cap \mathbb{R}^{d-1} : G > C \lambda r^d \log \frac{1}{r} \} \Big) < \lambda^{-1}$$

for any $\lambda > 1$. Since $\mathrm{mes}[\Pi \cap \mathbb{R}^{d-1}] = C r^{d-1}$, one has in particular that for some large C

$$\mathrm{mes}\Big(\{ \Pi \cap \mathbb{R}^{d-1} : G > C\, r \log \frac{1}{r} \} \Big) < \frac{1}{2} \mathrm{mes}[\Pi \cap \mathbb{R}^{d-1}].$$

For small r this implies in conjunction with (8.28) that

$$g_{2^j r^d}(z_2, \dots, z_d) \quad < \quad C\Big[r^\beta + 2^j r \Big] \log \frac{1}{r} \ \text{ for all } z_2, \dots, z_d \in \Pi \setminus \mathcal{B}_1. \tag{8.29}$$

Recall

$$g_t(z_2, \dots, z_d) = \int_{|z_1 - x_1^{(0)}| < t} \log \frac{t}{|z_1 - x_1^{(0)}|}\ \mu(dz_1, z_2, \dots, z_d)$$

where $\mu(\cdot, z_2, \dots, z_d)$ is the Riesz measure of $u(\cdot, z_2, \dots, z_d)$. We therefore conclude from (8.29) that

$$\mu(D(x_1^{(0)}, 2^j r^d), z_2, \dots, z_d) < C\Big[r^\beta + 2^j r \Big] \log \frac{1}{r} \ \text{ for all } z_2 \in \Pi \setminus \mathcal{B}_1 \ \text{ and all } \ 0 \le j < C \log \frac{1}{r}.$$

Assuming as we may that $\beta < 1$ Lemma 8.2 implies that

$$|u(z_1, z_2, \dots, z_d) - u(z_1', z_2, \dots, z_d)| < C\Big[r^\beta \log \frac{1}{H_2} + r^{\beta - d} \log \frac{1}{r} |z_1 - z_1'| \Big] \log \frac{1}{r}$$

if $z_1, z_1' \in D(x_1^{(0)}, r^d/2) \setminus \mathcal{B}(z_2, \dots, z_d)$ where $\mathcal{B}(z_2, \dots, z_d) \in \mathrm{Car}_1(H_2)$. Setting $H_2 = \exp\big(-r^{-\beta/2}\big)$ one obtains for any $(z_2, \dots, z_d)$, $(z_2', \dots, z_d') \in \Pi \setminus \mathcal{B}_1$

$$|u(z_1, z_2, \dots, z_d) - u(z_1', z_2, \dots, z_d)| \quad < \quad C r^{\beta/2} \log \frac{1}{r} \ \text{ if } z_1, z_1' \in D(x_1^{(0)}, r^d/2) \setminus \mathcal{B}(z_2, \dots, z_d). \tag{8.30}$$

Combining the deviation estimate (8.28) for h with the following easy consequence of (8.30)

$$|u(z_1, z_2, \dots, z_d) - h(z_2, \dots, z_d)| < C\big[r^{\beta/2} + r^{-2d} H_2 \big]$$

yields (8.25) and (8.26) with $\frac{\beta}{2d} - \epsilon$ instead of β.

One easily checks that this argument can be applied to $u_1, \dots, u_k$ simultaneously and the theorem follows. $\square$

9. A large deviation theorem for monodromy matrices in the multifrequency case

In this section we consider the Schrödinger equation

$$-\psi_{n+1} - \psi_{n-1} + V(\theta_1 + n\omega_1, \dots, \theta_d + n\omega_d)\psi_n = E\psi_n \tag{9.1}$$

where $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d) \in \mathbb{T}^d$ is arbitrary, $\boldsymbol{\omega} = (\omega_1, \dots, \omega_d) \in \mathbb{T}^d$ is an ergodic shift, and V is a real–analytic function on $\mathbb{T}^d$. We assume further that V extends to an analytic function on $D(0,2)^d$. Let

$$A(\boldsymbol{\theta}, E) = \begin{pmatrix} V(\boldsymbol{\theta}) - E & -1 \\ 1 & 0 \end{pmatrix}.$$

As usual,

$$M_n(\boldsymbol{\theta}, E) = \prod_{j=1}^{n} A(\boldsymbol{\theta} + j\boldsymbol{\omega}, E) \tag{9.2}$$

is the associated monodromy matrix. Furthermore, let

$$L_n(E) = \frac{1}{n}\int_{\mathbb{T}^d} \log\|M_n(\boldsymbol{\theta}, E)\|\, d\boldsymbol{\theta} \;\text{ and }\; L(E) = \inf_n L_n(E) = \lim_{n\to\infty} L_n(E) \tag{9.3}$$

be the Lyapunov exponents.

In [4] Bourgain and Goldstein proved a large deviation theorem for $\log\|M_n\|$ based on a Fourier series expansion of this function and (3.2), see Lemma 8.1. We show in this section how to obtain a similar statement by means of the Cartan type estimate from the previous section. This approach has the advantage that it generalizes immediately to other types of dynamics than shifts whereas the method from [4] appears to be rather restrictive. It is essential, however, that the map underlying the dynamics extends to an analytic function Φ on a polydisk Π containing the torus in such a way that Φ does not expand in the imaginary direction. In particular, it seems that the skew shift requires ideas beyond those presented here.

We are going to assume that $\boldsymbol{\omega}$ satisfies

$$\|\boldsymbol{\omega} \cdot \boldsymbol{k}\| \ge \frac{C(\epsilon_1)}{|\boldsymbol{k}|^{d+\epsilon_1}} \;\text{ for all nonzero } \boldsymbol{k} \in \mathbb{Z}^d \tag{9.4}$$

where $\epsilon_1 > 0$ is small. It is well–known that a.e. $\boldsymbol{\omega}$ satisfies (9.4) for any $\epsilon_1 > 0$.

The main results of this section are as follows.

Proposition 9.1. *Let $\boldsymbol{\omega}$ be as in (9.4). Suppose the function u satisfies the hypotheses of Theorem 8.5. Assume furthermore that for some $n \ge 1$*

$$\sup_{\boldsymbol{\theta}\in\mathbb{T}^d} |u(\boldsymbol{\theta} + \boldsymbol{\omega}) - u(\boldsymbol{\theta})| < \frac{1}{n}. \tag{9.5}$$

Then there exist $\sigma > 0$, $\tau > 0$, and c_0 only depending on d and ϵ_1 such that

$$\mathrm{mes}\big(\{\boldsymbol{\theta} \in \mathbb{T}^d : |u(\boldsymbol{\theta}) - \langle u\rangle| > n^{-\tau}\}\big) < \exp(-c_0 n^{\sigma}). \tag{9.6}$$

Here $\langle u\rangle = \int_{\mathbb{T}^d} u(\boldsymbol{\theta})\, d\boldsymbol{\theta}$. If $d = 2$ then one obtains the range $0 < \tau < \frac{1}{3} - \epsilon_2$ and $\sigma = \frac{1}{3} - \tau - \epsilon_2$ where $\epsilon_2 \to 0$ as $\epsilon_1 \to 0$.

Proof. Let $r \in (0,1)$ be arbitrary. By Theorem 8.5 there is some rectangle $R = \Pi \cap \mathbb{T}^d$ on $\mathbb{T}^d$ with $\mathrm{diam}(R) = 2r$ and a set $\mathcal{B} \subset \mathbb{T}^d$ such that

$$|u(\boldsymbol{\theta}) - u(\boldsymbol{\theta}')| < r^{\beta} \;\text{ for any } \boldsymbol{\theta}, \boldsymbol{\theta}' \in R \setminus \mathcal{B} \tag{9.7}$$

$$\mathrm{mes}(\mathcal{B}) < \exp\big(-r^{-\beta}\big). \tag{9.8}$$

Notice that it is essential to know the structure of Cartan sets to deduce the estimate for $\mathrm{mes}(\mathcal{B})$, cf. 8.1. It is well–known that for any $\boldsymbol{\theta} \in \mathbb{T}^d$ there exists $0 \le k < k_0 = [Cr^{-d-\epsilon_1}]$ so that $\boldsymbol{\theta} + k\boldsymbol{\omega} \in R$. Since (9.5) implies that

$$|u(\boldsymbol{\theta} + k\boldsymbol{\omega}) - u(\boldsymbol{\theta})| < \frac{k}{n}$$

one obtains from (9.7)

$$|u(\boldsymbol{\theta}) - u(\boldsymbol{\theta}')| < r^{\beta} + \frac{k_0}{n} \quad \text{for any } \boldsymbol{\theta}, \boldsymbol{\theta}' \in \mathbb{T}^d \setminus \widetilde{\mathcal{B}}$$

where $\widetilde{\mathcal{B}} = \bigcup_{k=0}^{k_0} (\mathcal{B} + k\boldsymbol{\omega}) \mod \mathbb{Z}^d$. Letting $r = n^{-\frac{1}{d+\beta+\epsilon_1}}$ one now obtains (9.6) with $\sigma = \tau = \frac{\beta}{d+\beta+\epsilon_1}$.
If $d = 2$ Theorem 8.3 gives better results. Indeed, fix some $\gamma, r \in (0,1)$ and let Π be as in Theorem 8.3. Setting $\epsilon_1 = 0$ for simplicity, recall that for any $\boldsymbol{\theta} \in \mathbb{T}^2$ there is some integer $0 < k < k_0$ such that $\boldsymbol{\theta} + k\boldsymbol{\omega} \in \Pi$. Here k_0 needs to satisfy

$$r^{\gamma-2} k_0^{-1} + r^{-1} k_0^{-1} < c$$

for some small constant c. Since $\gamma < 1$ one can take $k_0 \asymp r^{\gamma-2}$. Therefore, for any $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \mathbb{T}^2$,

$$|u(\boldsymbol{\theta}) - u(\boldsymbol{\theta}')| < r^{1-2\gamma} + \frac{r^{\gamma-2}}{n} \quad \text{for any } \boldsymbol{\theta}, \boldsymbol{\theta}' \in \mathbb{T}^d \setminus \widetilde{\mathcal{B}}$$

with $\widetilde{\mathcal{B}}$ as above. Setting $r = n^{-\frac{1}{3(1-\gamma)}}$ yields $\sigma = \frac{1}{3} - \tau$, as desired. $\square$

Corollary 9.2. *Let $\boldsymbol{\omega}$ be as in* (9.4)*. Let $S_n(E)$ be a positive number satisfying*

$$S_n(E) \ge \sup_{(z_1,\dots,z_d) \in D(0,2)^d} \Big[\frac{1}{n} \log \|M_n(z_1,\dots,z_d,E)\| + 2 \log \|A(z_1,\dots,z_d,E)\| \Big]. \tag{9.9}$$

Then there exist $\sigma > 0$, $\tau > 0$, and c_0 only depending on d and ϵ_1 such that

$$\mathrm{mes}\big(\{\boldsymbol{\theta} \in \mathbb{T}^d : |\log \|M_n(\boldsymbol{\theta}, E)\| - nL_n(E)| > S_n(E) n^{1-\tau}\}\big) < \exp(-c_0 n^{\sigma}). \tag{9.10}$$

If $d = 2$ then one obtains the range $0 < \tau < \frac{1}{3} - \epsilon_2$ and $\sigma = \frac{1}{3} - \tau - \epsilon_2$ where $\epsilon_2 \to 0$ as $\epsilon_1 \to 0$.

Proof. Fix some dimension d and energy E. Define for any $(z_1,\dots,z_d) \in D(0,2)^d$

$$u_n(z_1,\dots,z_d) = \frac{1}{nS_n} \log \|M_n(z_1,\dots,z_d,E)\|.$$

Then u_n is a continuous subharmonic function bounded by one in $D(0,2)^d$. Furthermore, u_n satisfies (9.5). Hence (9.10) is an immediate consequence of (9.6). $\square$

Remark 9.3. Usually one has bounded potentials and energies so that basically $S_n(E) \asymp 1$. In proving positivity of the Lyapunov exponent, however, it will be necessary to consider large potentials and then the statement of (9.10) will be convenient.

The method from [4] yields exponents $\tau = \frac{3}{4}(\frac{1}{3} - \sigma) - \epsilon_2$ for $d = 2$. This can be easily checked by making appropriate choices for the parameters in the proof of Lemma 8.1 in [4]. Therefore, our method is slightly more economical here. Moreover, the approach in [4] seems to be rather restrictive in terms of the dynamics, whereas our argument applies to any transformation that does not stretch in the imaginary direction.

10. Modulus of continuity for the integrated density of states in the multifrequency case

Fix some dimension $d \geq 2$ and let N denote the integrated density of states for equation (9.1) with d frequencies. Let σ be the exponent arising in Theorem 9.2. It turns out that N has modulus of continuity $\exp(-|\log t|^{\sigma})$ on any interval on which the Lyapunov exponent is positive. To obtain Hölder continuity for N using the methods of this paper one would need to prove (9.10) with $\sigma = 1$ and deviations of size δ, cf. Lemma 4.1 and Theorem 6.1. In what follows let $S(E) = \sup_n S_n(E)$ where $S_n(E)$ is defined in (9.9). Before turning to the discussion of continuity we require a version of Lemma 4.2 for the multifrequental case.

Lemma 10.1. *Fix some dimension $d \geq 2$ and let σ, τ be as in Proposition 9.2. Suppose $L(E) > \gamma > 0$ where $L(E)$ is the Lyapunov exponent* (9.3)*. Then*

$$0 \leq L_n(E) - L(E) < C\frac{(\log n)^{1/\sigma}}{n},$$

where $C = C(\gamma, |E|, V, \sigma, \tau)$. In particular, $L_n(E) \to L(E)$ uniformly on any compact interval on which $L(E)$ is positive.

Proof. The proof is basically the same as that of Lemma 4.2. The only difference is that here $\ell = [C(\log n)^{1/\sigma}]$ with some large C and that one uses Proposition 9.2 instead of Lemma 4.1. We leave the details to the reader. □

Proposition 10.2. *Fix some dimension $d \geq 2$ and suppose $L(E) > \gamma > 0$ for all $E \in I$, where I is some interval. Then*

$$|L(E) - L(E')| + |N(E) - N(E')| \quad \leq \quad C\exp(-C^{-1}|\log|E - E'||^{\sigma})$$

for all $E, E' \in I$ with $C = C(\gamma, \sigma, \tau, S, I)$. Here σ and τ are the exponents from (9.10) *and we set $S = \sup_{E\in I} S(E)$. If $d = 2$ one can take $\sigma < \frac{1}{3}$.*

Proof. The proof is similar to that of Theorem 6.1. Fix some $E, E' \in I$ and let n be some large integer. Write $n = m\ell + r$ with $\ell = [C_1(\log n)^{1/\sigma}]$ and $0 \leq r < \ell$. We claim that for some large constant C_1 and all large n

$$\Big|\log\|M_n(\boldsymbol{\theta}, E)\| + \sum_{j=0}^{m-1} \log\|M_\ell(\boldsymbol{\theta} + j\ell\boldsymbol{\omega}, E)\| - \sum_{j=0}^{m-1} \log\|M_\ell(\boldsymbol{\theta} + (j+1)\ell\boldsymbol{\omega}, E)M_\ell(\boldsymbol{\theta} + j\ell\boldsymbol{\omega}, E)\|\Big| \leq C\ell \tag{10.1}$$

$$\Big|\log\|M_n(\boldsymbol{\theta}, E')\| + \sum_{j=0}^{m-1} \log\|M_\ell(\boldsymbol{\theta} + j\ell\boldsymbol{\omega}, E')\| - \sum_{j=0}^{m-1} \log\|M_\ell(\boldsymbol{\theta} + (j+1)\ell\boldsymbol{\omega}, E')M_\ell(\boldsymbol{\theta} + j\ell\boldsymbol{\omega}, E')\|\Big| \leq C\ell \tag{10.2}$$

for all $\boldsymbol{\theta} \in \mathcal{G}_n(E, E')$ where $\mathrm{mes}(\mathbb{T}^d \setminus \mathcal{G}_n(E, E')) < n^{-1}$. This follows from Proposition 2.2 with $A_j = A_j(\boldsymbol{\theta}) = M_\ell(\boldsymbol{\theta} + j\ell\boldsymbol{\omega}, E)$, $\mu = \exp(\ell\gamma/2)$ and similarly for E'. In fact, if $S\ell^{-\tau} \leq \gamma/2$

$$\min_{1\leq j\leq m} \|A_j(\boldsymbol{\theta})\| \geq \exp(\ell(L_\ell(E) - S\ell^{-\tau})) \geq \exp\Big(\gamma[C_1(\log n)^{1/\sigma}]/2\Big) = \mu > n^2 \tag{10.3}$$

up to a set of $\boldsymbol{\theta}$ of measure less than, see (9.10)

$$m\exp(-c_0\ell^{\sigma}) \leq n\exp(-c_0 C_1^{\sigma}\log n) \leq n^{-2}. \tag{10.4}$$

The second condition (2.2) of Proposition (2.2) is checked as in (4.8), and we skip the details. It is clear that

$$\sup_{\boldsymbol{\theta}\in\mathbb{T}^d} \Big\|\frac{d}{dE}M_\ell(\boldsymbol{\theta}, E)\Big\| \leq \exp(C_3\ell) \tag{10.5}$$

with a constant C_3 depending only on the potential and the size of E. Since $\|M_\ell\| \geq 1$ one therefore has

$$\Big|\log\|M_\ell(\boldsymbol{\theta}, E)\| - \log\|M_\ell(\boldsymbol{\theta}, E')\|\Big| \leq \Big|\log\Big[1 + \frac{\|M_\ell(\boldsymbol{\theta}, E) - M_\ell(\boldsymbol{\theta}, E')\|}{\|M_\ell(\boldsymbol{\theta}, E)\|}\Big]\Big| \leq \exp(C_3\ell)|E - E'| \tag{10.6}$$

for all $\boldsymbol{\theta} \in \mathbb{T}^d$ and similarly for $M_{2\ell}$. Subtracting (10.1) from (10.2) yields by means of (10.5) that for all $\boldsymbol{\theta} \in \mathcal{G}_n(E, E')$

$$\left|\frac{1}{n}\log\|M_n(\boldsymbol{\theta}, E)\| - \frac{1}{n}\log\|M_n(\boldsymbol{\theta}, E')\|\right| \le \exp(2C_3\ell)|E - E'| + \frac{C\ell}{n} \le \frac{C(\log n)^{1/\sigma}}{n}$$

provided $|E - E'| < \frac{1}{n}\exp(-2C_3\ell)$. Integrating over $\mathbb{T}^d$ and invoking Lemma 10.1 finally implies that

$$|L(E) - L(E')| \le \frac{C(\log n)^{1/\sigma}}{n} \quad \text{if } |E - E'| < \exp\Big(-C(\log n)^{1/\sigma}\Big).$$

This proves the statement of Proposition 10.2 on the Lyapunov exponent. The corresponding bound on the integrated density of states can be derived from it fairly easily via the Thouless formula, see Theorem 11.8 in [12],

$$L(E) = \int \log|E - E'|\, dN(E') \quad \text{for all } E \in \mathbb{R} \tag{10.7}$$

and some elementary properties of the Hilbert transform H. Since the modulus of continuity involved is not so common, we provide some details. Fix I as above and let $J \subset I$ be an interval with the same center as I but half the length. Pick a smooth cutoff function ψ with compact support so that $\psi = 1$ on I. Define $N_1 = \psi N$ and $N_2 = N - N_1$. Then for almost every E

$$HN_1(E) = \int N_1(E')\frac{dE'}{E - E'} = -\int \log|E - E'|\, dN_2(E') + L(E) = g(E) \quad . \tag{10.8}$$

This follows from (10.7) by replacing $\log|E - E'|$ with $\log(|E - E'| + \epsilon)$, integrating by parts, and then letting $\epsilon \to 0+$. Let J_0 be an interval centered at 0 with $|J_0| = |J|$ and pick a smooth cutoff function ϕ with support inside J_0 and $\phi(0) = 1$. Define H_J to be the operator with kernel $k_J(x) = \phi(x)\frac{1}{x} = \phi(x)k(x)$. The operator $H_J H$ has the kernel $(\phi k) * k$. Taking Fourier transforms one obtains

$$\widehat{(\phi k) * k}(\xi) = \hat{\phi} * \hat{k}(\xi) \cdot \hat{k}(\xi) = -1 + \hat{R}(\xi)$$

where $|\hat{R}(\xi)| \le C_m(1 + |\xi|)^{-m}$ for any positive m. This follows from $\hat{k}(\xi) = -i\,\mathrm{sign}(\xi)$, $\int \hat{\phi}(\xi)\, d\xi = 1$, and the fact that $\hat{\phi}$ has rapidly decreasing tails. Consequently, R is a smooth kernel. Applying H_J to (10.8) therefore leads to

$$H_J H\, N_1 = -N_1 + R * N_1 = H_J g.$$

Since $R * N_1$ is a smooth function, the theorem follows from the fact that

$$|g(E) - g(E')| \le C\exp(-c|\log|E - E'||^\sigma) \quad \text{on } I$$

and the following lemma. □

Let ρ be a modulus of continuity with the property that

$$\sum_{n \ge \ell} \rho(2^{-n}) \asymp \rho(2^{-\ell}) \quad \text{for any } \ell \in \mathbb{Z} \tag{10.9}$$

$$\sum_{n < \ell} 2^n \rho(2^{-n}) \asymp 2^\ell \rho(2^{-\ell}) \quad \text{for any } \ell \in \mathbb{Z}. \tag{10.10}$$

Examples of such ρ are $\rho(t) = t^\alpha$ with $0 < \alpha < 1$ and $\rho(t) = \exp(-c|\log t|^\sigma)$ with $\sigma > 0$, the latter one being relevant for Theorem 1. Let

$$\mathcal{C}_\rho = \{f : \mathbb{R} \to \mathbb{R} \,:\, |f(x) - f(y)| \le A\,\rho(|x - y|) \ \text{ for all } x, y \text{ and for some } A\}$$

and let $[f]_\rho$ denote the minimum of all such A. The following lemma provides a fairly standard characterization of the spaces $\mathcal{C}_\rho$ in terms of the Fourier transform and states that they are preserved under singular integrals, see [35], chapter VI, Section 5.3. It is formulated by means of the Littlewood–Paley projections Δ_n that localize the Fourier transform to a dyadic block of size 2^n at a distance 2^n from the origin.

Lemma 10.3. *A function $f : \mathbb{R} \to \mathbb{R}$ lies in $\mathcal{C}_\rho$ iff $\|\Delta_n(f)\|_\infty \le B\rho(2^{-n})$ for all n. In fact,*

$$[f]_\rho \asymp \quad \text{minimum of all such } B.$$

Moreover,

$$[Tf]_\rho \le C[f]_\rho$$

for any singular integral operator T.

Proof. This is a simple exercise and we will leave most details to the reader. One can write

$$\Delta_n f = \psi_{2^{-n}} * f$$

where ψ is a Schwartz function with mean zero and $\psi_{2^{-n}}(x) = 2^n\psi(2^n x)$. Thus

$$\begin{aligned} \|\Delta_n(f)\|_\infty &= \sup_x \left| \int \psi_{2^{-n}}(x-y)[f(y)-f(x)]\,dy \right| \\ &\le C \sum_{m \le n} 2^{m-n}\rho(2^{-m}) \asymp \rho(2^{-n}) \end{aligned} \tag{10.11}$$

by (10.10). Conversely, one writes $f = \sum_{m<n} \Delta_m(f) + \sum_{m \ge n} \Delta_m(f)$ modulo a constant which yields

$$|f(x) - f(y)| \le |\sum_{m<n} \Delta_m(f)(x) - \sum_{m<n} \Delta_m(f)(y)| + 2\|\sum_{m \ge n} \Delta_m(f)\|_\infty \le CB\rho(|x-y|).$$

To obtain the last inequality one sets $2^n \asymp |x-y|^{-1}$, takes the derivative of the first sum and then applies the assumption together with (10.10) and (10.9), respectively.

To prove the bound for singular integrals let $\tilde{\Delta}_n$ be another Littlewood–Paley projection chosen such that $\tilde{\Delta}_n \Delta_n = \Delta_n$ for all n. Then

$$\|\Delta_n T f\|_\infty = \|T\tilde{\Delta}_n \Delta_n f\|_\infty \le \|T\tilde{\Delta}_n\|_{\infty\to\infty} \|\Delta_n f\|_\infty.$$

The lemma now follows since the kernel of $T\tilde{\Delta}_n$ is bounded in L^1 uniformly in n. □

11. Positivity of the Lyapunov exponent

The main purpose of this section is to present a general mechanism that allows one to prove positivity of the Lyapunov exponent for large disorders. More precisely, consider a family of equations of the form

$$-\psi_{n+1} - \psi_{n-1} + \lambda V(T^n \boldsymbol{\theta})\psi_n = E\psi_n \tag{11.1}$$

where $\boldsymbol{\theta} \in \mathbb{T}^d$, $T : \mathbb{T}^d \to \mathbb{T}^d$ is an ergodic transformation, and V a nonconstant real–analytic function on $\mathbb{T}^d$. Let

$$A_j(\boldsymbol{\theta}, \lambda, E) = \begin{bmatrix} \lambda V(T^j\boldsymbol{\theta}) - E & -1 \\ 1 & 0 \end{bmatrix}.$$

The matrix $M_n(\boldsymbol{\theta}, \lambda, E) = \prod_{j=1}^n A_j(\boldsymbol{\theta}, \lambda, E)$ denotes the monodromy matrix of the equation (11.1) . As before,

$$L_n(\lambda, E) = \frac{1}{n}\int_{\mathbb{T}^d} \log \|M_n(\boldsymbol{\theta}, \lambda, E)\|\,d\boldsymbol{\theta}$$

and $L(\lambda, E) = \lim_{n\to\infty} L_n(\lambda, E)$ exists. Finally, let $S(\lambda, E)$ be a number satisfying

$$S(\lambda, E) \asymp \sup_{n \ge 1} \sup_{\boldsymbol{\theta} \in \mathbb{T}^d} \frac{1}{n} \log \|M_n(\boldsymbol{\theta}, \lambda, E)\|. \tag{11.2}$$

The main result of this section is as follows: If the large deviation theorem (with some $\sigma > 0$)

$$\int_{\mathbb{T}^d} \left| \frac{1}{n} \log \| M_n(\boldsymbol{\theta}, \lambda, E) \| - L_n(\lambda, E) \right| d\boldsymbol{\theta} \quad \leq \quad C\, S(\lambda, E) n^{-\sigma} \tag{11.3}$$

holds for all $n = 1, 2, \ldots$, then

$$\inf_E L(\lambda, E) > 0 \quad \text{for all} \quad \lambda > \lambda_0(V, d, \sigma).$$

Suppose $T(\boldsymbol{\theta}) = \boldsymbol{\theta} + \boldsymbol{\omega}$ is simply a shift. Since (11.3) is a much weaker version of the large deviation theorems from Section 9, we thus get an independent proof of the Herman–Sorets–Spencer result [21],[34], and also the multifrequency version of it that was established in [4]. Our approach is different from that in [4] as it relies on the avalanche principle and the weak form of the large deviation theorem (11.3). The basic idea behind our argument is that Proposition 2.2 allows one to control the distances between various Lyapunov exponents, cf. Proposition 5.2. Therefore, as soon as one of them is sufficiently large, the positivity should follow. Throughout this section we assume that the potential V is nonconstant. For the following lemma we set $\lambda = 1$. This is no loss of generality, as one can replace V with λV (but only in this lemma). Hence we will suppress λ in our notation.

Lemma 11.1. *Suppose that* (11.3) *holds for all n with some choice of $\sigma > 0$. Then there exists a positive integer $\ell_0 = \ell_0(\sigma)$ such that if*

$$L_\ell > S(E)\ell^{-\sigma/4} \quad \text{and } L_\ell(E) - L_{2\ell}(E) < \frac{L_\ell(E)}{8} \tag{11.4}$$

for some $\ell \geq \ell_0$, then $L(E) > L_\ell(E)/2$.

Proof. Let $\ell_1 = \ell$ satisfy (11.4) and define inductively

$$\ell_{j+1} = \left[\ell_j^{1+\tau}\right] \quad \text{for } j = 1, 2, \ldots. \tag{11.5}$$

Here we have set $\tau = \frac{3}{8}\sigma$. For simplicity we shall drop E for the rest of this proof. Let C_1 be a large constant that will be determined below. We denote by A_j the statement

$$(A_j) \begin{cases} L_{\ell_j} - L_{2\ell_j} < L_{\ell_j}/8 \\ L_{\ell_j} > S\ell_j^{-\sigma/4}. \end{cases}$$

Notice that by hypothesis A_1 holds. Furthermore, B_j will denote the statement

$$(B_j) \begin{cases} |L_{\ell_{j+1}} - 2L_{2\ell_j} + L_{\ell_j}| < C_1\, S \frac{\ell_j}{\ell_{j+1}} \\ L_{\ell_{j+1}} - L_{2\ell_{j+1}} < C_1\, S \frac{\ell_j}{\ell_{j+1}}. \end{cases}$$

We shall show that $A_j, B_j \Longrightarrow A_{j+1}$ and that $A_j \Longrightarrow B_j$. Notice that this will give B_j and A_j for all j. The first implication is easy. Indeed,

$$\begin{aligned} L_{\ell_{j+1}} \quad &> \quad L_{\ell_j} - 2(L_{\ell_j} - L_{2\ell_j}) - C_1 S \frac{\ell_j}{\ell_{j+1}} > L_{\ell_j} - 2\frac{L_{\ell_j}}{8} - C_1 S \frac{\ell_j}{\ell_{j+1}} \\ &> \quad \frac{3}{4} S \ell_j^{-\sigma/4} - C_1 S \frac{\ell_j}{\ell_{j+1}} > S \ell_{j+1}^{-\sigma/4} \end{aligned} \tag{11.6}$$

where the latter inequality is an immediate consequence of (11.5) provided l_0 is large. Hence the second inequality from A_{j+1} holds. To obtain the first it suffices to prove the second inequality in

$$L_{\ell_{j+1}} > S\ell_{j+1}^{-\sigma/4} > 8C_1 S \frac{\ell_j}{\ell_{j+1}},$$

the first one being (11.6). ¿From (11.5) and $\tau > \sigma/4$ it is again evident that this will hold provided ℓ_0 is large. To show that $A_j \Longrightarrow B_j$ one uses Proposition 2.2. Fix some j and let $\ell_{j+1} = n\ell_j + r$ with $0 \le r < \ell_j$. In view of (11.3)

$$\mathrm{mes}\Big(\Big\{\boldsymbol{\theta} \in \mathbb{T}^d : |\log\|M_{\ell_j}(\boldsymbol{\theta}, E)\| - \ell_j L_{\ell_j}(E)| > S\delta\ell_j\Big\}\Big) \le C\delta^{-1}\ell_j^{-\sigma}. \tag{11.7}$$

Applying this with $S\delta = L_{\ell_j}/100$ shows that

$$\min_{0\le k<n} \|M_{\ell_j}(\boldsymbol{\theta} + k\ell_j\boldsymbol{\omega})\| \ge \exp\Big(3\ell_j L_{\ell_j}/4\Big) = \mu$$

for all $\boldsymbol{\theta} \in \mathcal{G}_1$ where

$$\mathrm{mes}(\mathbb{T}^d \setminus \mathcal{G}_1) < Cn\delta^{-1}\ell_j^{-\sigma} < C\ell_{j+1}\ell_j^{-1-3\sigma/4} = C\ell_j^{-\tau}.$$

One checks that $\mu > n$ provided ℓ_0 is large. Moreover, by (11.7) there exists a set $\mathcal{G}_2$ with $\mathrm{mes}(\mathbb{T}^d \setminus \mathcal{G}_2) < C\ell_j^{-\tau}$ so that

$$\max_{0\le k<n}\Big[\log\|M_{\ell_j}(\boldsymbol{\theta} + (k+1)\ell_j\boldsymbol{\omega})\| + \log\|M_{\ell_j}(\boldsymbol{\theta} + k\ell_j\boldsymbol{\omega})\| - \log\|M_{\ell_j}(\boldsymbol{\theta} + (k+1)\ell_j\boldsymbol{\omega})M_{\ell_j}(\boldsymbol{\theta} + k\ell_j\boldsymbol{\omega})\|\Big] \le$$

$$\le 2\ell_j(L_{\ell_j} + S\delta) - 2\ell_j(L_{2\ell_j} - S\delta) = 2\ell_j(L_{\ell_j} - L_{2\ell_j} + 2S\delta) < \frac{1}{3}\ell_j L_{\ell_j}. \tag{11.8}$$

Since this is clearly less than $\frac{1}{2}\log\mu$, (2.2) is satisfied. Therefore, Proposition 2.2 applies to all $\boldsymbol{\theta} \in \mathcal{G}_1 \cup \mathcal{G}_2$ and hence

$$\Big|\log\|M_{\ell_{j+1}}(\boldsymbol{\theta})\| - \sum_{k=0}^{n-1}\log\|M_{\ell_j}(\boldsymbol{\theta} + k\ell_j\boldsymbol{\omega})\| + \sum_{k=0}^{n-1}\log\|M_{\ell_j}(\boldsymbol{\theta} + (k+1)\ell_j\boldsymbol{\omega})M_{\ell_j}(\boldsymbol{\theta} + k\ell_j\boldsymbol{\omega})\|\Big|$$

$$\le \quad 2S\ell_j + C\frac{n}{\mu} \tag{11.9}$$

for all such $\boldsymbol{\theta}$. Integrating (11.9) over $\mathcal{G}_1 \cap \mathcal{G}_2$ yields

$$|L_{\ell_{j+1}} - 2L_{2\ell_j} + L_{\ell_j}| < C\Big[S\frac{\ell_j}{\ell_{j+1}} + \frac{n}{\mu\ell_{j+1}} + \ell_j^{-\tau}\Big] \le C_1 S\frac{\ell_j}{\ell_{j+1}} \tag{11.10}$$

with an appropriate choice of C_1. To complete the proof of $A_j \Longrightarrow B_j$, one simply applies the same reasoning to $M_{2\ell_{j+1}}$ and then subtracts the resulting inequality from (11.10). We skip the details.

Since B_j now holds for all j, one concludes that

$$\begin{aligned} L_{\ell_{j+1}} &> L_{\ell_j} - 2(L_{\ell_j} - L_{2\ell_j}) - C_1 S\frac{\ell_j}{\ell_{j+1}} > L_{\ell_j} - 3C_1 S\frac{\ell_j}{\ell_{j+1}} \\ &> \frac{3}{4}L_{\ell_1} - 3C_1\, S\Big[\frac{\ell_1}{\ell_2} + \frac{\ell_2}{\ell_3} + \ldots + \frac{\ell_j}{\ell_{j+1}}\Big] \end{aligned} \tag{11.11}$$

In view of (11.5) it is clear that

$$\sum_{j=1}^{\infty}\frac{\ell_j}{\ell_{j+1}} \asymp \ell_1^{-\tau}$$

provided ℓ_0 is large. Since $L_{\ell_1} > S\ell_1^{-\sigma/4}$ the lemma follows from (11.11). □

Remark 11.2. It is possible to prove a version of this lemma under a weaker condition than (11.3). More precisely, one can replace $n^{-\sigma}$ by $(\log n)^{-2-\delta}$, but we do not elaborate on this point.

In order to use this lemma to prove positivity of $L(\lambda, E)$ one needs to insure that the initial conditions (11.4) are satisfied. This will be accomplished by means of the following lemma. First we need to introduce some further notation. Let

$$f_n(\boldsymbol{\theta},\lambda,E) = \det \begin{bmatrix} \lambda V_1(\boldsymbol{\theta}) - E & 1 & 0 & 0 & . & . & . & . & 0 \\ 1 & \lambda V_2(\boldsymbol{\theta}) - E & 1 & 0 & 0 & . & . & . & 0 \\ 0 & 1 & \lambda V_3(\boldsymbol{\theta}) - E & 1 & 0 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & . & 1 & \lambda V_n(\boldsymbol{\theta}) - E \end{bmatrix} \tag{11.12}$$

where $V_j(\boldsymbol{\theta}) = V(T^j\boldsymbol{\theta})$. Recall the simple property

$$M_n(\boldsymbol{\theta},\lambda,E) = \begin{bmatrix} f_n(\boldsymbol{\theta},\lambda,E) & f_{n-1}(T\boldsymbol{\theta},\lambda,E) \\ f_{n-1}(\boldsymbol{\theta},\lambda,E) & f_{n-2}(T\boldsymbol{\theta},\lambda,E) \end{bmatrix}. \tag{11.13}$$

Finally, let

$$D_n(\boldsymbol{\theta},\lambda,E) \quad = \quad \mathrm{diag}(\lambda V_1(\boldsymbol{\theta}) - E, \ldots, \lambda V_n(\boldsymbol{\theta}) - E). \tag{11.14}$$

Lemma 11.3. *Let $0 < \epsilon < 1$. Then there is a constant C_v depending only on V such that*

$$(1-\epsilon)\log\lambda - C_v\epsilon^{-1} \quad \le \quad \frac{1}{n}\int_{\mathbb{T}^d} \log\|M_n(\boldsymbol{\theta},\lambda,E)\|\,d\boldsymbol{\theta} \le \log\lambda + C_v \tag{11.15}$$

for all $|E| \le 2\lambda\|V\|_\infty$, and $\lambda > \lambda_0(V,d,n,\epsilon)$.

Proof. The upper bound in (11.15) is simple. In fact,

$$\log\|M_n(\boldsymbol{\theta},\lambda,E)\| \le \sum_{j=1}^{n} \log\|A_j(\boldsymbol{\theta},\lambda,E)\| \le n\log\lambda + nC_v,$$

as claimed. Now fix some $0 < \epsilon < 1$ and any E as above. The matrix on the right–hand side of (11.12) can be written in the form $D_n + B_n$, where D_n is given by (11.14). Clearly, $\|B_n\| = 2$ and

$$\frac{1}{n}\log|\det D_n(\boldsymbol{\theta},\lambda,E)| = \log\lambda + \frac{1}{n}\sum_{j=1}^{n}\log|V_j(\boldsymbol{\theta}) - E/\lambda|.$$

By the Dunford–Schwarz maximal ergodic theorem

$$\mathrm{mes}\Big(\Big\{\boldsymbol{\theta}\in\mathbb{T}^d : \Big|\frac{1}{n}\sum_{j=1}^{n}\log|V_j(\boldsymbol{\theta}) - E/\lambda|\Big| > \rho\Big\}\Big) < \frac{C}{\rho}\int_{\mathbb{T}^d}\Big|\log|V(\boldsymbol{\theta}) - E/\lambda|\Big|\,d\boldsymbol{\theta}$$

with an absolute constant C. Since by the following lemma

$$\sup_{|E|\le 2\lambda\|V\|_\infty}\int_{\mathbb{T}^d}\Big|\log|V(\boldsymbol{\theta}) - E/\lambda|\Big|\,d\boldsymbol{\theta} \le C_v$$

one therefore has

$$\mathrm{mes}\Big(\Big\{\boldsymbol{\theta}\in\mathbb{T}^d : \frac{1}{n}\log|\det D_n(\boldsymbol{\theta},\lambda,E)| \ge \log\lambda - C_v\epsilon^{-1}\Big\}\Big) > 1 - \frac{\epsilon}{2} \tag{11.16}$$

for an appropriate choice of C_v. Clearly,

$$\|D_n(\boldsymbol{\theta},\lambda,E)^{-1}\| \quad \le \quad \lambda^{-1}\sup_{1\le j\le n}|V_j(\boldsymbol{\theta}) - E/\lambda|^{-1}. \tag{11.17}$$

By the following lemma there is $\delta = \delta(V, n, \epsilon) > 0$ such that

$$\operatorname{mes}\Big(\Big\{\boldsymbol{\theta} \in \mathbb{T}^d : \min_{1 \le j \le n} |V(T^j \boldsymbol{\theta}) - E/\lambda| < \delta\Big\}\Big) \le n \sup_{|E| \le 2\lambda \|V\|_\infty} \operatorname{mes}(\{\boldsymbol{\theta} \in \mathbb{T}^d : |V(\boldsymbol{\theta}) - E/\lambda| < \delta\}) < \frac{\epsilon}{2}.$$

Combining this with (11.17) yields

$$\operatorname{mes}(\{\boldsymbol{\theta} \in \mathbb{T}^d : \|D_n(\boldsymbol{\theta}, \lambda, E)^{-1}\| \ge \lambda^{-1}/\delta\}) < \frac{\epsilon}{2}$$

and thus

$$\operatorname{mes}\big(\{\boldsymbol{\theta} \in \mathbb{T}^d : 2\|D_n(\boldsymbol{\theta}, \lambda, E)^{-1} B_n\| < 1\}\big) > 1 - \frac{\epsilon}{2} \tag{11.18}$$

provided $\lambda > \lambda_0(V, n, \epsilon)$. Let $\mathcal{G} \subset \mathbb{T}^d$ be the intersection of the sets on the left–hand sides of (11.16) and (11.18). Then $\operatorname{mes}(\mathbb{T}^d \setminus \mathcal{G}) < \epsilon$, and for any $\boldsymbol{\theta} \in \mathcal{G}$,

$$\begin{aligned} f_n(\boldsymbol{\theta}, \lambda, E) &= \frac{1}{n} \log|\det D_n(\boldsymbol{\theta}, \lambda, E)| + \frac{1}{n} \log|\det(I + D_n(\boldsymbol{\theta}, \lambda, E)^{-1} B_n)| \\ &\ge \log \lambda - C_v \epsilon^{-1} - \log 2. \end{aligned}$$

Since $\|M_n\| \ge 1$ and $\|M_n(\boldsymbol{\theta})\| \ge |f_n(\boldsymbol{\theta}, \lambda, E)|$ (see (11.13)),

$$\frac{1}{n} \int_{\mathbb{T}^d} \log \|M_n(\boldsymbol{\theta}, \lambda, E)\| \, d\boldsymbol{\theta} \ge \frac{1}{n} \int_{\mathcal{G}} \log \|M_n(\boldsymbol{\theta}, \lambda, E)\| \, d\boldsymbol{\theta} \ge (1 - \epsilon) \log \lambda - C_v \epsilon^{-1}$$

provided $\lambda > \lambda_0(V, n, \epsilon)$, as claimed. □

The following technical lemma about real–analytic functions was used in the previous proof.

Lemma 11.4. *Suppose V is a nonconstant real–analytic function on $Q_0 = [-2, 2]^d$ with $\sup_{Q_0} |V| \le 1$. Then there exist $\epsilon = \epsilon(V, d) > 0$ and $C = C(V, d)$ so that*

$$\operatorname{mes}(\{(x_1, \ldots, x_d) \in [-1, 1]^d : |V(x_1, \ldots, x_d) - E| < t\}) \le C t^\epsilon$$

for all $-1 \le E \le 1$ and $0 < t < 1$.

Proof. It is not hard to derive this result from Theorem 8, part (B) in [29], see also Theorem 4 in that paper. Moreover, this statement is also contained in a forthcoming paper by A. Brudnyi. However, since Lemma 11.4 is much simpler than the results in [29], we give a short self–contained proof. We use the following fact about analytic functions of one complex variable, see Theorem 4, section 11.3 in [27]:

Let $f(z)$ be an analytic function in the disk $\{z : |z| \le 2e\}$ bounded by M and assume $|f(0)| = 1$. Then

$$\operatorname{mes}(\{z \in D(0, 1) : |f(z)| \le \lambda\}) \le C \exp\Big(2 \frac{\log \lambda}{\log M}\Big)$$

for any $\lambda > 0$. In fact, the set on the left–hand side can be covered by a family of disks $\{D_j\}_j$ so that

$$\sum_j \operatorname{diam}(D_j) \quad \le \quad C \exp\Big(\frac{\log \lambda}{\log M}\Big). \tag{11.19}$$

To apply this fact, consider a covering

$$[-1, 1]^d \subset \bigcup_{\ell=1}^m B(\boldsymbol{p}_\ell, r_\ell)$$

where $m = m(d)$ such that $|\nabla V(\boldsymbol{p}_\ell)| > g_0 = g_0(V) > 0$ and $r_\ell < 1/10$ for every ℓ. Suppose that $|\frac{\partial}{\partial x_1} V(\boldsymbol{p}_1)| > g_0/d$ and define $f_{\boldsymbol{u}}(z) = \frac{\partial}{\partial x_1} V(\boldsymbol{p}_1 + z\boldsymbol{u})$ where $\boldsymbol{u}$ is a unit vector in $\mathbb{R}^d$ and $z \in \mathbb{C}$. In view of (11.19), there is some $\epsilon = \epsilon(V, d)$ such that

$$\operatorname{mes}(\{x \in [-r_1, r_1] : |f_{\boldsymbol{u}}(x)| \le t^\delta\}) \le C t^{\epsilon\delta} \tag{11.20}$$

for any $\delta > 0$ and $0 < t < 1$. Integrating this over $\boldsymbol{u}$ and summing over $\ell = 1, \dots, m$ one obtains

$$\mathrm{mes}(\{(x_1, \dots, x_d) \in [-1,1]^d : |\nabla V(x_1, \dots, x_d)| < t^\delta\}) \le C t^{\delta\epsilon}.$$

Suppose $|\nabla V(\boldsymbol{p})| > t^\delta$ for some point $\boldsymbol{p} \in [-1,1]^d$. Then clearly

$$|\nabla V(\boldsymbol{p}')| > \frac{1}{2} t^\delta \quad \text{for all} \quad |\boldsymbol{p}' - \boldsymbol{p}| < c t^\delta.$$

One therefore concludes that

$$\begin{aligned} \mathrm{mes}(\{[-1,1]^d : |V - E| < t\}) &\le \mathrm{mes}(\{[-1,1]^d : |V - E| < t, |\nabla V| > t^\delta\}) + \mathrm{mes}(\{[-1,1]^d : |\nabla V| \le t^\delta\}) \\ &\le C t^{-d\delta} t^{1-\delta} + C t^{\epsilon\delta}. \end{aligned}$$

Choosing $\delta = \frac{1}{4d}$, say, implies the lemma. □

The following proposition is the main result of this section.

Proposition 11.5. *Suppose* (11.3) *holds. With* $L(\lambda, E)$ *as defined above,*

$$\inf_E L(\lambda, E) \quad > \quad \frac{1}{4} \log \lambda \tag{11.21}$$

provided $\lambda > \lambda_0(V, d)$. *In particular,* (11.21) *holds in case of an ergodic shift on* $\mathbb{T}^d$.

Proof. Consider first the case $|E| < 2\lambda \|V\|_\infty$. Clearly,

$$S(\lambda, E) = \log(C\lambda \|V\|_\infty + 1)$$

satisfies the requirement (11.2) for all n. To obtain the first condition in (11.4) one needs to insure that (setting $\epsilon = \frac{1}{2}$ in (11.15))

$$\frac{1}{2} \log \lambda - C_v \quad > \quad 40 \log(C\lambda \|V\|_\infty + 1) \ell^{-\sigma/4}. \tag{11.22}$$

Fixing some $\ell > \max(\ell_0, 100^{\frac{4}{\sigma}})$ with ℓ_0 as in Lemma 11.1, and taking $\lambda > \lambda_0(V, d, \ell)$ sufficiently large yields (11.22). To obtain the second condition in (11.4), one applies (11.15) with $\epsilon = 1/32$, say. For large λ the proposition now follows from Lemma 11.1 for energies as above.

Now suppose that $|E| \ge 2\lambda \|V\|_\infty$. Then

$$|\lambda V(T^j \boldsymbol{\theta}) - E| > \lambda \|V\|_\infty.$$

Let D_n be as in (11.14). Then $|\det D_n(\boldsymbol{\theta}, \lambda, E)| \ge (\lambda \|V\|_\infty)^n$ and $\|D_n(\boldsymbol{\theta}, \lambda, E)^{-1}\| \le (\lambda \|V\|_\infty)^{-1} \le \frac{1}{4}$ provided $\lambda \ge 4 \|V\|_\infty^{-1}$. Writing the matrix on the right–hand side of (11.12) as $D_n + B_n$, leads to

$$f_n(\boldsymbol{\theta}, \lambda, E) = \det(D_n) \det(I + D_n^{-1} B_n).$$

One therefore has (since $\|B_n\| \le 2$)

$$\inf_{\boldsymbol{\theta} \in \mathbb{T}^d} \frac{1}{n} \log |f_n(\boldsymbol{\theta}, \lambda, E)| \ge \frac{1}{n} \log[(\lambda \|V\|_\infty)^n 2^{-n}] = \log \lambda + \log \|V\|_\infty - \log 2.$$

Hence

$$\frac{1}{n} \int_{\mathbb{T}^d} \log \|M_n(\boldsymbol{\theta}, \lambda, E)\| \, d\boldsymbol{\theta} \ge \log \lambda - C_v$$

which implies that

$$\inf_{|E| \ge 2\lambda \|V\|_\infty} L(\lambda, E) > \log \lambda - C_v$$

and the proposition follows. □

References

1. Anderson, P. *Absence of diffusion in certain random lattices.* Phys. Rev. **109** (1958), 1492–1501.
2. Aizenman, M., Molchanov, S. *Localization at large disorder and at extreme energies: an elementary derivation.* Comm. Math. Phys. 157 (1993), 245–278.
3. Avron, J., van Mouche, P., Simon, B. *On the measure of the spectrum for the almost Mathieu operator.* Comm. Math. Phys. 132 (1990), 103–118.
4. Bourgain, J., Goldstein, M. *On nonperturbative localization with quasiperiodic potential.* preprint 1999.
5. Campanino, M., Klein, A. *A supersymmetric transfer matrix and differentiability of the density of states in the one–dimensional Anderson model.* Comm. Math. Phys. 104 (1986), 227–241.
6. Carmona, R., Lacroix, J. *Spectral theory of random Schrödinger operators.* Birkhäuser, Boston 1990.
7. Constantinescu, F., Fröhlich, J., Spencer, T. *Analyticity of density of states and replica method for random Schrodinger operators on a lattice.* J. Stat. Phys.34 (1984), 371–396.
8. Cycon, H. L., Froese, R. G., Kirsch, W., Simon, B. *Schrödinger operators.* Springer Verlag, Berlin, Heidelberg, New York 1987.
9. Dinaburg, E. I., Sinai, Y. G. *The one dimensional Schrödinger equation with quasiperiodic potential.* Funkt. Anal. i. Priloz. **9** (1975), 8–21.
10. Dunford, N., Schwartz, J. *Linear operators I, General theory.* Interscience Publishers, Inc., New York, 1958.
11. Eliasson, H. *Floquet solutions for the 1–dimensional quasiperiodic Schrödinger equation.* Comm. Math. Phys. 146 (1992), 447–482.
12. Figotin, A., Pastur, L. *Spectra of random and almost–periodic operators.* Grundlehren der mathematischen Wissenschaften 297, Springer 1992.
13. Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T. *Constructive proof of localization in the Anderson tight binding model.* Comm. Math. Phys. 101 (1985), 21–46.
14. Fröhlich, J., Spencer, T. *Absence of diffusion in the Anderson tight binding model for large disorder or low energy.* Commun. Math. Phys. **88** (1983), 151–189.
15. Fröhlich, J., Spencer, T. *A rigorous approach to Anderson localization.* Phys. Rep. **103** (1984), no. 1–4, 9–25.
16. Fröhlich, J., Spencer, T., Wittwer, P. *Localization for a class of one dimensional quasi-periodic Schrödinger operators.* Commun. Math. Phys. **132** (1990), 5–25.
17. Fürstenberg, H. *Noncommuting random products.* Trans. AMS 108 (1963), 377–428.
18. Goldsheid, I. Ya., Molchanov, S. A., Pastur, L. A. *A pure point spectrum of the stochastic one-dimensional Schrödinger equation.* Funkt. Anal. Appl. **11** (1977), 1–10.
19. Gordon, A., Jitomirskaya, S., Last, Y., Simon, B. *Duality and singular continuous spectrum in the almost Mathieu equation.* Acta Math. 178 (1997), 169–183.
20. Harper, P. G. *Single band motion of conduction electrons in a uniform magnetic field.* Proc. Phys. Soc. London **A68** (1955), 874–892.
21. Herman, M. *Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le charactère local d'un theoreme d'Arnold et de Moser sur le tore de dimension* 2. Comment. Math. Helv. 58 (1983), no. 3, 453–502.
22. Jitomirskaya, S. *Metal–insulator transition for the almost Mathieu operator.* Annals of Math. 150, no. 3 (1999).
23. Khinchin, A. Ya. *Continued fractions.* Dover 1964.
24. Koosis, P. *The logarithmic integral II.* Cambridge Studies in Advanced Mathematics, 21. Cambridge University Press, Cambridge, 1992.
25. Last, Y. *A relation between a.c. spectrum of ergodic Jacobi matrices and the spectra of periodic approximants.* Comm. Math. Phys. 151 (1993), 183–192.
26. Last, Y. *Zero measure spectrum for the almost Mathieu operator.* Comm. Math. Phys. 164 (1994), 421–432.
27. Levin, B. Ya. *Lectures on entire functions.* Transl. of Math. Monographs, vol. 150. AMS, Providence, RI, 1996.
28. Oseledec, V. *The multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems.* Trudy Mosc. Mat. Obz. 19 (1968), 197–231.
29. Phong, D. H., Stein, E. M., Sturm, J. A. *On the growth and stability of real–analytic functions.* Amer. J. Math. 121 (1999), no. 3, 519–554.
30. Shnol, I. *On the behavior of the Schrödinger equation.* Mat. Sb., 273-286 (1957), Russian.
31. Simon, B. *Spectrum and continuum eigenfunctions of Schrödinger Operators.* J. Funct. Anal. 42 (1981), 66–83.
32. Simon, B., Taylor, M. *Harmonic analysis on* $Sl(2,\mathbb{R})$ *and smoothness of density of states in the one–dimensional Anderson model.* Comm. Math. Phys. 125 (1985), 113–125.
33. Sinai, Y. G. *Anderson localization for one-dimensional difference Schrödinger operator with quasi-periodic potential.* J. Stat. Phys. **46** (1987), 861–909.
34. Sorets, E., Spencer, T. *Positive Laypunov exponents for Schrödinger operators with quasi–periodic potentials.* Comm. Math. Phys. 142 (1991), no. 3, 543–566.
35. Stein, E. *Harmonic analysis.* Princeton Mathematical Series 43, Princeton University Press, Princeton, 1993.
36. Thouless, D. *Scaling for the discrete Mathieu equation.* Comm. Math. Phys. 127 (1990), 187–193.

37. Wegner, F. *Bounds on the density of states in disordered systems.* Z. Phys. **B44** (1981), 9–15.

Institute for Advanced Study, Olden Lane, Princeton, N.J. 08540, U.S.A.
Permanent Address: Dept. of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 1A1
E-mail address: mgold@math.ias.edu

Department of Mathematics, Princeton University, Fine Hall, Princeton N.J. 08544, U.S.A.
E-mail address: schlag@math.princeton.edu